ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Changing Value Relevance

The European Perspective

ALEXANDER ZAPANTOULIS

ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Changing Value Relevance

The European Perspective

ALEXANDER ZAPANTOULIS

Supervisors: Dr. C.D. Knoops
Dr. D. Veenman

In partial fulfilment of the requirements
for the degree of
Master of Science in Economics and Business

8 August 2012

Abstract

This paper investigates value relevance in France, Germany, the Netherlands and the United Kingdom over the period 1991-2011. In contrast to most previous researches, value relevance is not only measured by coefficients of determination, but also by abnormal pricing errors (Gu, 2007). Three main findings are reported. First, value relevance appears to be stable in the sample countries over the period 1991-2004, after which there is a significant increase in value relevance that appears to be caused by the mandatory adoption of IFRS in 2005. Second, firms experiencing high amounts of business change have significantly lower value relevance than firms experiencing little business change. Finally, there does not appear to be a difference in the value relevance of financial statements in the Netherlands and the United Kingdom, countries classified as having a market-oriented financial system, compared to that of financial statements in Germany and France, countries classified as having a bank-oriented financial system.

Contents

1. Introduction and structure

Value relevance has been and still is an important topic in accounting research in the United States. Value relevance, as the name suggests, concerns the relevance of financial statement information to users of those financial statements. A great deal has been written on changes in value relevance over time in the U.S. Previous researchers have reported a decline in value relevance in the U.S. over the period 1953-1996. This decline in value relevance is a concern to policy makers, as it indicates that accounting information no longer fully serves the purpose it was designed to meet..

In all of the research on declining value relevance, Europe has however been overlooked. To the author's best knowledge, no substantive paper exists on temporal changes in value relevance in Europe, despite the fact that European countries have seen some very interesting developments in their accounting systems, e.g. through the mandatory introduction of IFRS in 2005. This study attempts to fill part of this gap in value relevance research and will shed some light on changes in value relevance in Europe.

The key research question of this study is:

"Have there been any changes in value relevance in France, Germany, the Netherlands and the United Kingdom over time and what are plausible explanations for these possible changes?"

The structure of this thesis is as follows: the key research question is introduced in this chapter. Next, in chapter two, a background on the definition of value relevance and the theoretical framework of this research are presented. In chapter three, a number of metrics to measure value relevance over time are discussed, as well as the drawbacks and advantages of these different measures. Chapter four presents empirical results of previous researchers on changes in value relevance over time and on a number of explanatory variables that might have driven these changes. Since a cross-country sample is used in this research, chapter five elaborates on international differences in value relevance. In chapter six the different hypotheses that are tested in this research are presented together with the research methodology. Chapter seven discusses the sample and sample selection procedures, and in chapter eight results are presented. In chapter nine tests of data robustness are performed and the results are further analysed. Chapter 10 presents the conclusion of this research.

2. Value relevance and the usefulness of financial statement information

2.1 Value Relevance

One of the primary objectives of financial reporting is to assist in the decision making process of users (Scott, 2006). In this sense important users are investors. If financial statements lead to improved decision making by investors, they can be deemed useful. Two perspectives exist about the usefulness of financial statements, namely the information perspective and the measurement perspective (Scott, 2006). The information perspective assumes that the usefulness of financial information is independent of the form of disclosure (Scott, 2006). To justify this, the information perspective relies heavily on the efficient market hypothesis. To test the usefulness of financial statements and other information, research based on the information perspective often uses event studies to investigate whether the release of this information leads to changes in security prices. Another important note about the information perspective is that it justifies the use of historic cost based accounting, provided that there is full disclosure (Scott, 2006).

The measurement perspective, on the other hand, relies on the efficient market hypothesis. Instead of historical costs, the measurement perspective emphasizes the use of fair value to assist users in the decision making process. This can also be seen in the definition of the measurement perspective as put forward by Scott (2006, 157): *"The measurement perspective on decision usefulness is an approach to financial reporting under which accountants undertake a responsibility to incorporate fair values into the financial statements proper, provided that this can be done with reasonable reliability, thereby recognizing an increased obligation to assist investors to predict firm value."* An important prerequisite for the use of fair value accounting is that fair values can be estimated with reasonable reliability, as stated in the definition above. For this reason, it is unlikely that fair value based accounting, and thus the measurement perspective, will replace historical cost based accounting completely, but there seems to be a shift in that direction (Scott, 2006).

Scott (2006) provides three arguments which could be made in favour of the measurement perspective and thus fair value based accounting. These are (i) inefficient security markets, (ii) the Ohlson clean surplus relation and (iii) practical considerations. First of all, if security markets are not efficient, investors might need additional information to properly estimate firm value. This information can be obtained by fair value measures of assets and liabilities, provided that these estimates are reliable. Secondly, the Ohlson clean surplus relation states that firm value can be expressed in terms of accounting variables. The difference between the book value and the market value of a firm is reduced if accounting standards are "unbiased" (Scott, 2006), i.e. measure fair value. This notion is consistent with the measurement perspective, which also assumes accounting variables should indicate firm value. Because of the importance of the Ohlson clean surplus relation and the related accounting based valuation model in value relevance studies, these will be discussed more extensively in the next section. Finally, with practical considerations, Scott (2006) refers to recent lawsuits against auditors in which the audited firms had seriously overstated their net assets. This is especially the case for financially distressed firms, which should not be valued at historical costs, but at liquidation value. In these circumstances an accounting framework based on fair value might be more appropriate.

An important concept in the measurement perspective is value relevance. This is defined by Scott (2006, 171) as: *"The extent to which financial statements information has a material effect on share returns and prices."* Although this definition makes intuitive sense, it is hard to empirically measure value relevance and especially changes in value relevance over time. An extensive discussion of interpretations of value relevance is provided by Francis and Schipper (1999). The following section shortly summarizes their discussion.

Interpretation 1 of value relevance provided by Francis and Schipper (1999, 325) is that "*financial statement information leads stock prices by capturing intrinsic share values toward which stock prices drift."* However, to empirically test this interpretation, strong assumptions on both share prices and financial information are required. For these reasons Francis and Schipper (1999) reject this interpretation. Interpretation 2 of value relevance is that "*financial statement information is value relevant if it contains the variables used in a valuation model or assists in predicting those variables*" (Francis and Schipper, 1999, 325). This interpretation is not widely applied, but it can be found in Kim and Kross (2005).

The other interpretations of value relevance are concerned with statistical associations between financial statement information and share prices or returns. First of all, Interpretation 3 questions whether investors actually use financial statement information to set prices (Francis and Schipper, 1999). Again there are considerable difficulties to measure this in an empirical setting. Note for example that this requires that financial statement information is the earliest source of information, which is hard to verify. Alternatively, Interpretation 4 (Francis and Schipper, 1999) merely questions whether there exists a statistical association between financial information and share prices or returns. In contrast to Interpretation 3, this does not require that financial statement information is the earliest source of information. Also, this interpretation of value relevance is somewhat weaker than the one defined by Scott (2006). Under Interpretation 4, financial information can be deemed relevant if it captures and confirms earlier sources of information. It is not required that financial statement information has an impact on either share prices or returns.

Interpretation 4 is widely used; see for example Collins et al. (1997), Francis and Schipper (1999), Chang (1999), Lev and Zarowin (1999) and Gu (2007). It should be noted that under this interpretation, there still exist several ways of measuring value relevance. Before these will be explained in more detail in chapter three, the theoretical foundation of one of these measures is discussed, namely the Ohlson clean surplus relation.

2.2 The Ohlson clean surplus relation and the accounting based valuation model

Ohlson (1995, 661-662) tries to articulate *"a cohesive theory of a firm's value that relies on the clean surplus relation to identify a distinct role for each of the three variables, earnings, book value, and dividends."* To do so Ohlson (1995) makes three basic assumptions; i) regular owners' equity accounting applies, ii) shares are valued by the present value of future dividends and iii) the stochastic process of abnormal earnings can be described by a linear model. These assumptions require some additional elaboration.

The first assumption states that regular owner's equity accounting applies (Ohlson, 1995). This entails the following two concepts. First of all, it means that the clean surplus relation applies and secondly that dividends influence the book value of equity, but do not influence current earnings. These assumptions can also be stated mathematically:

$$BV_{t-1} = BV_t + D_t - X_t \tag{2.1}$$

and

$$\frac{\partial BV_t}{\partial D_t} = -1 \tag{2.2}$$

$$\frac{\partial X_t}{\partial D_t} = 0 \tag{2.3}$$

The following notation applies: BV_t is the net book value of equity for the period t, X_t are the earnings for the period t and D_t are the net dividends paid at period t. Equation (2.1) states the clean surplus relation, which means that the change in book value of equity over a period is equal to earnings minus net dividends paid that period. In this equation dividends can be both negative and positive. Negative dividends occur when capital contributions in a year exceed paid dividends. Equation (2.2) and (2.3) state respectively, that dividends decrease net book value of equity on a one-to-one basis and that current earnings are not affected by dividends.

The second assumption, which states that the value of a firm is equal to the net present value of future (expected) dividends, can be stated as:

$$P_t \equiv \sum_{\tau=1}^{\infty} \frac{E_t[\widetilde{D}_{t+\tau}]}{(1+R_f)^{\tau}} \tag{2.4}$$

Here is P_t the market value of the firm at date t, R_f the risk-free rate and $E_t[.]$ is used to indicate that expected values are applied, conditional on the information available at date t.[1]

Assumption (2.4) is straightforward and is widely applied in finance. It should be noted that it is a very simplified version of the dividend-discount model. The simplification comes from additional assumptions made by Ohlson (1995), namely that risk neutrality and homogeneous beliefs apply. Furthermore it is assumed that the term structure of interest rates is flat and non-stochastic. Because of these assumptions the risk-free rate is used in the model, instead of the cost of equity capital. These additional assumptions can be relaxed to allow a more appropriate reflection of reality, but for matters of simplicity, these assumptions will also be applied in this discussion.

Before the final assumption of Ohlson (1995) will be discussed, abnormal earnings (X_t^a) have to be defined. Abnormal earnings are earnings minus a charge for the use of capital. The charge for the use of capital is described as the cost of equity capital, which is R_f in our risk-neutral world, multiplied by the book value of equity at the beginning of a period (Ohlson, 1995). This means:

$$X_t^a \equiv X_t - R_f BV_{t-1} \tag{2.5}$$

By combining (2.1), (2.4) and (2.5), Ohlson (1995) derives one of the main results of his paper, namely:

[1] The tilde above abnormal dividends is used to indicate that this is a random variable at time t.

$$P_t = BV_t + \sum_{\tau=1}^{\infty} \frac{E_t[\tilde{X}^a_{t+\tau}]}{(1+R_f)^\tau} \tag{2.6}$$

Relation (2.6) states that the market value of a firm is equal to the book value of equity plus the present value of expected abnormal earnings. The latter is also called "goodwill" (Feltham and Ohlson, 1995). The valuation model of (2.6) is called the accounting based valuation model.

The final assumption of Ohlson is that the stochastic process of abnormal earnings can be described by a linear model. The main points of his discussion are:

$$\tilde{X}^a_{t+1} \equiv \theta X^a_t + v_t + \tilde{\varepsilon}_{1,t+1} \tag{2.7}$$

$$\tilde{v}_{t+1} \equiv \gamma v_t + \tilde{\varepsilon}_{2,t+1} \tag{2.8}$$

Assumption (2.7) describes that next year abnormal earnings are dependent on current year's abnormal earnings plus information other than abnormal earnings, i.e. v_t, and a disturbance term, $\varepsilon_{1,t+1}$, which is zero-mean and unpredictable.[2] Also, the "other information" in year $t+1$ is dependent on the previous year's other information and again a disturbance term, $\varepsilon_{2,t+1}$, as described by (2.8).[3] The parameters θ and γ are assumed to be constant and are restricted to be non-negative and smaller than one. These dynamics are also described by Feltham and Ohlson (1995), who call it earnings persistence. How much earnings persist depends on the scale of parameter θ, which in turn is dependent on a firm's accounting principles and economic environment (Ohlson, 1995).

Because of the specification of (2.7) and (2.8), Ohlson (1995, 682) is able to simplify his accounting based valuation model (2.6) to a linear model. This is stated as:

$$P_t = BV_t + \alpha_1 X^a_t + \alpha_2 v_t \tag{2.9}$$

[2] Meaning that $E_t[\varepsilon_{1,t+\tau}] = 0$ for $\tau \geq 1$

[3] This disturbance term $\varepsilon_{2,t+1}$ is similar to $\varepsilon_{1,t+1}$ in that it also is zero-mean and unpredictable, i.e. $E_t[\varepsilon_{2,t+\tau}] = 0$ for ≥ 1 . Note that no assumptions are made about the underlying distributions, variances or covariances of the disturbance terms.

in which:

$$\alpha_1 = \frac{\theta}{\left(R_f - \theta\right)} \geq 0 \tag{2.10}$$

$$\alpha_2 = \frac{R_f}{\left(R_f - \theta\right)}\left(R_f - \gamma\right) > 0 \tag{2.11}$$

$$R_f \neq \theta \tag{2.12}$$

This linear specification is especially useful for empirical testing as will be discussed in the following chapter. Equation (2.9) states again that the market value of a firm is dependent on a firm's book value of equity, and its abnormal earnings, but also on information not captured in this year's abnormal earnings. It differs from (2.6) because only current year's information is necessary to value a firm. However this specification of a firm's value is more restrictive than (2.6), because assumptions (2.7) and (2.8) have to hold.

3. Measures of value relevance

3.1 Introduction

To investigate changes in value relevance over time, a number of methodologies have been developed. As was discussed in chapter two, value relevance can be interpreted in several ways. The most common used interpretation is that of association amongst financial information (i.e. accounting variables) and share prices or returns (Interpretation 4). This specific interpretation of value relevance still allows for different operationalisations. Some questions that rise are: how does one choose the variables of interest, how does one measure the strength of the relations between these variables and how can one assess changes in value relevance over time? The answers to these questions will be discussed in detail in the following sections.

It should be noted that this chapter provides an extensive overview of measures of value relevance, but is not exhaustive. A number of alternative methodologies to measure changes in value relevance also exist, which are either not directly related to the methodology and objective of this research or are concerned with specific issues of value relevance.[4] A short discussion of the results of these studies as well as a comparison with the methodologies described in this chapter will instead be presented in chapter four.

The remainder of this chapter is organized as follows. Sections 3.2 to 3.5 are concerned with operationalisations of value relevance under Interpretation 4. In section 3.2 the most common used metric is discussed; the coefficient of determination (R^2). This metric has been heavily scrutinized over the past years and a thorough review of its shortcomings as well as possible remedies is presented in section 3.3. In section 3.4 and 3.5 other measures of value relevance are introduced, namely the coefficients metric and portfolio metric. These metrics can still be classified under Interpretation 4, but they are fundamentally different from those discussed in section 3.2 and 3.3, because they do not rely on the variance of residuals. Section 3.6 then briefly introduces an alternative operationalisation of value relevance under Interpretation 2, although it will become clear that this does not completely alter the applied methodology. Section 3.7 then concludes.

[4] Examples of studies which are not directly related to the methodology and objective of this research are those of Landsman and Maydew (2002), Francis et al. (2002) and Collins et al. (2009). These studies focus on the information content, or equivalently the informativeness, of earnings announcements over time. Informativeness in this context is defined as the capital market reaction following earnings announcements, which can be measured by abnormal trading volume, abnormal returns or abnormal return volatility around earnings announcements (Collins et al, 2009). Strictly speaking these studies are also not referred to as value relevance studies, but informativeness studies. Studies focusing on particular issues of value relevance are, amongst others, Ryan and Zarowin (2003) and Dontoh et al. (2007).

3.2 The coefficient of determination metric

One of the first studies to investigate changes in value relevance over time was performed by Collins et al. (1997). Their study was inspired by claims that value relevance has declined over the years. To investigate this, Collins et al. (1997) estimate the value relevance of book values and earnings combined and the value relevance of earnings and book values on a standalone basis for each year in the period 1953-1993. In their study, Collins et al. (1997) clearly apply Interpretation 4; value relevance is measured by the association between accounting variables and share prices. The question then is which relations should be estimated. As a theoretical basis for their estimated relations, Collins et al. (1997) use the Ohlson (1995) framework discussed in the previous chapter.

The adopted version of Ohlson's model, also called the accounting based valuation model in subsequent chapters, used by Collins et al. (1997) is expressed as:

$$P_{i,t} = \alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}BV_{i,t} + \varepsilon_{i,t} \tag{3.1}$$

In this expression $P_{i,t}$ is the price of a share of firm i, three months after fiscal year end t, $X_{i,t}$ are the earnings per share of firm i in year t, $BV_{i,t}$ is the book value of equity per share of firm i at the end of year t and $\varepsilon_{i,t}$ represents all other value relevant information of firm i for year t not included in either $BV_{i,t}$ or $X_{i,t}$.

Collins et al. (1997) do not explicitly state how this expression relates to the Ohlson (1995) framework, but it is insightful to make their assumptions explicit. First of all, the basis of (3.1) is provided by expression (2.9). For (2.9) to hold, assumptions (2.7) and (2.8), which describe the stochastic process of abnormal earnings, also have to hold. It is questionable whether this is true in an empirical setting.

Another important notion is that by estimating (3.1) it is assumed that parameters $\alpha_{1,t}$ and $\alpha_{2,t}$ are constant among all firms. If parameter $\alpha_{1,t}$ is related to (2.10), this implies that earnings persistence, θ, also has to be constant among all firms. Furthermore, in the discussion of the accounting based valuation model, risk neutrality and homogeneous beliefs were assumed. In an empirical setting we face unique risk-adjusted discount rates for each firm. Incorporating this in the accounting based valuation model would affect equation (2.10), making it unique for each firm as well. This is not incorporated in (3.1), which imposes that the cost of capital is constant among all firms (Chang, 1999). Finally, in the discussion of the Ohlson (1995) framework, no assumption was made about earnings growth. With (3.1), Collins et al. (1997) implicitly assume that expected earnings growth rates are constant among all firms (Chang, 1999).

Since (3.1) is a very restricted version of the accounting based valuation model, it is not clear whether this makes it unsuitable for the purpose of value relevance studies. As a first indicator, Collins et al. (1997) note that allowing for firm specific discount rates does not significantly improve the explanatory power of the model. Chang (1999) directly compares inferences about value relevance from (3.1) and a more extensive accounting based valuation model. This extensive model does allow for industry specific earnings persistence, firms specific growth rates and firm specific costs of capital. Chang (1999) finds that this extended model is superior, but it does not alter his inferences about value relevance. This evidence might not be decisive, but it provides some faith in the suitability of (3.1).

Estimating relation (3.1) for each year t is only the first step. The variable of interest is the strength of this relation. A common used metric for this is the coefficient of determination.[5] The coefficient of determination measures the linear association between dependent and independent variables (Greene, 2008). This makes it a logical choice as a metric for value relevance in this linear system. As an additional note, Collins et al. (1997) are interested not only in the value relevance of book values and earnings combined, but also of the two components individually. To measure this, the following relations are also estimated:

$$P_{i,t} = \beta_{0,t} + \beta_{1,t} X_{i,t} + \epsilon_{i,t} \tag{3.2}$$

$$P_{i,t} = \gamma_{0,t} + \gamma_{2,t} BV_{i,t} + \varphi_{i,t} \tag{3.3}$$

Also, define the estimated coefficients of determination for year t from (3.1)-(3.3) as $\hat{R}^2_{1,t}$, $\hat{R}^2_{2,t}$ and $\hat{R}^2_{3,t}$.[6] To determine the explanatory power of either book values or earnings on a standalone basis Collins et al. (1997) use:

$$\hat{R}^2_{BV,t} \equiv \hat{R}^2_{1,t} - \hat{R}^2_{2,t} \tag{3.4}$$

$$\hat{R}^2_{E,t} \equiv \hat{R}^2_{1,t} - \hat{R}^2_{3,t} \tag{3.5}$$

[5] Either the “standard” coefficient of determination or the “adjusted” coefficient of determination; see Greene (2008) for the difference between these concepts.
[6] The hats above variables indicate sample estimates.

$$\hat{R}^2_{C,t} \equiv \hat{R}^2_{1,t} - \hat{R}^2_{E,t} - \hat{R}^2_{BV,t} \tag{3.6}$$

In these expressions $\hat{R}^2_{BV,t}$ and $\hat{R}^2_{E,t}$ are the estimated explanatory powers for book values and earnings for year t and $\hat{R}^2_{C,t}$ is the estimated explanatory power common to book values and earnings. This decomposition technique was first derived by Theil (1971).

The final step is to determine if there have been any changes in value relevance over time. This is straightforward once a measure of value relevance is obtained. A simple approach is to regress the measure of value relevance on a time variable. This is exactly the approach followed by Collins et al. (1997), i.e.:

$$\hat{R}^2_t = \delta_0 + \delta_1 Time_t + \vartheta_t \tag{3.7}$$

The variable $Time_t = 1, \ldots, 41$ for Collins et al. (1997) since their sample represents the years 1953-1993. This regression is done for $\hat{R}^2_{1,t}$, $\hat{R}^2_{BV,t}$ and $\hat{R}^2_{E,t}$. The variable of interest here is $\hat{\delta}_1$. If $\hat{\delta}_1$ is significantly smaller than zero, this would indicate that value relevance has decreased over time. Conversely, a sufficiently large positive $\hat{\delta}_1$ would indicate an increase of value relevance over time.

This specific methodology, across sample comparison of coefficients of determination, has been widely applied in the literature. It can also be found in for example Francis and Schipper (1999), Lev and Zarowin (1999), Ely and Waymire (1999), Brown et al. (1999) and Kim and Kross (2005). Furthermore, these studies do no confine themselves to estimating (3.1); alternative relations have also been estimated. Lev and Zarowin (1999) for example estimate the relation between returns, cash flows and accruals as well as the relation between returns, earnings and earnings changes. This last relation has been studied for quite some time, see Lev (1989). Although these are valid tests of value relevance under interpretation 4, it should be noted that some of these relations lack the rigorous theoretical underpinning (3.1) has.

3.3 Comments on the use of coefficients of determination

3.3.1 Introduction

One of the fundamental assumptions made in the above discussed methodology is that coefficients of determination can be compared across samples. This might seem reasonable at first sight, but a word of caution can be found in a standard econometrical textbook such as Greene (2008, 38): *"Little can be said about the relative quality of fits of regression lines in different contexts or in different data sets even if supposedly generated by the same data generating mechanism."* In this section a review is provided of several issues concerning the use of coefficients of determination as a measure of value relevance. Also, alternative measures proposed to levy these issues are presented. First, in section 3.3.2 scale effects are extensively discussed as are possible remedies. In section 3.3.3 a general issue related to across sample comparisons of coefficients of determination is reviewed. Section 3.3.4 examines the variance of the logarithm metric proposed by Chang (1999) and section 3.3.5 discusses the closely related residual dispersion metric proposed by Gu (2007). Finally, section 3.3.8 concludes.

3.3.2 Scale effects

An important issue in any cross-sectional regression using accounting information and share prices are "scale effects". These effects have frequently been studied; see for example Barth and Kallapur (1996), Brown et al. (1999) and Christie (1987). Scale effects will be discussed in detail since these do not only influence estimates for coefficients of determination, but also regression coefficient estimates. The latter are of importance in an alternative measure of value relevance.

To see where scale effects come from, among other reasons, consider the following example (Brown et al., 1999). A researcher wants to investigate value relevance by regressing share prices on earnings per share. Assume that all firms are exactly identical and that no relationship exists between share prices and earnings per share. Subsequently this regression should yield $R^2 = 0$. In a subsequent period, some of the firms have undertaken share splits while others have not. A similar regression analysis would then yield $R^2 > 0$ even though the underlying economic relation has not changed. This difference in scale among firms and the resulting consequences for regression analysis are referred to as scale effects (Brown et al., 1999 and Barth and Kallapur, 1996). The scale effect is caused by a variety of things, such as dividend pay-out policy, stock splits and differences in performance. The problem is that financial data free of scale effects is often unobservable.

The influence of scale effects on regression analysis can be described as follows (Brown et al., 1999). Assume there exists a hypothetical economic relation described by[7]:

[7] Equivalently: $z_i = \beta_0 + \beta_1 u_i + \beta_2 w_i + \epsilon_i$. However, in this section matrix notation will be used since the discussion here is for expositional purposes only and matrix notation greatly simplifies the used expressions.

$$\boldsymbol{z} = \beta_0 \boldsymbol{\iota} + \beta_1 \boldsymbol{u} + \beta_2 \boldsymbol{w} + \boldsymbol{\epsilon} \tag{3.8}$$

In (3.8) $\boldsymbol{z}$ is a (N x 1) vector of dependent variables, $\boldsymbol{\iota}$ is an (N x 1) vector of ones, $\boldsymbol{u}$ and $\boldsymbol{w}$ are (N x 1) vectors of independent variables and $\boldsymbol{\epsilon}$ is a (N x 1) vector of error terms. It is assumed that $\boldsymbol{z}$, $\boldsymbol{u}$ and $\boldsymbol{w}$ are unobservable. Relation (3.8) depicts the real economic process underlying $\boldsymbol{z}$ in which a researcher is interested. This could be the association between the price of a share ($\boldsymbol{z}$) and earnings per share ($\boldsymbol{u}$) and book values per share ($\boldsymbol{w}$), free of scale effects. In this sense (3.8) could be considered as a scale free version of (3.1) with ϵ_i the pricing error for firm i.

Consequently, consider the following adaption of (3.8) incorporating a specific type[8] of scale effect and scaled data[9]:

$$\boldsymbol{Sz} = \beta_0 \boldsymbol{S\iota} + \beta_1 \boldsymbol{Su} + \beta_2 \boldsymbol{Sw} + \boldsymbol{S\epsilon} \tag{3.9}$$

In this $\boldsymbol{S}$ is an (N x N) diagonal matrix with elements $s_{i,i}$, which is the scale factor for firm i. Relation (3.9) should then be estimated with the following regression[10]:

$$\boldsymbol{y} = \beta_0 \boldsymbol{S\iota} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{3.10}$$

In (3.10) some variables have been renamed to their observable counterparts. Here $\boldsymbol{y} = \boldsymbol{Sz}$ the observable share prices; $\boldsymbol{X} = (\boldsymbol{\iota}\ \boldsymbol{Su}\ \boldsymbol{Sw})$ which is the data matrix containing observable earnings per share $\boldsymbol{Su}$ and observable book values per share $\boldsymbol{Sw}$; $\boldsymbol{\beta} = (\alpha\ \beta_1\ \beta_2)'$ and $\boldsymbol{\varepsilon} = \boldsymbol{S\epsilon}$. Explicitly note that the scale variables $\boldsymbol{S}$ are unobservable. This regression includes the constant α for econometrical reasons (Greene, 2008), but based on (3.9) $\alpha = 0$.

Since (3.10) cannot be estimated, in empirical work the following relationship is often estimated:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{3.11}$$

This gives rise to the following coefficient estimates by OLS:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{3.12}$$

[8] In this example it is assumed that scale effects are proportional, which does not have to be the case empirically, see section 3.3.4 and Gu (2007).
[9] Equivalently: $s_i z_i = \beta_0 s_i + \beta_1 s_i u_i + \beta_2 s_i w_i + s_i \epsilon_i$
[10] Equivalently: $y_i = \alpha + \beta_0 s_i + \beta_1 s_i u_i + \beta_2 s_i w_i + \varepsilon_i$

It can be shown that these coefficient estimates are biased. Substituting (3.10) and solving gives:

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \beta_0(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{S}\boldsymbol{\iota} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} \tag{3.13}$$

Assuming $\mathbb{E}[\boldsymbol{\epsilon} \mid \boldsymbol{X}, \boldsymbol{S}] = 0$, the conditional expectation of (3.13) gives:

$$\mathbb{E}\big[\widehat{\boldsymbol{\beta}} \mid \boldsymbol{X}, \boldsymbol{S}\big] = \boldsymbol{\beta} + \beta_0(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{S}\boldsymbol{\iota} \tag{3.14}$$

This bias in coefficients arises in this example because of an omitted variable (Greene, 2008 and Brown et al., 1999), which is $\boldsymbol{S}\boldsymbol{\iota}$. Barth and Kallapur (1996) show that scale effects also give rise to heteroskedasticity. To investigate the influence of scale effects on coefficient bias and standard errors, which are affected due to heteroskedasticity, Barth and Kallapur (1996) use simulated data. They show that the effects can be severe and might lead to wrong inferences. To mitigate scale effects Barth and Kallapur (1996) propose that empirical researchers incorporate a proxy for scale and that White (1980) standard errors are used. Deflating variables by a proxy for scale only decreases coefficient bias by a small amount or might even increase bias. Depending on the proxy for scale, standard errors from deflated regression could be over- or underestimated, leading to invalid inferences (Barth and Kallapur, 1996).

Furthermore, Brown et al. (1999) show that the coefficient of determination is also influenced by scale effects. To see this, note that the coefficient of determination is defined as:

$$R^2 \equiv (\boldsymbol{y}'\boldsymbol{M}\boldsymbol{y})^{-1}\widehat{\boldsymbol{y}}'\boldsymbol{M}\widehat{\boldsymbol{y}} \tag{3.15}$$

The matrix $\boldsymbol{M}$ is used to create a variable's variation around its mean, e.g. $\boldsymbol{M}\boldsymbol{y} = \boldsymbol{y} - \overline{\boldsymbol{y}}$, where $\boldsymbol{M} \equiv \left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}'\right)$, with $\boldsymbol{I}$ the (N x N) identity matrix. Using the fact that:

$$\boldsymbol{M}\widehat{\boldsymbol{y}} = \boldsymbol{M}\boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{3.16}$$

And the fact that $\boldsymbol{M}$ is an idempotent and symmetric matrix, equation (3.15) can be rewritten as:

$$R^2 = (\boldsymbol{y}'\boldsymbol{M}\boldsymbol{y})^{-1}\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{M}\boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{3.17}$$

From (3.17) it is intuitively clear that the coefficient of determination depends on the biased estimates $\widehat{\boldsymbol{\beta}}$ which in turn are dependent on scale effects.[11]

Brown et al. (1999) also show that even in the case of unbiased coefficient estimates the coefficient of determination might be influenced by scale effects. This result holds both in finite samples as when $N \to \infty$. More precisely, when the coefficient of variation[12] of scale, ω_s, changes, the coefficient of determination also changes although the direction of change is unclear (Gu, 2005). Since it is unlikely that ω_s is constant over time, comparisons of coefficients of determination over time might yield invalid results. This result also holds for cross sectional comparisons.

To correct for the effect of scale, Brown et al. (1999) propose two alternative ways. First of all (3.7) can be altered to include proxies for ω_s:

$$\hat{R}_t^2 = \delta_0 + \delta_1 Time_t + \delta_2 \widehat{\omega}_{P,t} + \delta_3 \widehat{\omega}_{BV,t} + \vartheta_t \qquad (3.18)$$

In this, $\widehat{\omega}_{P,t}$ and $\widehat{\omega}_{BV,t}$ are the estimated coefficients of variation for year t of share prices and book value of equity per share respectively. Alternatively, Brown et al. (1999) suggest that deflated variables are used when estimating (3.1). Note that (3.1) already is deflated by number of shares. Brown et al. (1999) argue that it should also be deflated by the price per share at $t-1$, as they find that market value is the most appropriate proxy for scale. As mentioned earlier, deflating variables might not levy coefficient bias, while incorporating a proxy for scale in (3.1) would (Barth and Kallapur, 1996). Since incorporating additional variables would affect the coefficient of determination and since regression coefficients are not of viable interest in this type of value relevance studies, Brown et al. (1999) reject this approach. Finally, the deflated version of (3.1) is referred to as a return model, while (3.1) is referred to as a level model (Gu, 2005).

Gu (2005) further researches the two corrections proposed by Brown et al. (1999). To properly assess his remarks, an important concept is the derivative of the probability limit of R^2 with respect to ω_s^2, i.e. $\frac{\partial plim(R^2)}{\partial \omega_s^2}$. Brown et al. (1999) derive an expression for $\frac{\partial plim(R^2)}{\partial \omega_s^2}$ and argue that:

$$\forall\, \omega_s : \frac{\partial plim(R^2)}{\partial \omega_s^2} > 0 \qquad (3.19)$$

[11] Note that this is only indicative; to fully assess the impact of scale effects on R^2 the behaviour of $plim(R^2)$, i.e. R^2 for $N \to \infty$, is also required. An expression for this is derived by Brown et al. (1999).

[12] The coefficient of variation is defined as $\omega = \left|\frac{\sigma}{\mu}\right|$ (Brown et al., 1999).

Gu (2005) and Brown et al. (2002) note however that this does not hold. The sign of $\frac{\partial plim(R^2)}{\partial \omega_s^2}$ is dependent on ω_s.

Gu (2005) then shows that small variations in ω_s can have a profound impact on the sign of $\frac{\partial plim(R^2)}{\partial \omega_s^2}$ especially when $\omega_s < 1$. This is of interest because for early sample years the proxy $\widehat{\omega}_{BV} < 1$ (Brown et al., 1999). It therefore is an important question how well $\widehat{\omega}_{BV}$ proxies ω_s because small errors might render the proxy useless. Another remark related to this issue is the unilateral way in which $\widehat{\omega}_P$ and $\widehat{\omega}_{BV}$ are included in (3.18). This imposes the restriction that an increase in one of these proxies always has the same effect on R^2. Again since (3.19) is not true, this restriction is invalid even if either $\widehat{\omega}_{BV} = \omega_s$ or $\widehat{\omega}_P = \omega_s$. Furthermore, Gu (2005) argues that the proxies $\widehat{\omega}_P$ and $\widehat{\omega}_{BV}$ include information both about scale effects as well as value relevance. Including them in (3.18) and attributing all their contribution to the scale effect only, understates the change in value relevance. The actual change in value relevance would in such a case be captured not only by δ_1, but also by δ_2 and δ_3 in (3.18).

Finally, Gu (2005) comments on the use of deflated variables to correct for the scale effect. One of the assumptions made by Brown et al. (1999) to derive their analytical results is that the scale factor is independent from the scale free variables. If share prices at $t-1$ are used as a proxy for the scale factor this assumptions is likely to be violated. Notice that share prices at t and $t-1$ are generally not independent. This holds similarly for share prices at $t-1$ and either earnings per share or book values per share (Gu, 2005). Gu (2005) also shows that deflating variables might lead to a new scale factor instead of eliminating scale effects. The influence of these issues on the coefficient of determination is however not clear.

3.3.3 General incomparability of coefficients of determination

It can be shown that coefficients of determination are incomparable across samples even in the absence of scale effects (Gu, 2007). Recall the definition of the coefficient of determination given by (3.17). This definition can be applied much more widely than only for the two factor model discussed in section 3.3.2. It holds for any OLS regression including a (N x K) data matrix $\boldsymbol{X}$ with a (K x 1) vector of estimated coefficients $\widehat{\boldsymbol{\beta}}$. To see why coefficients of determination are incomparable across samples, consider the following (Gu, 2007). Equation (3.17) can be rewritten as:

$$R^2 = 1 - \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\left(\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{M}\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\right)^{-1} \tag{3.20}$$

Which is equivalent to:

$$R^2 = 1 - \hat{\sigma}_\varepsilon^2\left(\widehat{\boldsymbol{\beta}}'\widehat{\boldsymbol{\sigma}}_{\boldsymbol{X}}^2\widehat{\boldsymbol{\beta}} + \hat{\sigma}_\varepsilon^2\right)^{-1} \tag{3.21}$$

In which $\hat{\sigma}_\varepsilon^2$ is the variance of the estimated errors:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N-1}\sum_{i=1}^{N}\hat{\varepsilon}_i^2 \tag{3.22}$$

And $\widehat{\boldsymbol{\sigma}}_{\boldsymbol{X}}^2$ is the estimated variance covariance matrix of the sample data:

$$\widehat{\boldsymbol{\sigma}}_{\boldsymbol{X}}^2 = \frac{1}{N-1}(\boldsymbol{X} - \overline{\boldsymbol{X}})'(\boldsymbol{X} - \overline{\boldsymbol{X}}) \tag{3.23}$$

Gu (2007) argues that the inclusion of $\widehat{\boldsymbol{\sigma}}_{\boldsymbol{X}}^2$ in the definition of the coefficient of determination makes it inappropriate for across sample comparisons. If we assume that scale effects are not present and that $\widehat{\boldsymbol{\beta}}$ and $\hat{\sigma}_\varepsilon^2$ are unbiased and consistent estimators of respectively $\boldsymbol{\beta}$ and σ_ε^2, even then will the coefficient of determination depend on the variance covariance matrix of the sample data. Since this is a property inherent to the population under review, across sample comparisons of the coefficient of determination are only relevant when the population is unchanged and samples are drawn randomly only from this population. Furthermore, Gu (2007) remarks that the effect of changes in $\widehat{\boldsymbol{\sigma}}_{\boldsymbol{X}}^2$ on the coefficient of determination, i.e. $\frac{\partial plim(R^2)}{\partial \boldsymbol{\sigma}_{\boldsymbol{X}}^2}$, are practically unpredictable. Finally, since the population of firms has significantly changed over the last 30 years, the coefficient of determination is not fitted for across sample comparisons in this setting. Gu (2007) instead proposes an alternative measure based on the residual variance σ_ε^2. This measure will be reviewed more thoroughly in section 3.3.5.

3.3.4 The variance of the logarithm metric

To assess the work of Chang (1999), who proposed the variance of the logarithm metric, it is necessary to reconsider (3.1). After careful consideration of the specification of

(3.1) a natural question arises. For two firms A and B with share prices of \$ 10 and \$ 100 respectively, is a pricing error of \$ 1 per share equivalent? The answer to this question is obviously no, since for firm A this represents a much larger relative pricing error. Note that (3.1) treats such a situation as equivalent (Chang, 1999). This comment is clearly related to the scale effect discussed in the previous section, although Chang (1999) does not refer to it explicitly. Instead, Chang (1999) argues that (3.1) is incorrectly specified based on additive errors, instead of multiplicative errors. This is similar to stating that the scale effect leads to heteroskedasticity of error terms (Ye, 2007).

To remedy the concerns above consider the following. Assume some estimated firm price $\hat{P}_{i,t}$ for firm i at t and the actual market value of the firm $P_{i,t}$. The estimated firm price can either come from (3.1) or some more advanced accounting based valuation model. The latter is what Chang (1999) recommends although he finds that estimating firm value based on (3.1) does not alter his results. Finally consider the pricing error $\varepsilon_{i,t} = P_{i,t} - \hat{P}_{i,t}$. These errors do not take into effect firm size, so Chang (1999) suggests that $\varepsilon_{i,t}$ is scaled by $P_{i,t}$. It should be clear that this is almost equivalent to deflating as suggested by Brown et al. (1999). One nice insight provided by Chang (1999) is that this kind of scaling does not put the same weight on over- and under pricing of firms. As a solution to this problem, Chang (1999) proposes that the natural logarithm of the ratio $\hat{P}_i/P_i$ is used. Finally the variance of the logarithm of the ratio $\hat{P}_i/P_i$ is then an inverse metric for value relevance. More clearly:

$$variance\ of\ the\ logarithm\ metric\ for\ year\ t \equiv Var\left\{\ln\left(\frac{\hat{P}_{i,t}}{P_{i,t}}\right)\right\} \qquad (3.24)$$

The interconnection between this metric and the earlier discussed coefficient of determination can also be shown more clearly, see Chang (1999). The final step of this methodology is to regress the variance of the logarithm metric over time, in a fashion similar to (3.7).

It is insightful to compare the variance of the logarithm metric to the earlier comments on scale effects. If the firm values $\hat{P}_{i,t}$ are obtained by estimating some version of (3.1) scale effects are again clearly present. The estimated coefficients would be biased and the estimated firm values are not in line with the true underlying economic relation. Deflating variables decreases this effect but it is better to incorporate a proxy for scale which should be uncorrelated with the other independent variables (Barth and Kallapur, 1996). As was stated earlier, deflating variables could also lead to new scale effects and the deflator is unlikely to be statistically independent of the scale factor (Gu, 2005). Since the variance of logarithm metric is closely related to the coefficient of determination, these issues are likely to be relevant for the variance of logarithm metric as well, but their effect is not clear.

Following Chang (1999) and estimating firm value with a more elaborate framework was found to be more appropriate, but there is a caveat to this technique. The extended accounting based valuation model discussed by Chang (1999) requires a lot of inputs such as industry earnings persistence and expected return on equity capital. These inputs are also not readily observable and have to be estimated, which could lead to errors-in-variables complications (Campbell et al., 1997).

3.3.5 The residual dispersion metric

As was noted in section 3.3.3 the coefficient of determination is unsuitable for across sample comparisons if those samples have been drawn from different populations. The reason for this is that the coefficient of determination not only measures changes in value relevance, but also changes in the variance of the independent variables. A better suited metric to assess changes in value relevance is the variance of the residual pricing error (Gu, 2007). The reason for this is that, assuming scale free data, σ_ε^2 will only change if the underlying economic relation changes. Furthermore if data is scale free then $\mathbb{E}[\hat{\sigma}_\varepsilon^2 \mid \boldsymbol{X}] = \sigma_\varepsilon^2$ and $plim(\hat{\sigma}_\varepsilon^2) = \sigma_\varepsilon^2$, that is $\hat{\sigma}_\varepsilon^2$ is an unbiased and consistent estimator of σ_ε^2 (Gu, 2007). The residual dispersion metric also makes intuitive sense; years with high value relevance should have a low residual dispersion and vice versa.

The residual dispersion metric is closely related to the variance of the logarithm metric. To see this note that (3.1) can also be specified with a multiplicative error (Chang, 1999):

$$P_{i,t} = (\alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}BV_{i,t})(1 + \varepsilon_{i,t}) \tag{3.25}$$

Taking logarithms:

$$\ln(P_{i,t}) = \ln(\alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}BV_{i,t}) + \ln(1 + \varepsilon_{i,t}) \tag{3.26}$$

The unknown parameters can be estimated by non-linear regression which would result in:

$$\ln(P_{i,t}) = \ln(\hat{\alpha}_{0,t} + \hat{\alpha}_{1,t}X_{i,t} + \hat{\alpha}_{2,t}BV_{i,t}) + \ln(1 + \hat{\varepsilon}_{i,t}) \tag{3.27}$$

Define:

$$\ln(\hat{P}_{i,t}) = \ln(\hat{\alpha}_{0,t} + \hat{\alpha}_{1,t}X_{i,t} + \hat{\alpha}_{2,t}BV_{i,t}) \quad (3.28)$$

Rewriting gives:

$$-\ln\left(\frac{\hat{P}_{i,t}}{P_{i,t}}\right) = \ln(1 + \hat{\varepsilon}_{i,t}) \quad (3.29)$$

Thus:

$$Var\left\{\ln\left(\frac{\hat{P}_{i,t}}{P_{i,t}}\right)\right\} = Var\{\ln(1 + \hat{\varepsilon}_{i,t})\} \quad (3.30)$$

Although either the variance of the logarithm metric or the residual dispersion metric is an improvement over the coefficients of determination metric, there still remains the problem of scaling. In the discussion of σ_{ε}^2 it was assumed that data is scale free, which obviously is not the case. The variance logarithm metric does take into account scaling, but similar to the improvements proposed by Brown et al. (1999), it can only incorporate proportional scale effects. Gu (2007) finds that the scale effect is not proportional; the pricing error for a firm with a share price of \$ 50 is not five times as large as the pricing error for a firm with share prices of \$ 10. This non-linearity in scale effects makes standardization of pricing errors, i.e. by deflating variables, unsuitable.

A number of alternative approaches to take into account non linear scale effects are proposed by Gu (2007), but only the abnormal pricing error methodology will be discussed in detail here.[13] The steps for this specific approach are as follows. First of all, estimate a model describing a relation between share prices or returns and accounting information such as (3.1).[14] Secondly, pool all data across years and sort it into deciles based on the absolute fitted values $|\hat{P}_{i,t}|$. The pricing errors resulting from OLS regressions are then calculated for each decile. The pricing error is defined very similar to the sample standard deviation $\hat{\sigma}_{\varepsilon}$, hence the name residual dispersion metric:

[13] The two alternative approaches proposed by Gu (2007) to correct for the non linearity of scale effects are to use matched samples or compare changes in value relevance over time only within deciles based on scale.

[14] Gu (2007) also estimates two return (deflated) models based on (3.1). The focus in this discussion will however solely be on the application of the residual dispersion metric to (3.1), since the methodology of the residual dispersion metric is independent of the estimated model.

$$pricing\ error\ (1) \equiv \sqrt{\frac{1}{n}\sum_{j=1}^{n}\hat{\varepsilon}_j^2} \qquad (3.31)$$

Here is n the number of observations in each decile and the subscript j is used to indicate a single observation in each decile. In the sample used by Gu (2007) these steps alone already clearly show the non-proportionality of scale effects. The mean fitted value of the tenth decile is approximately thirteen times as large as that of the first decile, but the pricing error calculated by (3.31) is less than three times as large for the tenth decile as for the first decile (Gu, 2007).

To correct for the non linearity of scale effects, Gu (2007) calculates abnormal pricing errors, which can be done in two distinct ways. The first alternative is to calculate the mean absolute fitted value for each year and identify the scale decile corresponding to this mean value. The pricing error of the decile is then used as a benchmark pricing error. Finally, the abnormal pricing error for year t is calculated as the "raw" pricing error of year t, calculated again by (3.31), minus the benchmark pricing error of the corresponding decile.

Alternatively, the benchmark pricing error for each decile can be calculated by:

$$pricing\ error\ (2) \equiv \frac{1}{n}\sum_{j=1}^{n}\left|\hat{\varepsilon}_j\right| \qquad (3.32)$$

Using this definition of pricing errors, Gu (2007) calculates the abnormal pricing errors on an observation by observation basis. That is; the abnormal pricing error for a single observation is calculated as the absolute raw pricing error for that observation minus the benchmark pricing error, calculated by (3.32), for the decile corresponding to that observation. For each year the mean abnormal pricing error is then the average of the abnormal pricing errors of that year.

The final step is to see if there have been any changes in abnormal pricing errors over the years. Again note that an increase in abnormal pricing errors would arguably indicate a decrease in value relevance over time. To asses this, Gu (2007) regresses the abnormal pricing error against a time variable as was done in (3.7). This step is exactly similar to that of Collins et al. (1997) only the coefficient of determination is replaced by the abnormal pricing error for year t. Therefore a positive (negative) estimate of δ_1 in this specification indicates a decrease (increase) in value relevance.

3.4 The coefficients metric

The measures of value relevance discussed in the previous sections all are related in some sense to the variance of the residuals resulting from various relations. This metric made intuitive sense, because an increasing (decreasing) variance of residuals indicates a deteriorating (improving) relationship. There is however another metric resulting from regression analysis that can be used as an indicator for the strength of this relationship, which are the coefficient estimates. An application of this can be found in Lev and Zarowin (1999).

In addition to changes in the coefficient of determination Lev and Zarowin (1999) also report changes in both earnings response coefficients and cash flow response coefficients. These metrics are obtained by the coefficient estimates from the following two returns models (Lev and Zarowin, 1999):

$$R_{i,t} = \alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}\Delta X_{i,t} + \varepsilon_{i,t} \tag{3.33}$$

Where $R_{i,t}$ are the share returns of firm i in year t, $X_{i,t}$ are the earnings before extraordinary items of firm i in year t and $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$ which is the change in earnings. This regression shall be referred to in subsequent chapters as the earnings-returns relation Also:

$$R_{i,t} = \beta_{0,t} + \beta_{1,t}CF_{i,t} + \beta_{2,t}\Delta CF_{i,t} + \beta_{3,t}AAC_{i,t} + \beta_{4,t}\Delta AAC_{i,t} + \epsilon_{i,t} \tag{3.34}$$

In (3.34); $CF_{i,t}$ are the cash flows from operations of firm i in year t, $\Delta CF_{i,t} = CF_{i,t} - CF_{i,t-1}$ which is the change in cash flows from operations, $AAC_{i,t}$ are the annual reported accruals of firm i in year t and $\Delta AAC_{i,t} = AAC_{i,t} - AAC_{i,t-1}$ is the change in annual accruals. Finally the annual accruals are defined as $AAC_{i,t} = X_{i,t} - CF_{i,t}$.

Using the output of these regressions Lev and Zarowin (1999) define:

$$ERC_t \equiv (\hat{\alpha}_{1,t} + \hat{\alpha}_{2,t}) \tag{3.35}$$

$$CFRC_t \equiv (\hat{\beta}_{1,t} + \hat{\beta}_{2,t}) \tag{3.36}$$

In (3.35) ERC_t stands for the earnings response coefficient in year t and in (3.36) $CFRC_t$ stands for the cash flow response coefficient in year t. To assess changes in value relevance these response coefficients are regressed against a time variable in a similar way as in (3.7).

Finally, although this has not been done by Lev and Zarowin, one could consider if this methodology might also be applied to (3.1).

It should be clear that some comments can be made about the coefficients metric of value relevance. If (3.33) or (3.34) describe the true data generating model underlying returns, the coefficient estimates resulting from these regressions are consistent and unbiased. A short reflection on the earlier discussion of scale effects poses serious doubts to these claims and as has been shown, scale effects lead to biased coefficient estimates. This result holds true even for the deflated regressions (3.33) and (3.34) (Barth and Kallapur, 1996). This unknown bias renders any across time comparison of coefficients useless.

As a way to circumvent this problem, a proxy for scale could be included in either (3.33) or (3.34). It is shown by Barth and Kallapur (1996) that including a proxy for scale is more effective in reducing bias from scale effects than deflating variables. Including a proxy for scale in this setting might however lead to other problems, namely multicollinearity. Although coefficient estimates are still unbiased and consistent in the presence of multicollinearity, they could be over- or underestimated in finite samples (Greene, 2008). Since the magnitude of coefficient estimates is very important in this setting, this measure of value relevance does not seem to be the most appropriate.

3.5 Portfolio metrics

Francis and Schipper (1999) propose an alternative measure of value relevance, which is based on the market-adjusted returns that could be earned with perfect foresight of accounting information. The rationale behind this metric is that accounting information is value relevant if perfect foresight of this information produces high market-adjusted returns. Francis and Schipper (1999) favour this measure of value relevance, because it controls for changes in the volatility of returns over time, while other metrics do not. A similar methodology to assess the value relevance of accounting information has also been used by Alford et al. (1993) in an across countries comparison.

As a general note, Francis and Schipper (1999) use market adjusted returns in their research design, calculated as the cum-dividend compounded 15-month return starting at the beginning of fiscal year t. These returns for firm i in year t will be denoted by $R_{i,t}$. At the start of each year hedge portfolios are formed based on the accounting information available at the end of year t. The return that could be earned with these accounting based hedge portfolios is then compared to the return of a hedge portfolio based on perfect foresight of returns, i.e. a portfolio taking long positions in shares with positive returns and short positions in shares with negative returns (Francis and Schipper, 1999).

Furthermore, Francis and Schipper (1999) form five different portfolios based on accounting information to evaluate changes in value relevance over time. One of these portfolios will be discussed in detail to illustrate this methodology. This specific hedge portfolio is based on the sign of earning changes in year t. More clearly, this hedge portfolio takes a long position if $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$ is positive and a short position if $\Delta X_{i,t}$ is negative. Here $X_{i,t}$ is again defined as the earnings for firm i in year t. After the long and short positions for this portfolio are determined, one can calculate the return this portfolio would have earned in year t, defined as R_t^*. Then also the return for a hedge portfolio based on perfect foresight of returns is calculated, which is defined as R_t^H. The portfolio metric is defined as:

$$portfolio\ metric\ for\ year\ t \equiv R_t^* / R_t^H \tag{3.37}$$

An increase (decrease) in this metric would indicate an increase (decrease) of value relevance over time. To assess this, the portfolio metric is regressed against a time variable similarly to (3.7). Alternatively, Francis and Schipper (1999) also test a non linear specification of (3.7) including squared time as an independent variable.

Besides the above described hedge portfolio, Francis and Schipper (1999) also investigate changes in value relevance based on the following portfolios:

- Earnings portfolio; this portfolio takes positions based on the magnitude and sign of $\Delta X_{i,t}$. Each year firms are ranked according to $\Delta X_{i,t}$ and this portfolio takes a long (short) position in the highest (lowest) 40%.
- Cash portfolio; similarly to the earnings portfolio, but for changes in cash flows. Cash flows are approximated by net income before extraordinary items plus depreciation, amortization deferred tax expenses and minority interest and minus changes in working capital.
- Ratio portfolio; this portfolio is based on the financial ratio model described by Lev and Thiagarajan (1993). Long and short positions are taken respectively in the top and bottom 40% of returns predicted by this model.
- Ratio 2 portfolio; this portfolio is based on the predicted returns from a deflated adoption of (3.1) also including $\Delta X_{i,t}$ as an independent variable. Again long (short) positions are taken in the top (bottom) 40% of predicted values by this model.

An extension of the portfolio metric is provided by Chang (1999), who proposes a hedge portfolio based on return on equity. This selection criterion both incorporates earnings and book values, making it a multivariate extension of the univariate earnings portfolio (Chang, 1999). Finally, a comment about the portfolio metric is provided by Gu (2007), who argues that this metric is also sensitive to changes in the variance of data. Since the return on any of these portfolios is based on the variance between the long and short positions, an increase (decrease) in value relevance could occur because of a mere increase (decrease) in the variance of returns around zero. This effect is reduced by scaling R_t^* by R_t^H, but it is not eliminated (Gu, 2007).

3.6 Alternative interpretation of value relevance

All of the measures of value relevance discussed in the previous sections are operationalisations of value relevance under Interpretation 4. In this section, an operationalisation of value relevance under Interpretation 2 will be presented. As was discussed in chapter two, Interpretation 2 states that accounting information is value relevant if it contains the variable used in a valuation model or assists in predicting those variables (Francis and Schipper, 1999). One of the main inputs in a standard valuation model such as the discounted cash flow model, are future cash flows (Koller et al., 2005). As an operationalisation of Interpretation 2, Kim and Kross (2005) therefore investigate the ability of current earnings to predict future cash flows.

The methodology applied by Kim and Kross (2005) is similar to that of Collins et al. (1997). First of all, the following regression, taken from Dechow et al. (1998) is estimated:

$$CF_{i,t+1} = \alpha_0 + \alpha_1 CF_{i,t} + \alpha_2 X_{i,t} + \varepsilon_{i,t} \tag{3.38}$$

In (3.38) $CF_{i,t}$ are the cash flows of firm i for year t and $X_{i,t}$ is again defined as earnings. Cash flows are defined by Kim and Kross (2005) as income before depreciation minus interest expense, taxes and changes in working capital plus interest revenue. This definition of cash flows can also be found in Dechow et al. (1998). Secondly, to investigate changes in value relevance over time Kim and Kross (2005) compare the coefficients of determination resulting from (3.38) using (3.7), including decomposition between earnings and cash flows. This decomposition is similar to that performed by Collins et al. (1997), described in section 3.2.

After the discussion of the shortcomings of the methodology of Collins et al. (1997) in section 3.3, this research design does not appear to be very solid. However, the results obtained by Kim and Kross (2005) are robust to most of the comments made in section 3.3. For example, Kim and Kross (2005) deflate the independent variables in (3.38) by total assets to reduce scale effects, as proposed by Brown et al. (1999). To levy the concerns of Gu (2007) about changes in sample variance, Kim and Kross (2005) also use rank regression. Their results are robust to these alterations and, more interestingly, contradict the general finding of decreasing value relevance.

3.7 Summary

One of the most important issues in value relevance studies is empirically measuring value relevance. Since value relevance is not readily observable this becomes a complicated matter. As will be shown in the next chapter, the choice of metric can have substantial consequences for empirical results. At the introduction of this chapter three main questions were posed about measuring changes in value relevance. The answer to these questions is unfortunately not straightforward and depends on a variety of factors.

First of all, in choosing the variables to consider there are several guidelines. The first consideration is the interpretation of value relevance that will be applied. Generally speaking, the most common interpretation of value relevance is Interpretation 4, while Interpretation 2 is also applied in the literature. To recapitulate, Interpretation 2 is concerned whether financial statement information contains the variables used in a valuation model or assists in predicting those variables, while Interpretation 4 is concerned with the statistical association between financial information and share prices or returns. Interpretation 4 still allows for a wide array of operationalisations, since it merely states that there should be a statistical association between accounting variables and share prices or returns. A guideline to choose which variables should be considered is provided by Ohlson's (1995) accounting based valuation model. This framework provides a theoretical underpinning for (3.1), which, in returns or level specification, is the most used framework.

Secondly, to measure the strength of the association between accounting variables and share prices or returns, a number of metrics exist. There is not a single best metric, but it should be clear from the previous discussion that some metrics are superior to others. The metrics which are most robust to scale effects and related problems of heteroskedasticity and bias are the residual dispersion metric and the portfolio metric. Finally, to assess changes in value relevance over time, all discussed research relies on some version of (3.7).

As a general note, value relevance researchers should be particularly aware of the influence of methodology on obtained results. It is therefore worthwhile to control obtained results for robustness across model specifications, alternative metrics and ideally also interpretations.

4. Overview of empirical findings

4.1 Introduction

In this chapter an overview of empirical findings concerning changes in value relevance is provided. This discussion will highlight the importance of the used metric for empirical inferences. Since most researchers are also concerned with determining the causes of changes in value relevance, the importance of the used metric is further emphasized. The empirical findings in this chapter are mostly based on the measures of value relevance discussed in chapter three, but also some additional research will be presented.

The outline of this chapter is as follows. Section 4.2 presents findings on changes in value relevance with section 4.2.1 focusing on the metrics discussed in chapter three and section 4.2.2 focusing on additional findings.[15] In section 4.3 an overview of possible causes for changes in value relevance is provided as well as empirical results concerning these causes. Section 4.4 summarizes these results and provides an overview of relevant research.

4.2 Changes in value relevance

4.2.1 Empirical findings of changing value relevance

As was stated in chapter three, one of the earliest studies investigating changes in value relevance over time was performed by Collins et al. (1997). Using their coefficient of determination metric, Collins et al. (1997) find that the value relevance of earnings and book values combined has not decreased over the years. This result is based on the methodology described in section 3.2. If there has been any change in value relevance, Collins et al. (1997) indicate that this is more likely to be an increase than a decrease. Furthermore, Collins et al. (1997) argue that the value relevance of earnings has declined, but that this decline is compensated by an increase of value relevance of book values. The samples under investigation in this research are U.S. based firms over the period 1953-1993.

[15] As a general note, the metrics discussed in chapter three will often be referred to in both this chapter and chapter five. For the sake of readability, the exact specification of each researcher will not be elaborated on. It should be noted however, that small differences do exist, for example specifying earnings with or without extraordinary items or returns calculated over different time intervals.

The results of Collins et al. (1997) are reinforced by Francis and Schipper (1999). Francis and Schipper (1999) performed an analysis comparable to that of Collins et al. (1997) and also found an increase in value relevance of earnings and book values combined. Besides estimating (3.1), Francis and Schipper (1999) also estimate a regression similar to (3.33), linking earnings and returns, and an alternative regression linking share prices to per share values of assets and liabilities. Again, a decrease of value relevance of earnings is found, while the value relevance of balance sheet information has increased. The sample under investigation are U.S. firms from 1952-1994. Finally, these results are also found by Ely and Waymire (1999) using data from 1927-1993.

Francis and Schipper (1999) find some contradicting results using their portfolio metric. Of the hedge portfolios described in section 3.5 the earnings sign portfolio and the two ratio portfolios perform progressively worse over the years. The performance of the hedge portfolios based on the magnitude and sign of earnings change and on cash flows does not change across the sample period. The portfolio metric has also been applied by Chang (1999) using both a portfolio based on the magnitude and sign of earnings and a hedge portfolio based on return on equity. The latter hedge portfolio performs progressively worse over his sample period of 1953-1996, while the former does not. These results provide some evidence for declining value relevance.

Lev and Zarowin (1999), using the coefficients of determination metric in combination with (3.1), find a decrease of value relevance of earnings and book values which contradicts the findings of Collins et al. (1997) and Francis and Schipper (1999). Lev and Zarowin (1999) attribute this result to their differing sample from 1977-1996. They also find a decrease in the value relevance of both earnings and cash flows separately, both using the coefficients metric and the coefficient of determination metric. For these last results relations (3.33) and (3.34) are used.

As was discussed in section 3.3, the coefficient of determination metric in combination with a levels based regression, might suffer from scale effects and other econometrical issues. To investigate the results of these issues Brown et al. (1999) replicate the analysis of Collins et al. (1997), only using (3.18) instead of (3.7), i.e. incorporating proxies for scale in the regression of coefficients of determination against time. This alteration significantly influences the results of Collins et al. (1997); Brown et al. (1999) find that the value relevance of earnings and book values combined has decreased over time. The value relevance of book values on a standalone basis has however increased, while the value relevance of earnings has decreased. This result is further confirmed by a returns based version of (3.1) (Brown et al., 1999).

A decrease of value relevance of earnings and book values is also established by Chang (1999) using the variance of the logarithm metric. Chang (1999) investigates where the difference between his results of declining value relevance and those of Collins et al. (1997) and Francis and Schipper (1999) arise from. He concludes that this is due to the multiplicative error assumption, thereby again highlighting the problem of scale effects in value relevance studies. Chang (1999) uses a sample period of 1953-1996 for this analysis.

Gu (2007) argues that studies of value relevance based on coefficients of determination are inappropriate even when properly accounted for scale effects. Gu (2007) also comments on the proportional scale effect assumption made by both Brown et al. (1999) and Chang (1999). His study based on the residual dispersion metric, does however not change the findings of these earlier studies. Gu (2007) also concludes that value relevance has declined over his sample period of 1953-1998, especially after 1970. This result indicates that some structural changes in value relevance could be of interest.

Interestingly, the study performed by Kim and Kross (2005) finds an increase of value relevance over the period 1972-2001. This result is obtained using the methodology described in section 3.6 which is based on Interpretation 2 instead of Interpretation 4, i.e. it is researched whether current cash flows are predictive for future cash flows. Kim and Kross (2005) also investigate if value relevance increases if (3.38) is changed to include prices instead of future cash flows as a dependent variable. This model specification changes their results and indicates a decrease in value relevance over time. Unfortunately, Kim and Kross (2005) are not able to provide a complete explanation for these differing results. Their conclusion is that the ability of earnings to predict future cash flows has increased, while the ability of earnings to explain stock prices has decreased over the years.

The research above can be summarized as follows. Value relevance of earnings and book values has most likely decreased in the period 1950-2000 in the U.S. The differing initial results of Collins et al. (1997), Francis and Schipper (1999) and Ely and Waymire (1999) are most likely due to scale effects, as these are not controlled for in these researches. Furthermore, the finding of increasing value relevance has not been established after these studies. The results of Kim and Kross (2005) are an exception to this, which is probably due to their alternative interpretation of value relevance.

4.2.2 Additional findings on changing value relevance

Ryan and Zarowin (2003) investigate changes in the earnings returns relation in a sample of U.S. firms in the period 1966-2000. For this they use a methodology distinctly different from those discussed in chapter three. The main objective of their paper is to investigate what drives the declining earnings return relation; a topic further explored in section 4.3. To analyse if there has been any change in the relation between earnings and returns, Ryan and Zarowin (2003) regress current year's earnings on current and lagged returns.[16]

The model applied by Ryan and Zarowin (2003) is an extension of the work of Ryan (1995) and Pope and Walker (1999) and it relates current year's earnings to a moving average process of shocks to market value and expected permanent earnings. This model is derived from three main assumptions, namely no arbitrage, the clean surplus relation and a specific assumption about the nature of accrual measurement.[17] To analyse changes in value relevance over time, Ryan and Zarowin (2003) investigate changes in estimated regression coefficients and coefficients of determination. They find that the value relevance of earnings has significantly decreased in their sample period.

Dontoh et al. (2007) investigate yet another methodology of measuring value relevance. Their research is mainly concerned with the ability of earnings and share prices to predict future firm payoffs. Intuitively, to examine this predictive ability of prices and earnings Dontoh et al. (2007) use a regression with either prices or net income as the dependent variable, regressed against future firm payoffs, such as dividends, as independent variables. The coefficients of determination of these regressions are then used as a proxy for the predictive content of either earnings or prices.[18] Dontoh et al. (2007) find that both of these have declined over time. The main contribution of Dontoh et al. (2007) is however that any decrease in value relevance over time might be induced by non-information-based trading. This issue will be discussed more extensively in section 4.3.

A concept closely related to value relevance is the informativeness, or information content, of earnings announcements. An operationalisation of this concept is provided by Landsman and Maydew (2002), who investigate abnormal trading volume and abnormal stock price volatility around quarterly earnings announcements. With this, Landsman and Maydew (2002) report on a phenomenon earlier established by Beaver (1968), namely that there is increase in both of these metrics around earnings announcements. Notice the difference between earnings announcement informativeness and value relevance. The former focuses mainly on the short horizon effects of accounting information on capital markets, while the latter investigates long term, mostly yearly, associations between accounting information and share prices or returns. Earnings announcement

[16] Explicitly note the difference with the models discussed in chapter three, which regress returns on earnings contemporaneously.
[17] See Ryan and Zarowin (2003) for further information.
[18] For a complete explanation of this measure see Dontoh et al. (2007).

informativeness studies normally apply an event study methodology.[19] Based on the metrics mentioned above, Landsman and Maydew (2002) do not find a decrease in the informativeness of earning announcements over time. If anything, there seems to be an increase in informativeness. This conclusion still holds when Landsman and Maydew (2002) control for a number of variables such as firm size and non-recurring items.

Francis et al. (2002) further investigate the increased informativeness of earnings announcements reported by Landsman and Maydew (2002). The research method used by Francis et al. (2002) is based on abnormal stock returns around earnings announcements. With their alternative metric, Francis et al. (2002) also find an increase in the informativeness of earnings announcements. This increased informativeness is further confirmed by Collins et al. (2009), who use a methodology similar to that of Landsman and Maydew (2002). Contrary to the general findings about value relevance, the informativeness of earning announcements thus seems to be increasing.

4.3 Possible causes for changing value relevance

4.3.1 Introduction

In this section a number of possible causes for changing value relevance will be discussed. Section 4.3.2 to 4.3.5 will start off with four causes for changing value relevance proposed by Collins et al. (1997), namely intangibles, nonrecurring items, negative earnings and size. In section 4.3.6 the influence of business changes, as proposed by Lev and Zarowin (1999), is discussed, while section 4.3.7 presents findings on the influence of conservatism. Section 4.3.8 discusses the effect of standard setting bodies on value relevance, paying attention to both specific research done by Ely and Waymire (1999) in the U.S., as well as to recent research on the result of the introduction of IFRS. In section 4.3.9 attention is paid to the findings of Dontoh et al. (2007) and the concept of non-information-based trading.

4.3.2 Intangibles and technology based firms

One of the factors influencing value relevance could be investments in intangibles. A study of Amir and Lev (1996) indicates that accounting information is not very relevant when valuing service or technology based companies. Intangibles include, but are not limited to, human capital, brand development and research and development. These investments increase the market value of firms, but are not properly recorded on balance

[19] For a description of the event study methodology see Campbell et al. (1997).

sheets (Collins et al, 1997). Amir and Lev (1996) confirm a low usefulness of financial statement information when valuing cellular telephone companies. If these findings hold for other industries as well and if the number of firms investing in intangibles has increased over time, this could have induced a decline in value relevance (Collins et al., 1997).

Collins et al. (1997) find partial support for these hypotheses. They do conclude that the number of firms investing in intangibles has increased, but do not find a significant decrease of value relevance over time due to this factor. If anything, the increased propensity of firms investing in intangibles has increased the value relevance of book values. This is highly contradictive to the a priori hypothesis of Collins et al. (1997) and unfortunately they do not provide a satisfactory explanation for this finding. Arguably this counterintuitive result could be due to a flawed methodology for measuring value relevance, as was discussed in chapter 3.

In a similar argument, Francis and Schipper (1999) investigate whether there is a difference in the value relevance of high- and low technology firms. Using both the portfolio metric and the coefficient of determination metric, they find only weak evidence of a decrease in value relevance due to an increasing number of high technology firms. Francis and Schipper (1999) also do not find that the value relevance of high technology firms has decreased with a higher pace compared to that of low technology firms.

4.3.3 Nonrecurring items

Another possible explanation for changing value relevance, provided by Collins et al. (1997) is the propensity of nonrecurring items reported. Elliot and Hanna (1996) show that capital markets do not value nonrecurring items symmetrically to earnings before nonrecurring items. This could be explained by their special and transitory nature (Collins et al., 1997). Furthermore, Elliot and Hanna (1996) and Maydew (1997) find that most nonrecurring items are losses. If this is combined with research from Basu (1997), which states that good and bad news are treated differently by capital markets and an increased propensity of firms to report nonrecurring items (Elliot and Hanna, 1996), this could partially explain a decline in value relevance (Collins et al., 1997).

Alternatively, Collins et al. (1997) argue that firms divesting noncore businesses or firms in financial distress might have a higher frequency of reporting nonrecurring items. If abandonment value is of increased importance in these firms and if abandonment value is related to book values, this could increase the value relevance of book values. Collins et al. (1997) find support for both of above described hypotheses. Their results indicate that the value relevance of book values and earnings combined and the value relevance of earnings on a standalone basis decreases with the number of non recurring items reported. Also the value relevance of book values on a standalone basis increases with the number of nonrecurring items reported.

4.3.4 Negative earnings

As was stated above, research from Basu (1997) indicates that markets react differently on good and bad news. Basu (1997, 3) particularly investigates the role of conservatism in accounting, which he describes as "*earnings reflecting bad news more quickly than good news*". Basu (1997) finds that the degree of conservatism has increased over the years. Hayn (1995) finds that capital markets react to a lesser extent to negative than to positive earnings. The argument behind this is that shareholders always have the possibility of abandoning the firm and therefore negative earnings cannot persist. These findings, combined with results of Hayn (1995) that the number of firms reporting negative earnings has increased over time, could explain a decrease of value relevance of earnings (Collins et al., 1997).

Collins et al. (1997) further argue that negative earnings might lead to an increased value relevance of book values, because the abandonment option becomes more important. Collins et al. (1997) find some evidence for their hypotheses. The effect of negative earnings on the value relevance of book values and earnings combined is positive and significant. However, the effect of negative earnings on either earnings or book values is not significant.

4.3.5 Size

A final factor that could influence value relevance identified by Collins et al. (1997 is firm size. Collins et al. (1997) note that book values are of increased importance in valuation, when current earnings are an inappropriate proxy for future earnings and when firms face an increased likelihood of abandonment. They conjecture that both of these factors are related to firm size. Collins et al. (1997) argue that small firms are more likely to include start-ups whose value is driven by earnings growth potential. This makes current earnings a bad predictor for future earnings. Also, Hayn (1995) shows that small firms are more likely to report negative earnings compared to large firms. As discussed above, this would lead to a relative increase in the value relevance of book values compared to earnings. Finally, Collins et al. (1997) expect that small firms face a greater probability of financial distress, thereby increasing the importance of the abandonment option and the importance of book values for valuation. These arguments combined with an increase in the proportion of small firms in the dataset under investigation (Collins et al., 1997) could affect value relevance.

Collins et al. (1997) find some evidence confirming their hypotheses. The value relevance of earnings and book values combined is positively correlated with the average size of firms. Likewise, if average firm size decreases, value relevance also decreases. Collins et al. (1997) find no support for an increase in the value relevance of book values as average firm size decreases. Contradicting their a priori beliefs, the value relevance of earnings is negatively correlated with firm size.

4.3.6 Business change

Lev and Zarowin (1999) further investigate changes in value relevance and identify a broader influencing factor, namely business change. Lev and Zarowin (1999) conjecture that the rate of business change in the economy has been increasing. Key factors driving business change are deregulations, innovations and competition. Furthermore, Lev and Zarowin (1999) argue that these changes are not adequately captured by accounting standards. There is a mismatch between the recognition of the costs of business change, such as R&D and restructurings, and the benefits. The costs of change are expensed immediately, while the benefits are often recorded later. Lev and Zarowin (1999) finally state that capital markets fully incorporate these changes as they become known which decreases the connection between accounting values and share prices or returns.

One of the issues with business change is that it is not readily observable. This is not the case for the factors identified by Collins et al. (1997), which are either observable or good proxies are easily obtained. To measure business change, Lev and Zarowin (1999) therefore propose the following. Each year all firms are sorted into ten deciles based on book value of equity or market value of equity. For each year and each firm the absolute change in decile is then recorded, i.e. if a firm was in decile seven in 1987 and is in decile two in 1988 its rank change is five. The average absolute rank change is then calculated for each year and is used as a proxy for business change. Lev and Zarowin (1999) clearly show an increasing rate of business change based on this measure. Lev and Zarowin (1999) also find that increasing levels of business change are connected to decreasing levels of value relevance.

Finally, Lev and Zarowin (1999) consider the influence of nonrecurring items and negative earnings on value relevance when corrected for business change. These factors become insignificant and Lev and Zarowin (1999) infer that nonrecurring items and negative earnings are merely symptoms of business change and not causes for declining value relevance. A similar argument cannot be made for research and development or investments in intangibles, which are instead causes of business change (Lev and Zarowin, 1999).

4.3.7 Conservatism

The topic of conservatism was briefly touched upon in section 4.3.3; in this section some additional evidence on the effects of conservatism on value relevance will be presented. One of the main researches in this topic has been performed by Ryan and Zarowin (2003), who derive a model to capture particularly the effects of lags and asymmetry. Ryan and Zarowin (2003) define lags as the tendency of earnings to reflect news later than stock prices and asymmetry is defined as the asymmetric fashion in which earnings reflect good and bad news.

Ryan and Zarowin (2003) hypothesize that both asymmetry and lags have increased over the years. Asymmetry has increased due to increased conservatism, while lags have increased due to increasing uncertainty, competitive dynamics and an increase in intangibles (Ryan and Zarowin, 2003). The increase in lags is thus very much related to an increase in business change. Ryan and Zarowin (2003) find evidence for both these hypotheses. Asymmetry has increased over the years and specifically for accruals, while not so much for cash flows. Ryan and Zarowin (2003) conclude from this that the change in asymmetry is mostly attributable to accounting reasons and not economic changes.

Additional evidence on an increase in accounting conservatism is put forward by Kim and Kross (2005). As was explained in chapter three, Kim and Kross (2005) investigate the ability of earnings to predict future cash flows and found that this has been increasing. As an explanation for this result, Kim and Kross (2005) evaluate increased conservatism. Kim and Kross (2005) argue the following. If bad news is incorporated into financial statements more promptly and if cash flows are affected by bad news, then a more timely recognition of bad news increases the ability of financial statements to predict cash flows.

As an indicator of accounting conservatism, Kim and Kross (2005) rely on work of Givoly and Hayn (2000). One indicator for conservatism by Givoly and Hayn (2000) is the level of accumulated nonoperating accruals, e.g. loss provisions on inventory, assets write-downs and restructuring charges. Based on this measure, Kim and Kross (2005) find a link between increased conservatism and increased value relevance. To summarize, the effect of conservatism is twofold. First of all, it decreases the association between accounting variables and share prices or returns because of asymmetry (Ryan and Zarowin, 2003), but it increases the ability of earnings to forecast future cash flows, because of the timely recognition of bad news (Kim and Kross, 2005).

4.3.8 Standard setting bodies and IFRS

The effect of standard setting bodies on value relevance in the U.S has been studied by Ely and Waymire (1999). Ely and Waymire (1999) investigate value relevance of earnings and earnings and book values combined for three standard setting bodies in the U.S.; the Committee on Accounting Procedure (1939-1959), the Accounting Principles Board (1960-1973) and the Financial Accounting Standards Board (1974-present). The focus on earnings is intentional, since income measurement and disclosure has been a primary focus during the period under investigation. Ely and Waymire (1999) only find limited support for changes in value relevance subsequently to the empowerment of standard setting bodies.

Standard setting bodies in the U.S. are however not of critical importance in this research, since it is concerned with value relevance in Europe. Conversely, the mandatory transition from local GAAP to IFRS in 2005 for EU based companies could have had a significant impact on value relevance. For example, Barth et al. (2008) find that value relevance, measured by the association between accounting variables and share prices or returns, increases for firms adopting IAS, the predecessor of IFRS. This result is based on a global sample of firms in the period 1994-2003.

The findings of Barth et al. (2008) are further confirmed by Capkun et al. (2008), who find that value relevance increases for European firms after the mandatory adoption of IFRS. To establish these results, Capkun et al. (2008) specifically investigate the transition year 2004-2005, in which firms published both statements under local GAAP and IFRS. Related, Daske et al. (2008) finds that both market liquidity and equity value increases after the mandatory adoption of IFRS. Finally, Clarkson et al. (2011) report mixed results on the influence of IFRS on value relevance. Dependent on the model, they report either a decline or no change in value relevance subsequent to the introduction of IFRS. Clarkson et al. (2011) do note that the introduction of IFRS has increased comparability of financial statements.

4.3.9 Non-information-based trading

Dontoh et al. (2007) propose an alternative view on the cause of declining value relevance. Their main hypothesis is that an increase in non-information-based trading decreases the suitability of share prices or returns as a measure of respectively intrinsic firm value or performance. Examples of non-information-based trading are global wealth transfers. Dontoh et al. (2007) show, in an equilibrium framework, that this kind of trading might lead stock prices away from their intrinsic value. They furthermore notice that accounting information is independent of non-information-based trading and the corresponding noise in stock prices. These factors combined with an increase in non-information-based trading could induce a decrease in value relevance (Dontoh et al., 2007). This also implies that share prices have become a progressively worse predictor of future performance over time.

This last implication of non-information-based trading is also shown empirically by Dontoh et al. (2007). More precisely, they find a decrease in the predictive content of earnings, but an even larger decrease in the predictive content of share prices.[20] Based on their results, Dontoh et al. (2007, 38) conclude with: "*Our findings cast doubt on the appropriateness of using stock prices or returns as benchmark for evaluating the information content of accounting numbers in value-relevance studies.*" A deep review of this remark is deemed beyond the boundaries of this paper, but it should be noted that such findings are contrary to the efficient market hypothesis (Fama, 1970).[21]

4.4 Summary

This chapter presents an overview of empirical findings on changing value relevance. One of the main conclusions of section 4.2 is that value relevance measured under Interpretation 4 has likely declined in the U.S. The amount of evidence and empirical work for value relevance under Interpretation 2 is limited, but the extensive study of Kim and Kross (2005) indicates that value relevance has most likely increased under this interpretation. This chapter also provides a strong justification for the thorough discussion of value relevance metrics in chapter three; it has been shown that results can be sensitive to the choice of metric. In this respect the study of Brown et al. (1999) is informative; these authors show that the conclusions of value relevance studies can be altered by incorporating scale effects. The main studies focussing on changes in value relevance are further summarized in table 1.

Several factors have been identified over the years that might influence the value relevance of accounting information. A returning theme in these possible causes is the

[20] As was stated in section 4.2.2, the predictive content of earnings and prices is measured as their ability to capture future firm's dividend or earnings. For a complete explanation see Dontoh et al. (2007).

[21] As a note to this citation, the information content of earnings or prices is defined by Dontoh et al. (2007, 5) as the extent to which these metrics capture fundamental firm value. Note that this definition differs from that applied by Landsman and Maydew (2002).

change in the business environment; e.g. the increasing importance of high-technology firms in the U.S. (Dontoh et al., 2007), which would have led to a decrease in value relevance. There exists considerable support for this hypothesis. Related to this cause are increased investments in intangibles which are not adequately captured by accounting standards in the U.S. Alternative explanations include increased conservatism, changes in the average firm size across the population and an increase in non-information-based trading. Where appropriate, table 1 links these causes to the main studies discussed in this chapter. Finally, of particular interest for this study is the mandatory transition of local GAAP to IFRS, which might have significantly influenced value relevance. Based on the discussed research, the sign of this influence is however not clear.

Table 1

Overview of value relevance literature on U.S. samples

Author	*Methodology*	*Sample period*	*Conclusion*	*Influencing factors*
Collis et al. (1997)	The coefficient of determination metric estimated for the adopted accounting based valuation model without any corrections for scale effects.[a]	1953-1993	The value relevance of earnings has decreased The value relevance of book values and of book values and earning combined has increased	Increase in investments in intangibles Increase in nonrecurring items Increase in the number of firms reporting negative earnings Decrease in average firm size
Francis and Schipper (1999)	The coefficient of determination metric for an earnings–returns relation, a market value-book values relation and for the adopted accounting based valuation model. Only the earnings-returns relation is deflated by market value of equity. The portfolio metric is also used.[b]	1952-1994	The value relevance of earnings has decreased The value of book values and of book values and earnings combined has increased	An increase in high technology firms
Lev and Zarowin (1999)	The coefficient of determination metric for the adopted accounting based valuation model and the coefficient of determination and coefficients metric for an earnings –returns relation and for a cash flows-returns relation. These last two regressions are deflated by market value of equity.[c]	1977-1996	The value relevance of earnings, book values and of cash flows has decreased	An increase in the rate of business change
Ely and Waymire (1999)	The coefficient of determination metric estimated for an earnings-returns relation and estimated for the adopted accounting based valuation model, the former is deflated by market value of equity while the latter is uncorrected for scale effects.[d]	1927-1993	The value relevance of earnings is unchanged over the sample period The value relevance of earnings and book values combined has increased	Institutional changes, although these only weakly influence value relevance

Table 1 (cont'd)

Overview of value relevance literature on U.S. samples

Author	*Methodology*	*Sample period*	*Conclusion*	*Influencing factors*
Chang (1999)	The variance of the logarithm metric and the portfolio metric.[e]	1953-1993	▪ The value relevance of both earnings and book values has decreased	▪ Some evidence that the creation of the FASB coincides with a decrease in value relevance
Brown et al. (1999)	The coefficient of determination metric estimated for the adopted accounting based valuation model and corrected for scale effects by using proxies and deflating variables by market value of equity.[f]	1952-1994	▪ The value relevance of earnings and of earnings and book values combined has decreased. ▪ The value relevance of book values has increased.	▪ Not applicable
Gu (2007)	The residual dispersion metric.[g]	1977-1996	▪ The value relevance of both earnings and book values has decreased	▪ No specific factors tested, but a structural break in value relevance seems to be present in the early 1970s
Kim and Kross (2005)	The coefficient of determination metric estimated by regressing cash flows against lagged earnings and cash flows. Corrected for scale effects by deflating by total assets and by ranking variables.[h]	1927-1993	▪ The value relevance of earnings to predict future cash flows has increased	▪ An increase in accounting conservatism ▪ Inefficient capital markets

[a] See section 3.2.
[b] See section 3.2 for the coefficient of determination metric. The estimated earnings-returns relation is similar to that estimated by Ely and Waymire (1999). The estimated market values-book values relation is not described in this paper, see Francis and Schipper (1999, 332) for a description. The portfolio metric employed by Francis and Schipper is described in section 3.5.
[c] See section 3.2 and 3.4.
[d] The methodology applied for the coefficient of determination metric for the accounting based valuation model is comparable to that applied by Collins et al. (1997), see section 3.2. The coefficient of determination metric for an earnings-returns relation is not described particularly in this paper, but it is a straightforward adoption, estimated by using (3.33) instead of (3.1).
[e] See sections 3.3.4 and 3.5 and note that Chang (1999) uses the magnitude and sign of earnings, as well as the return on equity as selection criteria for the portfolio metric.
[f] See sections 3.2 and 3.3.2, specifically equation (3.18).
[g] See section 3.3.5.
[h] See section 3.6 for a description of the methodology of Kim and Kross (2005). The procedure of ranking variables to correct for scale effects as applied by Kim and Kross (2005) has not been described in this paper, see Kim and Kross (2005, 15) for more information.

5. Country-specific factors

5.1 Introduction

The main focus of this research is on differences across time in the value relevance of accounting information for a cross-country European sample, i.e. consisting of France, Germany, The Netherlands and the United Kingdom. This cross-country research design necessitates the implementation of some country specific controls (Ruland et al., 2007). To provide some background for the discussions in the remainder of this paper, this chapter presents a short overview of researches on international differences in accounting systems. Furthermore, the influence of country-specific factors on accounting quality in general and on value relevance specifically is also discussed.

An important distinction to be made at the outset of this chapter is that between accounting practices and accounting standards. Accounting standards are a subset of the accounting practices of a country, which also encompass, amongst others, the quality of the audit profession and the legal enforcement of accounting standards. This complete set of practices is commonly referred to as the accounting system of a country; see for example Nobes and Parker (2006) and Ruland et al. (2007). The importance of this distinction between practices and standards is amplified by research of e.g. Ball et al. (2003) and Leuz et al. (2003), which will be discussed later in this chapter.

The subsequent sections of this chapter are organised as follows. In section 5.2 some frequently cited factors linked to accounting practices in countries are given. These country-specific factors are related to the sample countries in section 5.3. Section 5.4 offers an overview of empirical findings on the influence of country specific factors on accounting quality, while maintaining a focus on value relevance. Finally, section 5.5 concludes.

5.2 Factors influencing national accounting practices

5.2.1 Introduction

It has long been recognized that the accounting system of a country is shaped by institutional and environmental factors (Mueller et al., 1994). In their recent work, Ruland et al. (2007) summarize a number of variables that seem to be correlated to the information produced by the accounting systems of countries. These variables are identified based on a study of international accounting literature. The following six country-specific characteristics are distinguished (Ruland et al., 2007, 101-102):

1. Legal origin, most notably common law versus code law
2. Shareholder protection
3. Accounting-related aspects
4. Information environment
5. Market-related characteristics
6. Political factors

In the following paragraphs these characteristics will be elaborated on.

5.2.2 Legal origin

The first characteristic, the legal origin of a country is often cited to influence accounting standards and practices, see e.g. Choi and Mueller (1992), Mueller et al. (1994), Ball et al. (2000) and Nobes and Parker (2006). The most common distinction between legal systems is that of code law and common law countries. The difference between these categories is elegantly put by Mueller et al. (1994, 6). These authors argue that laws in common law countries are series of *"thou shalt nots"* compared to *"thou shalts"* in code law countries. With this it is meant that laws in code law countries determine a minimum standard of expected behaviour, while laws in common law countries set up limits of legal behaviour. The former category of law is highly descriptive and detailed, while the latter leaves room for professional judgement. Examples of code law countries are Germany, France and the Netherlands. England is seen as the originator of common law and the United Kingdom and the United States, which historically is heavily influenced by the United Kingdom, therefore have common laws.

Nobes and Parker (2006) note that accounting standards as such do not have to be determined by the law system of a country. A primary illustration of this is the adoption of IFRS by both common and code law countries. However, the nature of accounting regulation obviously is correlated to the system of law. More generally, the legal origin of a country is also correlated to the other characteristics distinguished by Ruland et al. (2007) that influence accounting practices. To provide some examples; Ball et al. (2000) use the classification of common and code law countries as a proxy for the political influence on accounting practices. Mueller et al. (1994) argue that financial accounting in code law countries serves mostly tax collection purposes. Furthermore, La Porta et al. (1997) find a connection between legal origin and providers of finance and Hope (2003) uses this dichotomous classification as a proxy to determine the strength of legal enforcement in countries. These country-specific characteristics are classified by Ruland et al. (2007) respectively as political factors, accounting-related aspects, market-related characteristics and shareholder protection. To summarize, legal origin might not have a direct influence on the accounting practices of a country, but it is related to the development of these practices and can serve as a proxy for other country-specific characteristics.[22]

5.2.3 Shareholder protection

Shareholder protection is the second characteristic identified by Ruland et al. (2007) that is linked to accounting practices in countries. Shareholder protection includes the quality of legal enforcement, the efficiency of the judicial system, the system of shareholder voting right and the strength of insider trading laws. Leuz et al. (2003) propose that managers in countries with strong shareholder protection have fewer incentives to conceal true firm performance and are thus less likely to manage earnings. By altering the incentives of managers, shareholder protection can increase accounting quality. This argument has also been put forward by Hung (2001) and Ball et al. (2003). Shareholder protection however again seems to be related to the other characteristics. La Porta et al. (1997) find that strong shareholder protection increases the size of equity markets and Hung (2001) uses the classification common versus code law as a proxy for shareholder protection.

5.2.4 Accounting related aspects

Accounting related aspects, the third characteristic, clearly are related to the information produced by different accounting systems. These factors are mostly researched by Ali and Hwang (2000). Several aspects of accounting are considered by these authors, namely the conformity between tax and financial reporting, whether private or public bodies set standards, the orientation of accounting, classified as either “British-American” or “Continental”, and the amount of spending on auditing services. The classification of accounting cluster requires some additional elaboration. These accounting clusters are identified by Mueller et al. (1994) based on overall similarities in accounting practices. The differing factor is that accounting practices in the British-American cluster are mostly geared towards the decision-making needs of investors, while in the Continental cluster the

[22] An example of a country where there exists a separation between legal origin and accounting system is the Netherlands. Even though it has a code law system (Nobes and Parker, 2006, 28), its accounting system more closely resembles that of common law countries, see Arce and Mora (2002, 577).

purpose of financial accounting is primarily to compute income taxes or to demonstrate compliance with regulation. Ali and Hwang (2000) argue that value relevance should be higher for countries where financial reports mostly serve public investor's information needs. This recognises that financial statement information adapts to the requirements of users (Nobes and Parker, 2006) and that individual investors need value relevant information to mitigate moral hazard problems. If financial reports are also used for e.g. tax-calculating or regulatory purposes this decreases value relevance. The other accounting related aspects would appear to be correlated to this distinction and this is confirmed by Ali and Hwang (2000) through a principal factor analysis, which shows that one underlying factor accounts for most of the variation in these variables.

5.2.5 Information environment

The fourth characteristic is the information environment of a country. Ruland et al. (2007) name two variables belonging to this characteristic, namely a disclosure index used by Leuz et al. (2003) and the impact of networking and private debt on the demand for public disclosures used by Ball et al. (2003). The disclosure index by Leuz et al. (2003) measures the inclusion or omission of 90 items in financial reports for separate countries, and is based on La Porta et al. (1998). Interestingly, Leuz et al. (2003) find that this disclosure index is significantly correlated at the 5% level to outside investor rights, legal enforcement, importance of stock markets and ownership concentration. This finding again highlights the interrelatedness of the factors identified by Ruland et al. (2007).

Ball et al. (2003) consider the demand for public disclosure and its effect on accounting information. In their research, Ball et al. (2003) investigate accounting quality in East Asia, where a large amount of family owned businesses exists and personal networks are of comparative importance. These authors hypothesize and find that personal networks reduce the demand for public disclosure. The reason for this is that information asymmetries between providers of capital and firms are resolved through these private networks instead of through public disclosures. A similar argument applies for countries with relatively large quantities of private debt. Finally, Ball et al. (2003) note that the countries under investigation have high quality accounting standards, but, because of the little demand for public disclosure, this does not necessarily imply high quality accounting information.

5.2.6 Market related characteristics

The demand for public disclosure naturally introduces the fifth characteristic, which are market related characteristics and noteworthy in this respect are the providers of finance in an economy. A general distinction is made between market-oriented financial systems and bank-oriented financial systems; see for example Mueller et al. (1994), Ali and Hwang (2000) and Nobes and Parker (2006), although more detailed classifications also exist (Nobes and Parker, 2006). In market-oriented financial systems, equity markets are of higher importance and equity financing through private shareholders is more common. Historically, ownership of businesses is more dispersed in market-oriented financial systems (Nobes and Parker, 2006). This contrasts to bank-oriented financial systems, where a few large banks provide most of the capital to businesses (Mueller et al., 1994). The United Kingdom has a market-oriented financial system, while France and Germany have bank-oriented financial systems. The Netherlands is somewhat peculiar in this respect, but this issue will be discussed in the following section.

The investors in these two financial systems have separate information needs, and the rationale behind this is comparable to the hypothesis of Ball et al. (2003) on private networks. Small private shareholders and creditors have less access to insiders' information and the cost of personally obtaining this information is disproportionally high for each investor. For these reasons, they require relevant financial reports to monitor and value firms. On the other hand, banks holding large stakes of businesses can efficiently resolve their information needs through private networks and direct visits, reducing the demand for public financial reports (Mueller et al., 1994).[23] This shifts the main purpose of financial reports to that of tax collection and protection of creditors, inducing more conservative and prudent information (Nobes and Parker, 2006). A similar classification therefore is that between "outsider/equity" (market-oriented) and "insider/credit" (bank-oriented) economies and accounting practices, see Leuz et al. (2003) and Nobes and Parker (2006). A priori, one would expect market-oriented financial systems to prepare higher quality and more relevant information, which is confirmed by research of Ali and Hwang (2000) and Leuz et al. (2003).

[23] Nobes and Parker (2006, 31-32) note that the clear distinction between market-oriented and bank-oriented financial systems appears to be decreasing over recent years. For one, small private shareholders are being replaced by large institutional investors in market-oriented financial systems. Furthermore, private investors are becoming increasingly important in bank-oriented systems such as in France and Germany. Despite this, Nobes and Parker (2006) argue that the two way distinction seems to be intact.

5.2.6 Political factors

The sixth and final characteristic is political factors, including the political influence on financial reporting, the influence of family control and the level of corruption (Ruland et al., 2007). That political factors influence financial reporting is not a new thought, as is apparent in work of Watts and Zimmerman (1978, 1986). These authors argue that there are both political costs to reporting large losses as well as to reporting large profits. This provides managers with incentives to smooth earnings and bears on the information in financial reports. Ball et al. (2003) argue that managers in different countries face different political costs, leading to international differences in accounting practices. The influence of family control on accounting practices is similar to that of private networks mentioned earlier (Ball et al., 2003), while the level of corruption is used by Leuz et al. (2003) as a proxy to determine the strength of legal enforcement of accounting standards.

5.2.6 Summary

From the discussion above it is clear that the six country-specific characteristics are strongly interrelated. Ruland et al. (2007, 102) note that: *"These characteristics are not mutually exclusive, and there is no overriding consensus regarding the overall relation among these characteristics. (i.e., are these characteristics substitutes or complements?)"* Despite this, Nobes and Parker (2006) suggest that the differentiation between market-oriented and bank-oriented financial system is the key cause of international differences in accounting systems. Some previously mentioned researches lend support to the notion that this differentiation encompasses most other characteristics.[24] To recite, La Porta et al. (1997) and Leuz et al. (2003) find that capital markets are larger in countries with a strong legal environment, i.e. those with strong shareholder protection and common law. Furthermore, findings of Ali and Hwang (2000) show a strong correlation between the types of financial system, the amount of spending on auditing services, whether private or public bodies set standards, the orientation of accounting and the conformity between tax and financial reporting. A further evaluation of this hypothesis of Nobes and Parker (2006) is deemed beyond the boundaries of this paper, but on an intuitive level it appears plausible that the purpose of financial reports and the demand for public disclosure, characterized by different financial systems, are the main drivers of international differences in accounting practices.[25]

---

[24] It is not the intention to posit causality, but merely to show that the type of financial system is strongly correlated to the other country-specific characteristics mentioned in the literature.

[25] For a further discussion of this point, see Nobes (1998) and Nobes and Parker (2006, 29-32). This hypothesis does appear to be widely acknowledged, see for example Joos and Lang (1994)

5.3 Characteristics of the sample countries

This section relates some of the country-specific factors distinguished by Ruland et al. (2007) to the sample countries of this research.[26] As was already mentioned in section 5.2, of the four countries investigated in this research, only the United Kingdom has a common law legal system. France, Germany and the Netherlands all have code law legal systems (Nobes and Parker, 2006). Contrary to this, Ball et al. (2000) classify the Netherlands as common law, but this classification is based on the accounting model of this country.

To characterize countries on shareholder protection is a somewhat more elaborate job. Leuz et al. (2003) provide some guidance. These authors identify three clusters of countries based on the size of equity markets, ownership concentration, legal enforcement and shareholder protection. The first cluster, to which the United Kingdom belongs, has large equity markets, low ownership concentration, high levels of disclosure and strong shareholder protection. The second cluster, which includes France, Germany and the Netherlands, is characterized by smaller equity markets, higher ownership concentration, weaker shareholder protection, weaker legal enforcement and lower disclosure levels. The third cluster is basically similar to the second cluster, but with even weaker legal enforcement.

Ali and Hwang (2000) evaluate the accounting related aspects of France, Germany, the Netherlands and the United Kingdom as well as those of some additional countries. The Netherlands and the United Kingdom are very similar based on these aspects. They both have public and private standard setting bodies and a low alignment between tax and financial reporting. Furthermore, these countries spend a relatively large amount on auditing services and belong to the British-American accounting cluster. France and Germany are the opposite of this; only the government sets accounting standards in these countries, a high correspondence between fiscal and financial accounting exists, spending on auditing services are low and both countries belong to the Continental accounting cluster.

The final distinction made here is that between market-oriented and bank-oriented financial systems. As was stated before, the United Kingdom has a market-oriented financial system, while Germany and France have bank-oriented financial systems (Nobes and Parker, 2006). In general, see e.g. Mueller et al. (1994), Ball et al. (2000), Arce and Mora (2002) and Nobes and Parker (2006), accounting practices in the Netherlands are regarded

---

[26] The descriptive variables in this section are mostly on a macro-level. For a more micro-approach and detailed description, although a bit outdated, of the accounting practices in these countries, see Choi and Mueller (1992).

as similar to those of countries with market-oriented financial systems. The main purpose of financial reporting is geared towards the decision-making process of investors (Mueller et al. 1994). Nevertheless, the cluster analysis of Leuz et al. (2003) puts the Netherlands in the second cluster, which they describe as an "insider economy" cluster. Notwithstanding these findings, based on the more qualitative arguments of the researchers mentioned above the Netherlands is regarded as having a market-oriented financial system in the remainder of this paper. Or to be precise, its accounting practices are regarded as similar to those of a market-oriented country.

The classifications and distinctions described here are becoming somewhat less relevant following the mandatory introduction of IFRS for all listed firms in the European Union in 2005. The word somewhat should be emphasized, because there are still reasons to believe that the accounting practices in the sample countries will differ. First of all, the introduction of IFRS merely concerns accounting standards and not practices as a whole. That high quality accounting standards do not necessarily imply high quality accounting information is shown by Ball et al. (2003). Secondly, and related to the first point, the variables described in section 5.2 and listed here are institutional variables that have not changed after the introduction of IFRS. Their influence on reporting practices might thus still be present, for example through the incentives of managers. Finally, concerning the accounting standards themselves, Nobes (2008) argues that national IFRS "standards" are arising. IFRS leaves a considerable amount of choice in standards to managers and these choices are likely to be influenced by previous national accounting standards.

To conclude, significant differences in accounting practices existed between France, Germany, the Netherlands and the United Kingdom and these differences might be pervasive after the introduction of IFRS. A two-way split of these countries seems appropriate with respectively the market-oriented/outsiders focused practices of the Netherlands and the United Kingdom and the bank-oriented/insiders practices of France and Germany. In general, the former group is found to produce more value relevant information, a topic more deeply explored in the following section.

5.4 Empirical results on international differences in accounting quality

In this section a closer look will be taken at some of the research presented in section 5.2, as well as at some additional research. The focus of this discussion is on empirical results and methodological aspects. Where possible, results on international differences in value relevance between France, Germany, the Netherlands and the United Kingdom will be highlighted.

A logical starting point is the work of Alford et al. (1993), who clearly document international differences in value relevance. Their sample consists of seventeen countries, including the sample countries of this paper, while the sample period of Alford et al. (1993) is 1983-1990. The objective of these researchers is to compare the value relevance and timeliness of earnings of sixteen of these countries to that of the United States. To measure value relevance, Alford et al. (1993) use both the portfolio metric based on earnings changes and the coefficient of determination metric for an earnings-returns relation.[27] One of their conclusions is that earnings in France, the Netherlands and the United Kingdom are more value relevant than those in the United States while earnings in Germany are comparatively less value relevant. This conclusion is based on an analysis of the differences in the portfolio metric between country specific samples and a matched sample from the United States.[28]

A comparison of the coefficient of determination metric of these countries against that of the United States only partially underlines these findings. In this test the United Kingdom is the single country of these four that produces significantly more value relevant information than the United States. It should be noted however, that the application of the coefficient of determination metric by Alford et al. (1993), despite using scaled variables, is susceptible to some of the critiques mentioned in section 3.3. Also, an additional econometrical issue in this research design is that of serial correlation. Alford et al. (1993) use fifteen month return windows to estimate the coefficient of determination. Because for each country all observations are pooled across years, a three month overlap exists between data points and hence the error terms are serially correlated. Several ways to correct for this type of serial correlation exist; the method used by Alford et al. (1993) is generalized least squares.[29] Finally, an underlying assumption of this research design is that share prices reflect information equally efficient across their sample countries. This assumption can be scrutinized, but Alford et al. (1993) argue that tests of market efficiency for their sample countries generally find these markets to be efficient.

Some counterintuitive and contrary results are found by Joos and Lang (1994). Joos and Lang (1994) evaluate differences in value relevance among France, Germany and the United Kingdom for the period 1982-1990. For this they use the coefficient of determination metric estimated for an earnings-returns relation and for the accounting based valuation

[27] In general, when reference is made to an earnings-returns relation, the specification of (3.33) is meant, i.e. $R_{i,t} = \alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}\Delta X_{i,t} + \varepsilon_{i,t}$. In this $R_{i,t}$ are the share returns of firm i in year t, $X_{i,t}$ are the earnings of firm i in year t and $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$ is the change in earnings.

[28] To recite, the portfolio metric is defined as R_t^* / R_t^H, where R_t^* is the fifteen month market adjusted return of a portfolio based on perfect foreknowledge of accounting information, in the case of Alford et al. (1993) earnings changes, while R_t^H is the market adjusted return on a portfolio based on perfect knowledge of returns. If this metric is comparatively higher in one country compared to another, one could conclude that value relevance is also higher. See section 3.5 for further details.

[29] See Greene (2008) for additional information on generalized least squares.

model as specified in (3.1).[30] The expectation of these authors is that value relevance is the highest in the United Kingdom, followed by France and then Germany. In general, Joos and Lang (1994) find that value relevance is the highest in France and the lowest in the United Kingdom. Despite that this is contrary to their expectations; Joos and Lang (1994) do not provide a clear explanation for these results. A possible explanation might be model misspecification and econometric issues as presented in section 3.3. Especially their regression of the accounting based valuation model is not corrected for these earlier identified problems. An additional finding of Joos and Lang (1994) is that the introduction of legislation by the European Union, which was intended for convergence of reporting standards, has had little effect on value relevance.

Joos (1997) also focuses on the comparative value relevance of accounting information in France, Germany and the United Kingdom. After considering the institutional characteristics of these countries, he comes to three separate hypotheses. The first is that earnings are more value relevant than book values in the United Kingdom, while the converse should hold for France and Germany. This because of the focus on the information needs of shareholders in the United Kingdom and the focus on the information needs of creditors in France and Germany. The second hypothesis by Joos (1997) is that the combined value relevance of earnings and book values will be higher in the United Kingdom than in France and Germany. The first hypothesis is tested by comparing the coefficient of determination metric for an earnings-price relation to that of a book values-price relation within each country, whereas the second hypothesis is tested by comparing the coefficient of determination metric for the empirical accounting based valuation model across countries.[31] The final hypothesis of Joos (1997) concerns conservatism and he expects conservatism of both earnings and book values to be higher in France and Germany than in the United Kingdom. The sample used by Joos (1997) consists of the years 1982-1993.

Test results only partially confirm these hypotheses. Joos (1997) is able to establish that earnings are more value relevant than book values in the United Kingdom but he does not find the reverse relation in France and Germany. Similar to Joos and Lang (1994), the second hypothesis is rejected as the combined value relevance of earnings and book values is found to be higher in France than in the United Kingdom. Finally, Joos (1997) confirms

---

[30] The accounting based valuation model of (3.1) was specified as $P_{i,t} = \alpha_{0,t} + \alpha_{1,t}X_{i,t} + \alpha_{2,t}BV_{i,t} + \varepsilon_{i,t}$, where $P_{i,t}$ is the price of a share of firm i, $X_{i,t}$ is the earnings per share of firm i in year t, $BV_{i,t}$ is the book value of equity per share of firm i at the end of year t and $\varepsilon_{i,t}$ represents all other value relevant information of firm i for year t not included in either $BV_{i,t}$ or $X_{i,t}$

[31] As should be clear from section 3.3, comparing the coefficient of determination from two different regressions is not a trivial matter especially not if these are non-nested as is the case in testing hypothesis 1. To do so, Joos (1997) uses the Vuong likelihoodratio test, see Vuong (1989) for additional details.

that book values are more conservative in France and Germany than in the United Kingdom, but this result is not found to hold true for earnings.

In a related research, King and Langli (1998) use the same methodology as Joos (1997) to contrast international differences in value relevance across Germany, Norway and the United Kingdom for the period 1982-1996. For similar reasons as Joos (1997), these authors expect the combined value relevance of earnings and book values to be lower in Germany than in their other sample countries. Furthermore, they also expect book values to be more value relevant than earnings in Germany and vice versa for Norway and the United Kingdom. The results of King and Langli (1998) do not reject these hypotheses.

A more extensive research is undertaken by Arce and Mora (2002), who focus on differences in value relevance among eight European countries. Again, the research design and hypotheses of this study are much like those of Joos (1997). The larger number of countries examined, these include France, Germany and the United Kingdom, but also Belgium, Italy, the Netherlands, Spain and Switzerland, arguably allows for a better comparison between different types of accounting systems. The division of these countries in accounting systems is familiar; those of the United Kingdom and the Netherlands are classified as investor-oriented British-American accounting systems and the remainder is classified as having credit-oriented Continental accounting systems.[32] Arce and Mora (2002) use the period 1990-1998 as their sample. A main finding of these authors is that earnings are more value relevant than book values in the British-American accounting systems and that the opposite relation holds true for Continental accounting systems. An additional result is that both earnings and book values have incremental explanatory power to explain market prices, the exception being Germany and Spain where book values convey all information for share prices. Finally, despite using an extended number of countries, Arce and Mora (2002) are unable to establish that the combined value relevance of earnings and book values is higher for British-American accounting systems than for Continental accounting systems.

As was described in section 5.2, Ali and Hwang (2000) relate five different country-specific factors to international differences in value relevance. These are: the distinction between market-oriented and bank-oriented financial systems, whether public or private bodies set accounting standards, the orientation of accounting classified as either Continental or British-American, the conformity between tax and financial reporting and the amount of spending on auditing services. Sixteen countries are investigated by Ali and Hwang (2000) and these countries are not limited to Europe but also include for example

[32] To be precise, Arce and Mora (2002) distinguish between the Anglo-Saxon and Continental accounting systems, but for the sake of consistency the previously introduced names are used.

the United States and Singapore. The used data is from the period 1986-1995. Ali and Hwang (2000) measure value relevance by both the portfolio metric and the coefficient of determination metric. For this latter metric the common regressions for an earnings-returns relation and the accounting based valuation model are estimated plus an alternative specification wherein returns are regressed against accruals. Notable about the research design of Ali and Hwang (2000) are also their proxies for the differentiation between market-oriented and bank-oriented financial systems. Instead of using a dichotomous classification, Ali and Hwang (2000) use both the ratio of domestic firms to population and a specific debt to asset ratio to capture the degree of market or bank orientation.

A first important result of Ali and Hwang (2000) is the strong interdependence of these five country-specific factors. Using a principal component analysis, Ali and Hwang (2000) show that one single underlying factor accounts for almost 85% of the variation in these variables. Furthermore, all five of these factors are found to influence the value relevance of accounting information in the expected direction, i.e. value relevance is higher in countries with market-oriented financial systems, where private bodies set standards, which have a British-American accounting system, where a low alignment between tax and financial reporting exists and where auditing spending is high. These results are robust for the different metrics for value relevance.

Ball et al. (2000) do not specifically investigate the value relevance of accounting data, but consider the effect of institutional factors on the timeliness and conservatism of earnings. The main institutional variable studied by Ball et al. (2000) is the political influence on accounting standards for which they use the system of law as a proxy. Ball et al. (2000) regard common law systems as those with low political influence on accounting practices, while they argue that in code law countries a high political influence on accounting practices exists. In common law countries accounting practices are instead formed by the market demand for disclosures, which Ball et al. (2000) hypothesize leads to more timely and less conservative financial reporting. The sample used by Ball et al. (2000) is extensive and encompasses accounting data from 25 countries, including the sample countries of this paper, for the period 1985-1995. The obtained results are consistent with their hypothesis.

However, Ball et al. (2000) point out that there is a limitation to this research. They caution that the classification of code law and common law countries is not homogeneous; put differently this means that in none of the countries financial reporting is determined in merely a market or regulated manner. This issue is partly resolved in research of Leuz et al. (2003) who use a cluster analysis to identify three types of country cluster. These clusters were presented in section 5.3, but to recite the first cluster consist of countries with an outsider economy, i.e. focused on equity investors, and also with strong legal systems. The second and third clusters are characterized as insider economies with the distinction that the third cluster has a significantly weaker legal system than the second. Leuz et al. (2003) argue that this classification conveys significantly more information about the accounting practices of countries compared to the dichotomous classification of Ball et al. (2000). Furthermore, using a sample of 31 countries and data from 1990-1999, Leuz et al. (2003) find that earnings management is related to these three clusters, where earnings management is the least pervasive in cluster 1 and the most pervasive in cluster 3. Since the United Kingdom is classified in cluster 1 and France, Germany and the Netherlands in cluster 2, this implies that accounting information in the United Kingdom would be the least susceptible to earnings management. The main argument Leuz et al. (2003) propose for their findings are differences in the incentives of managers among clusters due to either strong or weak shareholder protection.

The final research to be discussed is that of Ball et al. (2003) who clearly show that the quality of accounting information is more than just a function of standards. Ball et al. (2003) examine the effect of incentives on the properties of accounting information in four East Asian countries over the period 1984-1996. These four countries are historically heavily influenced by the United Kingdom and have similar, generally regarded as high quality, accounting standards. Despite this, the incentives faced by managers are much more alike to those in code-law countries, e.g. through political influences on financial reporting, family control of businesses, privately held debt and weak shareholder protection. As a result, timely recognition of economic income is not part of the accounting practices of these countries, reducing the quality of financial reporting (Ball et al., 2003).

5.5 Summary

In this chapter a number of country-specific characteristics that bear on the information produced by national accounting practices are distinguished. The used framework for this is provided by Ruland et al. (2007), who identify the following six factors: legal origin, shareholder protection, accounting related aspects, information environment, market-related characteristics and political factors. These six factors are however strongly interrelated, for example the interconnection between accounting related aspects and market-related characteristics is that market-oriented financial systems generally have a low alignment between tax and financial reporting, while for bank-oriented financial systems the opposite holds (Ali and Hwang, 2000). Numerous other examples also spring to mind. Of these six factors the distinction between market-oriented and bank-oriented financial systems is arguably the most important (Nobes and Parker, 2006). In market-oriented financial systems accounting practices are mostly geared towards the information needs of investors while in bank-oriented financial systems accounting practices are developed to produce relevant information for creditors and governments. This distinction therefore directly captures the purpose of financial reports in different countries and the corresponding demand for public disclosure. A priori, one would expect value relevance to be higher in market-oriented financial systems than in bank-oriented financial systems.

This chapter also relates these six country-specific factors to the sample countries of this research, i.e. France, Germany, the Netherlands and the United Kingdom. Keeping close to the classification of financial systems, the Netherlands and the United Kingdom are classified in this paper as having accounting practices arising from market-oriented financial systems and France and Germany are classified as having accounting practices linked to bank-oriented financial systems. This classification is based on a study of previous literature.

Finally, some empirical evidence is provided on international differences in value relevance. The discussed literature generally expects value relevance to be higher in the United Kingdom and the Netherlands than in either France or Germany. The presented results are mixed on this issue, although this might be due to the applied methodology of some researchers which is open to the caveats discussed in section 3.3.

6. Hypotheses development and research design

6.1 Introduction

This study aims to give insight into the value relevance of financial statement information in Europe. More specifically, the main research question is *"Have there been any changes in value relevance in France, Germany, the Netherlands and the United Kingdom over time and what are plausible explanations for these possible changes"*. The literature review in the preceding chapters provides the background and tools to answer this question and in this chapter several hypotheses concerning value relevance are put forward, as well as the research methodology to test these hypotheses.

It is noted that, to the author's best knowledge, no empirical evidence is available on changes in value relevance over time for the sample countries under review. As such, the starting points for hypothesis development are the U.S. researches described in chapter four. However, it is questionable how closely the value relevance of European firms' financial statements resemble that of U.S. firms' financial statements. In section 6.2 this and other questions are explored more thoroughly and the different hypotheses concerning value relevance are put forward.

Section 6.3 provides a full description of the research methodology applied in this study. The main basis for the research methodology of this study is the work of Gu (2007) and his residual dispersion metric, as this appears to be the most comprehensive measure of value relevance. On the other hand, to test for robustness of results, several metrics of value relevance will be applied. Finally, in section 6.4 a summary of the previous sections is given.

6.2 Hypotheses development

In this section the different research hypotheses of this study are developed. To set the scope of this research, first the interpretation of value relevance should be defined. As is common in the literature, see chapter three and four, Interpretation 4 of value relevance is used. To recapitulate, Interpretation 4 of value relevance questions whether there exists a statistical association between financial information and share prices or returns. Ideally, Interpretation 2 would also have been used in this research, but this would have extended the scope beyond practicality and as such an investigation of changes in value relevance under Interpretation 2 is left for future research.

The first question this research tries to answer is whether value relevance has changed over time in France, Germany, The Netherlands and the United Kingdom. Since no evidence on changes in value relevance in these countries is available, it is hard to form any expectations. On the one hand, it seems plausible that similar forces are at play in France, Germany, The Netherlands and the United Kingdom as in the United States. As was described in chapter four, these consists among others of an increase in intangibles and technology based firms, the number of non-recurring items and negative earnings reported, the size of firms and an increase in the pace of business change more generally. If similar forces are at play in the sample countries of this research as in the United States, one would expect a decline in value relevance.

On the other hand, accounting standards in the sample countries have undergone a significant transformation in recent years, i.e. the introduction of IFRS, which is expected to influence value relevance as well. Most of the evidence on the influence of the introduction of IFRS on value relevance indicates that value relevance has increased afterwards; see e.g. Barth et al. (2008) and Capkun et al. (2008).

Given these opposing forces, ex ante no expectation is formed about the direction of change in value relevance. However, taking into consideration the dynamics described above, it seems most plausible that a change has occurred. As such the first research hypothesis is:

H1: Value relevance has changed significantly in the sample period in France, Germany, the Netherlands and the United Kingdom.

As is described in the following chapter, the initial sample period is 1988-2011, however, due to data limitations the first years are not taken into account.

After a possible change in value relevance has been established, the objective is to relate this change to a cause. In this study, two distinct factors are considered that might have influenced value relevance, i.e. an increase in the rate of business change and the introduction of IFRS.

Lev and Zarowin (1999) argue and find that the rate of business change in the United States has been increasing. If business change is not adequately reflected in financial statements, this will lead to a deterioration of value relevance, which is also found by Lev and Zarowin (1999). Here it is conjectured that a similar process has taken place in France, Germany, the Netherlands and the United Kingdom, leading to the following hypothesis:

H2: The rate of business change has been increasing in France, Germany, the Netherlands and the United Kingdom.

Following the logic of Lev and Zarowin (1999), if business change decreases the value relevance of financial statement information, it should be higher for firms experiencing no or little business change than for highly changing firms. This hypothesis is also confirmed by these authors. Furthermore, Lev and Zarowin (1999) find that the rate of decrease in value relevance is higher for firms experiencing a lot of business change. Based on these findings, the following two hypotheses will also be investigated:

H3: Value relevance is higher for firms experiencing little or no business change than for firms experiencing high amounts of business change.

H4: The rate of change in value relevance is smaller for firms experiencing a high amount of business change.

The choice of business change as one of the key variables under investigation is not arbitrary. Lev and Zarowin (1999) argue that other commonly mentioned factors for changes in value relevance, i.e. the increasing propensity of firms to report losses or non-recurring items, see chapter four and Collins et al. (1997), are merely manifestations of the failure of the accounting system to account for business change. Put differently, these factors do not cause a decline in value relevance, but are symptoms of the underlying increase in the rate of business change.

An additional possible explanation for changes in value relevance over time is the introduction of IFRS. One of the goals of the International Accounting Standards Board is to develop high quality accounting standards (Barth et al., 2008) and the introduction of IFRS is described by Capkun et al. (2008, 1) as *"the single largest event in recent history designed to increase the quality and consistency of reporting standards"*. Given the objective of the IASB and previous empirical findings of increasing accounting quality under IFRS, see e.g. Bart et al. (2007) and Capkun et al. (2008), the following hypothesis is put forward:

H5: Value relevance has increased in France, Germany, the Netherlands and the United Kingdom after the mandatory introduction of IFRS in 2005.

In chapter five, evidence is given that the quality of accounting practices, and value relevance more specifically, differs across countries. It was conjectured that the most important differentiating factor between countries is the distinction between market-oriented and bank-oriented financial systems, see also Nobes and parker (2006). Concerning the sample countries, the accounting practices of Germany and France were classified as stemming from bank-oriented financial systems, while those of the Netherlands and the United Kingdom were classified as stemming from market-oriented financial systems. Finally, given the focus of accounting practices on satisfying the information needs of investors in market-oriented financial systems, one would generally expect that value relevance is higher in market-oriented financial systems than in bank-oriented financial systems. This hypothesis is not always confirmed in empirical research, but nevertheless the following final hypothesis is put forward:

H6: Value relevance is higher in the United Kingdom and the Netherlands than in Germany and France.

6.3 Research design

6.3.1 Introduction

In this section the research design of this study is described. The research design of this study closely follows that of Gu (2007), i.e. the leading metric for value relevance is the residual dispersion metric. The reason for this is that this metric is least sensitive to the econometrical problems described in chapter three. However, to increase comparability with previous research and to test for robustness of results, the coefficient of determination metric and portfolio metric are also used. Besides a description of the different applied metrics for value relevance, also a description of the procedure to measure business change and the procedures for hypothesis testing are given. As a general note, in this study only the combined value relevance of book values and earnings is considered. The reason for this is that the residual dispersion metric is the leading measure of value relevance in this study and that there is no decomposition technique available for this.

The remainder of this section is organised as follows. In section 6.3.2 details on the coefficient of determination metric and residual dispersion metric are given. Section 6.3.3 elaborates on the portfolio metric, while in section 6.3.4 the procedure for measuring business change is provided. In section 6.3.5 the procedures for hypothesis testing are described.

6.3.2 The coefficient of determination metric and residual dispersion metric

The first step for both the coefficient of determination metric and residual dispersion metric is to decide on which relationships will be estimated. The relationships that are estimated are based on work by Gu (2007) and Francis and Schipper (1999):

$$P_{i,t} = \alpha_{0,t} + \alpha_{1,t} EPS_{i,t} + \alpha_{2,t} BVPS_{i,t} + \varepsilon_{i,t} \tag{6.1}$$

$$\frac{\Delta P_{i,t}}{P_{i,t-1}} = \beta_{0,t} + \beta_{1,t} \frac{EPS_{i,t}}{P_{i,t-1}} + \beta_{2,t} \frac{BVPS_{i,t}}{P_{i,t-1}} + \epsilon_{i,t} \tag{6.2}$$

$$R_{i,t} = \gamma_{0,t} + \gamma_{1,t} \frac{\Delta EPS_{i,t}}{P_{i,t-1}} + \gamma_{2,t} \frac{EPS_{i,t}}{P_{i,t-1}} + \gamma_{3,t} \frac{BVPS_{i,t}}{P_{i,t-1}} + \varphi_{i,t} \tag{6.3}$$

In these expressions is $P_{i,t}$ the price of a share of firm i in year t, $EPS_{i,t}$ and $BVPS_{i,t}$ are respectively the earnings and book value of equity per share, $R_{i,t}$ is the compound return and $\varepsilon_{i,t}$, $\epsilon_{i,t}$ and $\varphi_{i,t}$ are error terms. Finally, $\Delta P_{i,t} = P_{i,t} - P_{i,t-1}$, $\Delta EPS_{i,t} = EPS_{i,t} - EPS_{i,t-1}$ and $\Delta BVPS_{i,t} = BVPS_{i,t} - BVPS_{i,t-1}$.

It should be noted that all three of these models are derived from the accounting based valuation model of Ohlson (1995), providing a sound theoretical basis. Furthermore, (6.1) is described as a level model, while (6.2) and (6.3) are return models (Gu, 2007). For the coefficients of determination metric, models (6.2) and (6.3) should be preferred, because these partly eliminate scale effects, see Brown et al. (1999) and Gu (2007). However, since the level model is so widely used, it is included to provide comparability with previous research. The differing factors between model (6.2) and (6.3) is that (6.3) includes dividends in the dependent variable and changes in earnings in the independent variable. Note that (6.1) and (6.2) are taken directly from Gu (2007), while (6.3) is based on Francis and Schipper (1999).[33]

The cross-sectional regressions of (6.1), (6.2) and (6.3) are estimated for all of the sample years. The first measure of value relevance then is the resulting coefficient of

[33] More precisely, (6.3) is a small adaption from the relation used by Francis and Schipper (1999) to estimate their Ratio 2 portfolio.

determination for each year, i.e. $\hat{R}^2_{j,t}$ with $j \in \{1,2,3\}$ to indicate respectively models (6.1), (6.2) and (6.3).

To estimate the residual dispersion metric some additional steps have to be undertaken. As was described in section 3.3.5, Gu (2007) proposes a number of alternative methodologies to estimate abnormal pricing errors. In this research the following approach is taken:

- Estimate either (6.1), (6.2) or (6.3);
- Pool all data across years and sort these into deciles based on the absolute fitted values of the dependent variable;
- For each of the deciles calculate the pricing error defined as:

$$pricing\ error\ decile\ \equiv \sqrt{\frac{1}{N_d}\sum_{d=1}^{N_d} \hat{\varepsilon}_d^2} \tag{6.4}$$

- In (6.4), N_d is the number of observations in each decile and d is used to indicate a single observation in each decile. It should be noted that in (6.4) $\hat{\varepsilon}_n$ is used, but it can be interchanged for the errors of (6.2) and (6.3);
- For each year t the raw pricing error is calculated as:

$$pricing\ error\ year\ t\ \equiv \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t} \hat{\varepsilon}_{i,t}^2} \tag{6.5}$$

- In (6.5), N_t is the number of observations in each year. Gu (2007) imposes the restriction that $N_t \geq 50$. In this research the restriction is that $N_t \geq 150$, since the sample will be split up in subgroups later;
- For each year t the annual mean scale is calculated by the mean fitted value of that year. Then the corresponding decile in which the annual mean scale falls is identified; and
- The abnormal pricing error for year t is defined as the difference between the raw pricing error of year t and the pricing error of the decile to which year t corresponds.

Finally, the abnormal pricing error $AbPerr_{t,j}$, with $j \in \{1,2,3\}$ to indicate respectively models (6.1), (6.2) and (6.3), is used as a metric for value relevance in year t.

6.3.3 The portfolio metric

The portfolio metric used in this study is based on work by Francis and Schipper (1999), who propose, amongst others, an accounting-based hedge portfolio taking into consideration both the book value of equity, earnings and earnings change. Since the focus of this study is on the value relevance of earnings and book values combined, this hedge portfolio is the most appropriate starting point of the five hedge portfolios used by Francis and Schipper (1999).[34]

To form the accounting-based hedge portfolio, first the following regression is estimated for each sample year:

$$R_{i,t} - R_{i,t}^{m} = \delta_{0,t} + \delta_{1,t}\frac{\Delta EPS_{i,t}}{P_{i,t-1}} + \delta_{2,t}\frac{EPS_{i,t}}{P_{i,t-1}} + \delta_{3,t}\frac{BVPS_{i,t}}{P_{i,t-1}} + \vartheta_{i,t} \qquad (6.6)$$

Here $R_{i,t}^{m}$ is the market return for year t over the return window of firm i. The market return is defined as the return on the Dow Jones STOXX Europe 600 Index. Note that (6.6) is essentially the same as (6.3), except for this adjustment.[35]

After (6.6) is estimated, yearly observations are ranked based on the fitted value of the dependent variable. The accounting-based hedge portfolio for year t then takes long and short positions in respectively the top and bottom 40% of observations for a year. This leads to yearly observations of R_t^*, the return that could be earned with perfect foresight of accounting information. To correct for differences in the variation of market adjusted returns, R_t^* is scaled by R_t^P, which is the yearly return on a perfect foresight portfolio. To calculate R_t^P for each year, long and short positions are taken in the top and bottom 40% of returns in year t. Finally, R_t^* scaled by R_t^P is used as a metric for value relevance.

6.3.4 Measuring business change

Business change is not a variable that is directly observable. As such, a proxy has to be devised to measure business change. The proxy used in this research is taken from Lev and Zarowin (1999). The first step to measure business change is to rank yearly observation

34 Note that the other accounting-based hedge portfolios proposed by Francis and Schipper (1999) are based on earnings only, cash flows or financial signals, for the latter see Lev and Thiagarajan (1993). Finally, Chang (1999) also uses an accounting-based hedge portfolio that takes into consideration both earnings and the book value of equity, i.e. it is based on the return on equity. However, due to data limitations this portfolio is not implemented.

35 The specification of (6.6) is somewhat different than that applied by Francis and Schipper (1999), e.g. those authors don't deflate the book value per share and use the change in earnings deflated by the market value of equity at $t-1$ as an independent variable.

on the book value of equity at year end.[36] Then, the sample firms for each year are classified into equally sized deciles based on the rank of book value of equity. The following step is to measure the absolute change in portfolios for each firm, e.g. if firm i switches from portfolio seven to three, this would be indicated by a rank change of four. The next step is to calculate the mean absolute rank change for year t, defined as $MARC_t$. Note that $MARC_t$ increases as firms experience more change, i.e. jump across portfolios more often, and that it is zero in the limit if none of the firms experienced any change. $MARC_t$ is a measure of business change in year t and can be used to test if there has been an increase in business change over the years.

To test hypothesis H3 and H4, firms need to be separated in two groups, high changing firms and low changing firms (Lev and Zarowin, 1999). To do so, for each firm the across time absolute rank change, defined as $ATARC_i$, is calculated. This measure reflects the number of rank changes experienced by a firm, standardized by the number of years a firm is present in the sample. For example, if firm i is present in the sample in the period 1999-2002, and it has ranks six, six, four and seven respectively in these years, it will have an across time absolute rank change of 1.25. In the following section, the use of $ATARC_i$ to test H3 and H4 is explained.

6.3.5 Procedures for hypothesis testing

In this section, the procedures to test the various hypotheses are elaborated on. As a matter of notation, the various metrics for value relevance in year t will generally be referred to as $METRIC_t$, since the procedures for hypothesis testing do not differ per metric. Hence, $METRIC_t$ refers to the seven different metrics for value relevance used, i.e. the coefficient of determination, the abnormal pricing error or the portfolio metric for the different model specification.

To test hypothesis H1, on changes in value relevance over time, the following relation will be estimated:

$$METRIC_t = \omega_{0,1} + \omega_{1,1} TIME_t + \varepsilon_{t,1} \qquad (6.7)$$

In which $Time_t = 1, \dots, T$ and T is the number of years of data in the sample. If value relevance has increased (decreased) over time, $\widehat{\omega}_{1,1}$ should have a positive (negative) sign

[36] Lev and Zarowin (1999) use two methods to rank observations, i.e. by ranking them on the book value of equity or the market value of equity. Here only the book value of equity is used as a starting point, because i) the results of Lev and Zarowin (1999) are insensitive to the ranking method and ii) only limited data is available on the market value of equity in the Capital IQ database.

and be significantly different from zero for the coefficient of determination and portfolio metric and should have opposite signs for the residual dispersion metric. A similar procedure is employed to test hypothesis H2 concerning an increasing rate of business change:

$$MARC_t = \omega_{0,2} + \omega_{1,2}TIME_t + \varepsilon_{t,2} \tag{6.8}$$

As explained in chapter three, this procedure to test for changes in a variable over time is commonplace, see e.g. Collins et al. (1997), Francis and Schipper (1999), Gu (2007), Kim and Kross (2005) and Lev and Zarowin (1999).

Hypothesis H3 and H4 question whether there are significant differences in value relevance and the change in value relevance over time of firms experiencing high or low amounts of business change. To test these hypotheses is somewhat more elaborate. As is explained in the previous section, for each firm the across time absolute rank change is calculated. This is used to separate the sample firms in two groups; a high change group for which $ATARC_i > 0.1$ and a low change group for which $ATARC_i \leq 0.1$. For each of the two groups the different metrics of value relevance are then recalculated. That also means that for example the normal pricing errors for the deciles or the return on the perfect foresight portfolio are recalculated for each group. Based on this separation, the following observations of value relevance result: $METRIC_t^H$ for high change firms and $METRIC_t^L$ for low change firms.

To test hypothesis H3, the following relationship is estimated:

$$METRIC_t^g = \omega_{0,3} + \omega_{1,3}CHANGE_{g,t} + \varepsilon_{t,3} \tag{6.9}$$

In which $g \in \{H, L\}$ and $CHANGE_{g,t}$ is a dummy variable which takes value one if an observation belongs to the high change group and zero if an observation belongs to the low change group. If H3 is true, it is expected that $\widehat{\omega}_{1,3}$ is negative and significantly different from zero for the coefficient of determination and portfolio metric and vice versa for the residual dispersion metric.. The procedure described here is identical to that used by Lev and Zarowin (1999).

To test hypothesis H4, a coefficient dummy variable is used:

$$METRIC_t^g = \omega_{0,4} + \omega_{1,4}TIME_t + \omega_{2,4}CHANGE_{g,t} \cdot TIME_t + \varepsilon_{t,4} \tag{6.10}$$

In (6.10), the rate of change in value relevance over time for low business change firms is measured by $\omega_{1,4}$. The rate of change in value relevance over time of firms experiencing a high amount of business change is $\omega_{1,4} + \omega_{2,4}$. H4 conjectures that the rate of change will be lower for firms experiencing high amounts of business change. If value relevance has generally increased it is expected that this will be less the case for high change firms and if value relevance has decreased then this would be more sharply for high change firms. Hence, it is expected that $\widehat{\omega}_{2,4}$ is smaller than zero and significantly so for the coefficient of determination and portfolio metric, and vice versa for the residual dispersion metric.

Hypothesis H5 states that value relevance has increased in the sample countries after the mandatory introduction of IFRS in 2005. To test this hypothesis again a dummy variable is used:

$$METRIC_t = \omega_{0,5} + \omega_{1,5} IFRS_t + \varepsilon_{t,5} \tag{6.11}$$

Here $IFRS_t$ equals one for $t \geq 2005$. If H5 is true then it is expected that $\widehat{\omega}_{1,5}$ is significantly larger than zero for the coefficient of determination and portfolio metric and significantly smaller than zero for residual dispersion metric..

To test hypothesis H6 the same procedure as for hypothesis H3 is used. This means that the sample is again split up, only this time not based on high or low change, but on the country of origin. Firms from France and Germany are grouped together and this group represents firms operating in bank-oriented financial systems, while on the other hand the Netherlands and the United Kingdom are grouped together, representing firms in market-oriented financial systems. For each of the two groups all metrics of value relevance are recalculated. Also instead of estimating (6.9), the following is estimated:

$$METRIC_t^c = \omega_{0,6} + \omega_{1,6} COUNTRY_{c,t} + \varepsilon_{t,6} \tag{6.12}$$

In (6.12) $c \in \{bank, market\}$ and $COUNTRY_{c,t}$ is a dummy variable with value one for metrics of value relevance from market-oriented financial systems and zero otherwise. If hypothesis H6 is true, it is expected that $\widehat{\omega}_{1,6}$ is significantly larger than zero for the coefficient of determination and portfolio metric and significantly smaller than zero for residual dispersion metric.

6.4 Summary

In this chapter the hypotheses of this research were formulated and the research design was presented. To summarise, the hypotheses are:

- H1: Value relevance has changed significantly in the sample period in France, Germany, the Netherlands and the United Kingdom;
- H2: The rate of business change has been increasing in France, Germany, the Netherlands and the United Kingdom;
- H3: Value relevance is higher for firms experiencing little or no business change than for firms experiencing high amounts of business change;
- H4: The rate of change in value relevance is smaller for firms experiencing a high amount of business change;
- H5: Value relevance has increased in France, Germany, the Netherlands and the United Kingdom after the mandatory introduction of IFRS in 2005; and
- H6: Value relevance is higher in the United Kingdom and the Netherlands than in Germany and France.

The interpretation of value relevance used in this research is Interpretation 4, i.e. the focus is on the statistical association between financial information and share prices or returns. To test the hypotheses, three different metrics of value relevance will be used, namely the coefficient of determination, residual dispersion and portfolio metric. Of these three, the residual dispersion metric of Gu (2007) is leading, since this metric is least sensitive to the econometrical problems described in chapter three. For robustness and comparability, the coefficient of determination metric and portfolio metric are also included. Furthermore, for the residual dispersion metric and coefficient of determination metric, three different relations are estimated, again for robustness, while a separate market adjusted return model is estimated specifically for the portfolio metric.

7. Sample selection

7.1 Data and variable definitions

As is previously described, the sample of this research consists of firms in France, Germany, the Netherlands and the United Kingdom.[37] The starting principle for the sample selection was to obtain an as long and broad, i.e. number of years and firms respectively, sample as possible. Accounting data is taken from Compustat Global, which covers firms from 1988 to 2011. The following items were obtained from Compustat Global for all industrial firms in the sample countries for which data is available *(Compustat identifier in parentheses)*:

- International Security Identification Number;
- Fiscal year-end month;
- Total assets *(AT)*;
- Common/ordinary equity *(CEQ)*;
- Income before extraordinary items *(IB)*;
- Special items *(SPI)*;
- Extraordinary items and discontinued operations *(XIDO)*; and
- Common shares outstanding – Issue *(CSHOI)*.[38]

Based on the International Security Identification Numbers from Compustat Global share price and return data was obtained from the S&P Capital IQ database. Unfortunately, coverage of firms in this database before 1992 is rather limited, meaning that some of the first sample years were lost; see the subsequent sections for further details. The S&P Capital IQ database provides both 'adjusted' and 'raw' closing prices. Raw closing prices are merely adjusted for stock splits, while adjusted closing prices are also adjusted for cash dividends, rights offerings and spin-offs. Finally, market returns are defined as the return on the Dow Jones STOXX Europe 600 Index, data on which is available through www.stoxx.com.

[37] All used data is available upon request from the author.
[38] Industrial firms were selected by filtering on *INDL* in the screening variables. Accounting data was obtained on the 27th of June 2012. It should be noted that as such, the sample of firms for the year 2011 is incomplete. However, this year is nevertheless included in this research, as there is a priori no reason to suspect that the included firms are susceptible to sampling bias.

Table two provides the exact definitions of the variables used in this research and also relates the obtained data to the variable definitions of chapter six. Furthermore, the variable definitions applied in this research closely resemble those of Gu (2007).

Table 2

Variable definitions

Variable	*Definition*
$P_{i,t}$/price	▪ S&P Capital IQ closing price (not adjusted) three months after fiscal year end of firm i in year t
$R_{i,t}$/return	▪ The return of firm i in year t, calculated as the adjusted closing price from the S&P Capital IQ database three months after fiscal year end divided by the adjusted closing price nine months before fiscal year end minus one[39]
$R_{i,t}^{m}$/market return	▪ The market return in year t based on the Dow Jones Stoxx Europe 600 Index. Note that the market return is calculated per firm, i.e. it corresponds to the return window of nine months before fiscal year end up to the third month after fiscal year end of firm i
Earnings	▪ Earnings are defined as income before extraordinary items plus extraordinary items and discontinued operations from the Compustat Global database. Note that income before extraordinary items in the Compustat Global database also excludes discontinued operations
$EPS_{i,t}$/earnings per share	▪ Earnings as defined above in year t divided by common shares outstanding – issue in year t from the Compustat Global database
Book values	▪ Common/ordinary equity from the Compustat Global database
$BVPS_{i,t}$/book value per share	▪ Book values as defined above in year t divided by common shares outstanding – issue in year t from the Compustat Global database
Non-recurring items	▪ Special items plus extraordinary items and discontinued operations from the Compustat Global database.

7.2 Data filters

To correct for outliers a number of data filters are applied. The first requirement for the data is that each firm-year observation has complete information for the necessary variables. The number of data items necessary depends on the model estimated, e.g. for (6.2), (6.3) and (6.6) it is necessary that the share price at $t-1$ is available, while this is not the case for (6.1). The following filters are then applied to the data: i) all observations with negative values on book value or total assets are removed, ii) all observations that fall in the top or bottom one-half of earnings-to-price are removed, iii) all observations that fall in the top or bottom one-half percent of book value per share-to-share price are removed and iv) all observation that fall in the top or bottom one-half percent of non-recurring items-to-

[39] Return definitions, and especially the length of the window to calculate returns tend to vary across researches. The definition applied here is the same as that used by Gu (2007).

earnings are removed. These filters are similar to those applied by Collins et al. (1997) and Gu (2007) to remove outliers.

An additional filter, also applied by e.g. Collins et al. (1997), Brown et al. (1999) and Gu (2007), is that extreme observations based on studentised residuals are deleted from the sample. To do so, in this research an iterative process is applied. Each of the models (6.1), (6.2), (6.3) and (6.6) is estimated for their sample. Then, in each sample all observations that have studentised residuals larger than four are removed. This process is repeated until there are no more observations that have studentised residuals larger than four.[40] For all samples this process is also done for regressions of price against earnings per share, price against book values per share and price against earnings per share and book values per share. The final filter applied is that each of the years has at least a total number of 150 observations, while each of the subgroups based on change or financial system contains at least 50 observations.

7.3 Descriptive statistics

Table three reports the number of yearly observations per model after all filters have been applied. For the level model, the years 1988-1990 were removed because of an insufficient number of complete observations. The same is the case for the years 1988-1991 for the price change and return models.

Table 3

Number of yearly observations after filters have been applied

Year	**Level model (6.1)**	**Price change model (6.2)**	**Return model (6.3)**	**Market adjusted return model (6.6)**
1991	455	-	-	-
1992	696	437	443	443
1993	724	648	657	653
1994	772	692	700	700
1995	847	785	797	797
1996	981	880	884	884
1997	1,486	1,253	1,259	1,262
1998	1,606	1,308	1,318	1,313
1999	1,751	1,346	1,335	1,336

[40] Gu (2007) notes that outcomes may be sensitive to the number of iterations used in this process. He also notes that if outliers are a serious issue, the process described here should be preferred over fixing a small number of iterations. Since the dataset is fairly large, a priori not limiting the number of iterations thus seems the best approach; however as a robustness check outcomes have also been calculated based on a fixed number of iterations. Details on these are presented in chapter nine.

Table 3 (cont'd)

Number of yearly observations after filters have been applied

Year	**Level model (6.1)**	**Price change model (6.2)**	**Return model (6.3)**	**Market adjusted return model (6.6)**
2000	1,913	1,527	1,519	1,521
2001	1,722	1,450	1,453	1,452
2002	1,655	1,459	1,460	1,461
2003	1,603	1,383	1,379	1,374
2004	1,523	1,284	1,276	1,276
2005	1,464	1,217	1,222	1,221
2006	1,520	1,225	1,217	1,217
2007	1,250	1,086	1,086	1,085
2008	1,047	962	960	962
2009	1,100	1,000	1,000	996
2010	1,035	939	933	932
2011	794	721	715	715
Total	**25,944**	**21,602**	**21,613**	**21,600**

Note: The initial sample consisted of the years 1988-2011. For each year it is required that there are at least 150 complete observations available.

Table four presents descriptive statistics on the samples used for the different models. It can be seen that the samples do not significantly differ from one another, although there is a slight difference between the sample statistics of the level model and the sample statistics of the other models. What is very apparent from this table is that all variables are positively skewed; the mean is significantly larger than the median for all variables, each variable has a relatively high standard deviation and each variable has a positive skew.

Collins et al. (1997) and Gu (2007) provide the same descriptive statistics as reported in table three, except for skewness. It appears from their descriptive statistics that their samples are also positively skewed, however not as severely as the sample of this research. Possible explanations for this difference might be the larger dataset employed by these authors, or a structural difference in the samples. The skewness of this data could indicate that the sample is susceptible to outliers. Since rather strict filters have already been applied to remove outliers, no further measures have been taken.

Table 4

Descriptive statistics

Variables	*Mean*	*Standard deviation*	*Lower Quartile*	*Median*	*Upper Quartile*	*Skewness*
Panel A: Level model (6.1) – Sample size = 25,944						
Price	10.187	15.024	1.225	3.900	13.433	3.703
Earnings per share	2.167	8.770	0.008	0.173	0.968	9.205
Book value per share	22.088	71.900	0.650	2.099	10.239	8.619
Panel B: Price change model (6.2) – Sample size = 21,602						
Price	11.759	16.923	1.450	4.750	15.800	3.752
Earnings per share	2.413	9.870	0.009	0.210	1.342	9.312
Book value per share	25.574	79.481	0.806	2.823	13.559	8.309
Panel C: Return model (6.3) – Sample size = 21,613						
Price	11.705	16.461	1.460	4.793	15.764	2.920
Earnings per share	2.425	9.680	0.010	0.212	1.336	7.887
Book value per share	25.724	77.830	0.812	2.793	13.488	6.699
Panel D: Market adjusted return model (6.6) – Sample size = 21,600						
Price	11.712	16.469	1.460	4.800	15.772	2.918
Earnings per share	2.423	9.668	0.010	0.213	1.337	7.905
Book value per share	25.690	77.764	0.811	2.793	13.482	6.710

Note: Prices are S&P Capital IQ closing prices (not adjusted) three months after fiscal year end of firm i in year t. Earnings are defined as income before extraordinary items plus extraordinary items and discontinued operations from the Compustat Global database. Earnings per share are defined as earnings in year t divided by common shares outstanding – issue in year t from the Compustat Global database. Book value is defined as Common/ordinary equity from the Compustat Global database and book value per share as book value in year t divided by common shares outstanding – issue in year t from the Compustat Global database.

8. Results

8.1 Introduction

In this chapter the results of this research are presented. As was previously stressed and expected, the outcomes of the hypotheses tests are dependent on the model applied. Given the econometrical issues related to the coefficient of determination metric, the leading measure of value relevance is the residual dispersion metric or abnormal pricing error. From the three model specifications of chapter six, the price change and return model are preferred over the level model, because these models partly eliminate scale effects.

The remainder of this chapter is organised as follows; in section 8.2 the results of tests of hypothesis H1, concerning changes in value relevance over time, are presented. Section 8.3 then provides results on hypothesis H2 on the rate of business change over time, while section 8.4 relates the rate of business change to value relevance and results of hypotheses H3 and H4 are presented. Section 8.5 provides an overview of the results of hypothesis H5 on IFRS and section 8.6 of the results of hypothesis H6 on international differences in value relevance. Where necessary, these sections are split up for the coefficient of determination, residual dispersion and portfolio metric. Section 8.7 then summarises the findings of this research.

8.2 H1: Changing value relevance over time

8.2.1 Coefficient of determination metric

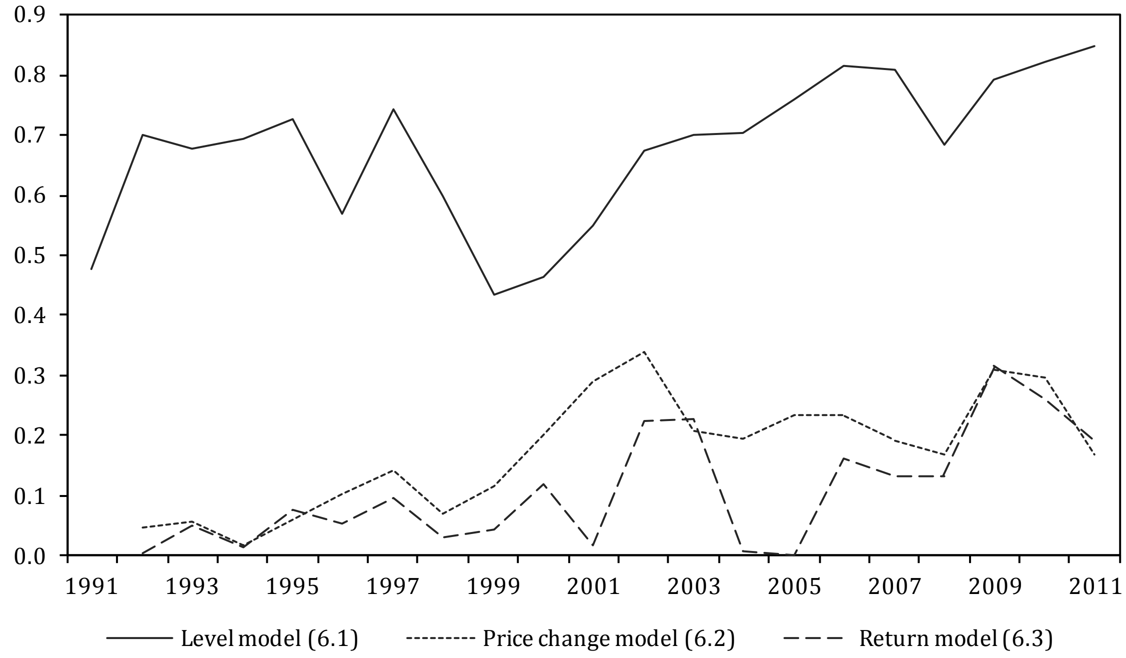
Figure one presents yearly estimates of coefficients of determination for the three different model specification; the level model, price change model and return model.[41] The range of the reported coefficients of determination for the level model reconcile to those reported in U.S. researches with a range of 0.435 in 1999 to 0.847 in 2011. The coefficients of determination for the price change and return model are substantially lower than those of the level model, something that is also generally found.

For the price change model, coefficients of determination in early years, i.e. from 1992 through 1995, can be considered low with values well below 0.1. The coefficients of determination for the return model shows a similar trend as those of the price change

[41] Please note that all coefficients of determination that are reported in this paper concern adjusted coefficients of determination.

Figure 1

Temporal patterns of coefficients of determination



Note: These coefficients of determination are obtained through the following regression; for the level model $P_{i,t} = \alpha_{0,t} + \alpha_{1,t}EPS_{i,t} + \alpha_{2,t}BVPS_{i,t} + \varepsilon_{i,t}$, for the price change model $\Delta P_{i,t}/P_{i,t-1} = \beta_{0,t} + \beta_{1,t}\,EPS_{i,t}/P_{i,t-1} + \beta_{2,t}\,BVPS_{i,t}/P_{i,t-1} + \epsilon_{i,t}$ and for the return model $R_{i,t} = \gamma_{0,t} + \gamma_{1,t}\,\Delta EPS_{i,t}/P_{i,t-1} + \gamma_{2,t}\,EPS_{i,t}/P_{i,t-1} + \gamma_{3,t}\,BVPS_{i,t}/P_{i,t-1} + \varphi_{i,t}$.

model, however this trend is somewhat distorted by the years 2004 and 2005. In these years the coefficient of determination for the return model is close to zero. Moreover, for 2005 the return model as a whole is statistically not significant with an F-statistic of 1.195 and corresponding p-value of 0.310. There is only one other regression which is statistically not significant as a whole, and that is for the year 1992 for the return model. This regression has an F-statistic of 1.818 and corresponding p-value of 0.143.

From figure one it appears that there has been an increase in the coefficients of determination over the years, especially for the period 2006 to 2011 the coefficients of determination are relatively high for all model specifications. In table five the results of a regression of the coefficients of determination against a time variable are presented. The regression specification is as in (6.7). As can be seen from this table, for all models $\widehat{\omega}_{1,1}$ is positive and significant at the 1% level. Based on table five there is only one possible conclusion to draw; there has been a significant increase in value relevance measured by the coefficient of determination metric over the period 1991-2011 or 1992-2011.

Table 5

Association between coefficients of determination and time trend
$METRIC_t = \omega_{0,1} + \omega_{1,1}TIME_t + \varepsilon_{t,1}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\widehat{\omega}_{0,1}$	0.557	0.049	-0.003
T-statistic	11.968	1.589	-0.080
p-value	0.000	0.129	0.937
$\widehat{\omega}_{1,1}$	0.011	0.012	0.011
T-statistic	2.977	4.531	3.663
p-value	0.008	0.000	0.002

Note: The table provides coefficient estimates, T-statistics and p-values (two-sided for a null hypothesis of equal to zero) for regression of coefficient of determination estimates against a time variable. The level model contains coefficient of determination estimates for the years 1991 up to 2011 and as such $TIME_t = 1, \dots, 21$. The price change and return model are estimated yearly for the period 1992 up to 2011 and for these models $TIME_t = 1, \dots, 20$. Note that a positive value for $\widehat{\omega}_{1,1}$ indicates an increase in value relevance.

8.2.2 Residual dispersion metric

The residual dispersion metric is preferred over the coefficient of determination metric, because it is not influenced by the variance-covariance matrix of the independent variables and because it can handle non-proportional scale effects. The estimated variance-covariance matrix of the independent variables is an inherent sample property and is likely to differ across samples drawn from the same population. When comparing coefficients of determination, samples are not drawn from the same population, and even the population variance-covariance-matrices are likely to differ as the population of firms changes over time. However, these sample properties are not further investigated in this research.

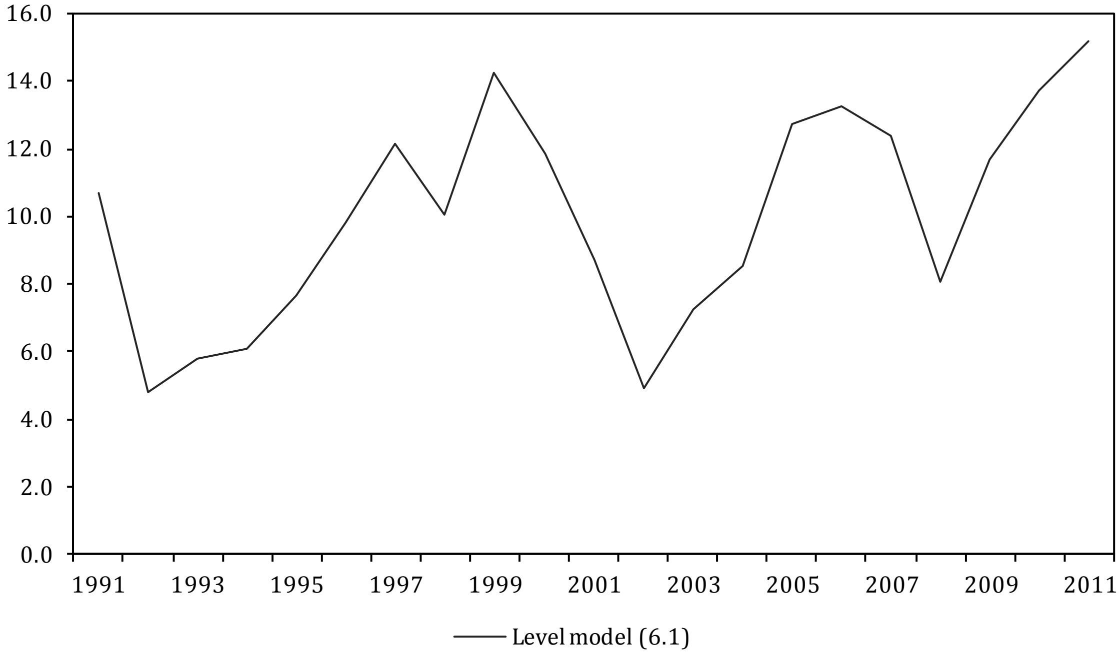
The second reason why the residual dispersion metric is preferred, i.e. that it can handle with non-proportional scale effects, will be investigated in more depth. Gu (2007) clearly documents the non-proportionality of scale effects, as well as a change in average scale over time. It is interesting to see if these characteristics are also present in this European sample.

Figure two depicts temporal patterns of the mean absolute fitted variables of the different models. From these figures it can be seen that the mean absolute fitted value of prices and returns, a proxy for average scale, has significant variability over time. As was discussed in chapter three, this could imply that scale effects are an issue for over time comparisons of coefficients of determination.

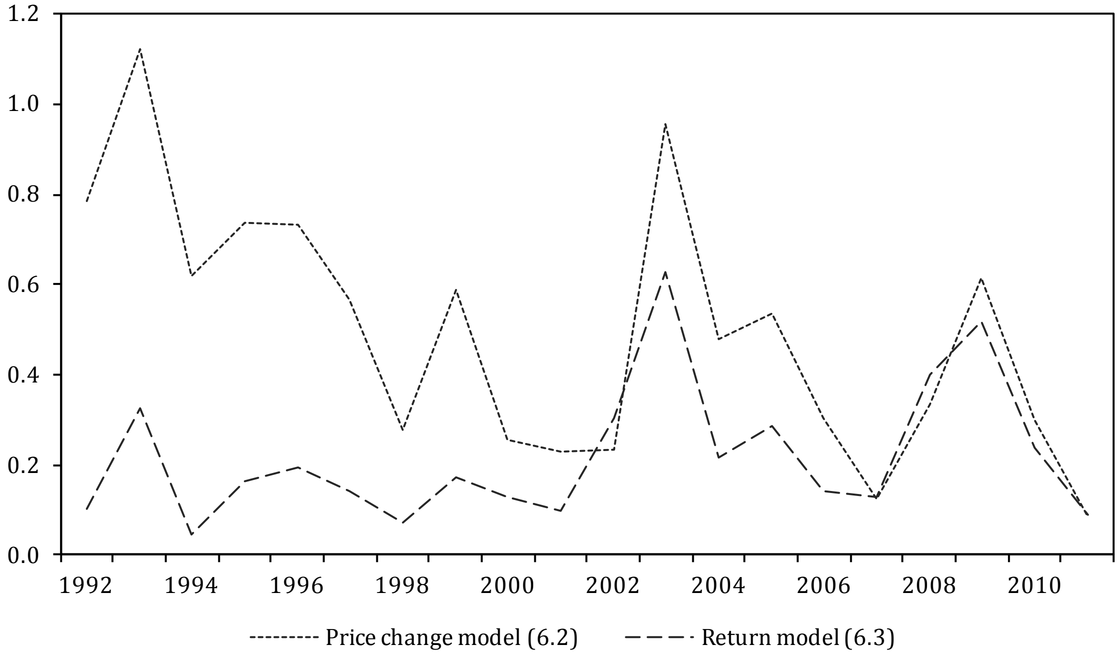
Figure 2

Temporal patterns of scale

2.a Mean $|\hat{P}|$



2.b Mean $|\Delta\hat{P}/P|$ and $|\hat{R}|$



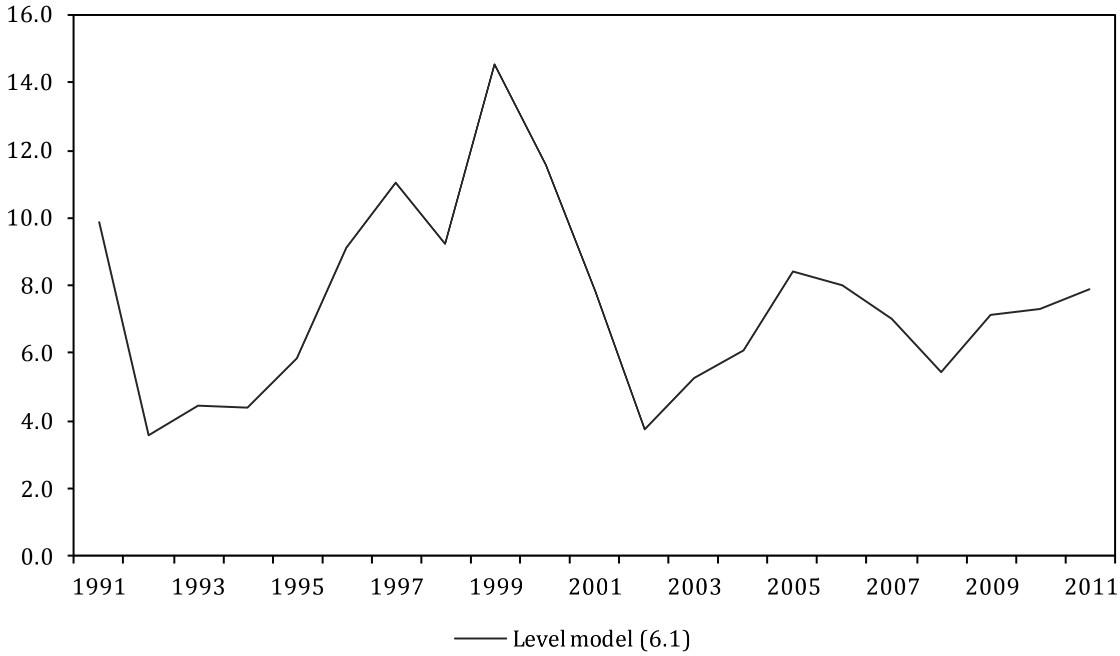
Note: These figures depict the temporal patterns of the mean absolute fitted values of the different models. The mean absolute fitted values serve as an indication of the average scale in a year.

Figure three depicts raw pricing errors, which are calculated as in (6.5), i.e. residual standard deviations. Similar as reported by Gu (2007) the temporal pattern of raw pricing errors closely follow those of scale depicted in figure two. This is expected and indicates that pricing errors are relatively high in years in which prices, price changes or returns are high. As such, any pattern present in raw pricing errors could be caused by a change in average scale, instead of by a change in value relevance.

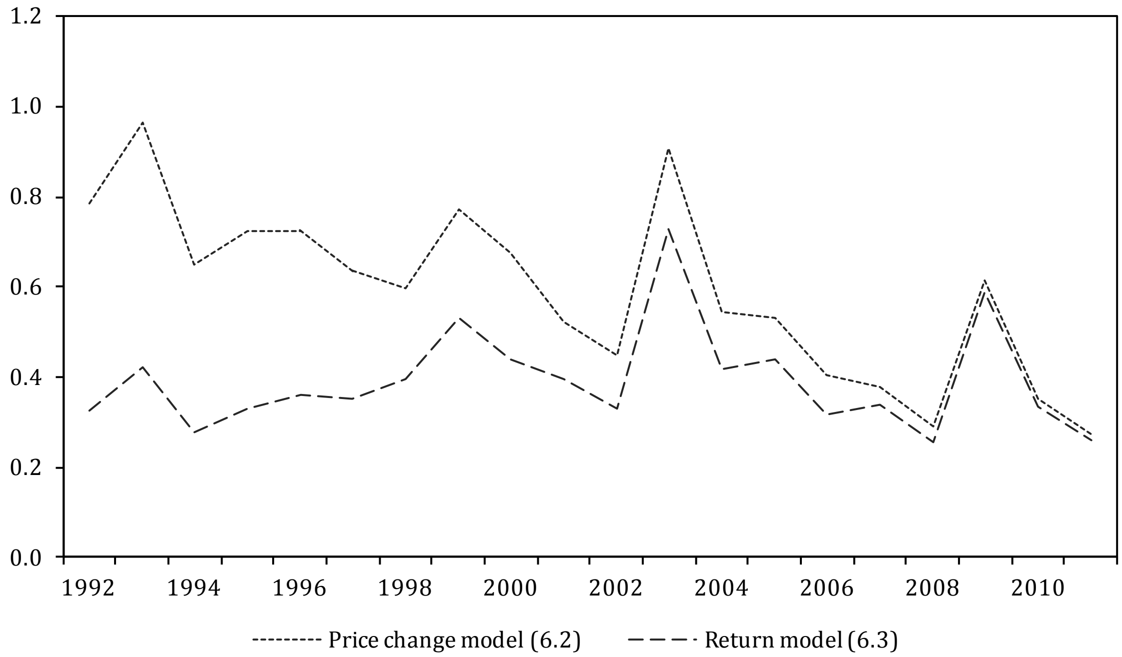
Figure 3

Temporal patterns of raw pricing errors

3.a



3.b



Note: Raw pricing errors are calculated yearly as the residual standard deviations of the different models.

Based on the analysis above, a similar conclusion as by Gu (2007) is reached; the average scale of prices, price changes and returns varies over time. Furthermore, the standard deviation of residuals, a key input for calculating the coefficient of determination, closely follows the pattern of average scale. This is especially the case for the level model, which is also most susceptible to scale effects. However, this finding would be less of an issue if scale effects are proportional. If this is the case, deflating the variables as suggested by Brown (1999) would remedy the problem.

Gu (2007) established that scale effects are non-proportional in the U.S. In table six it can be seen that this is also the case for this European sample. For example for the return model, the mean absolute fitted value in decile 10 is approximately 21 times the mean absolute fitted value of decile one, while the pricing error of decile 10 is not twice as large as the pricing error of decile one.

As is explained in chapter three, the non-proportionality of scale effects is corrected for by using abnormal pricing errors. These are depicted in figure four for the different models. From figure four it appears that abnormal pricing errors have been decreasing over the years for the level and price change model, especially in the final sample years. Note that since abnormal pricing errors are an inverse metric of value relevance; figure four hints towards an increase in value relevance over time for these models.

Table 6

Scale deciles and pricing errors

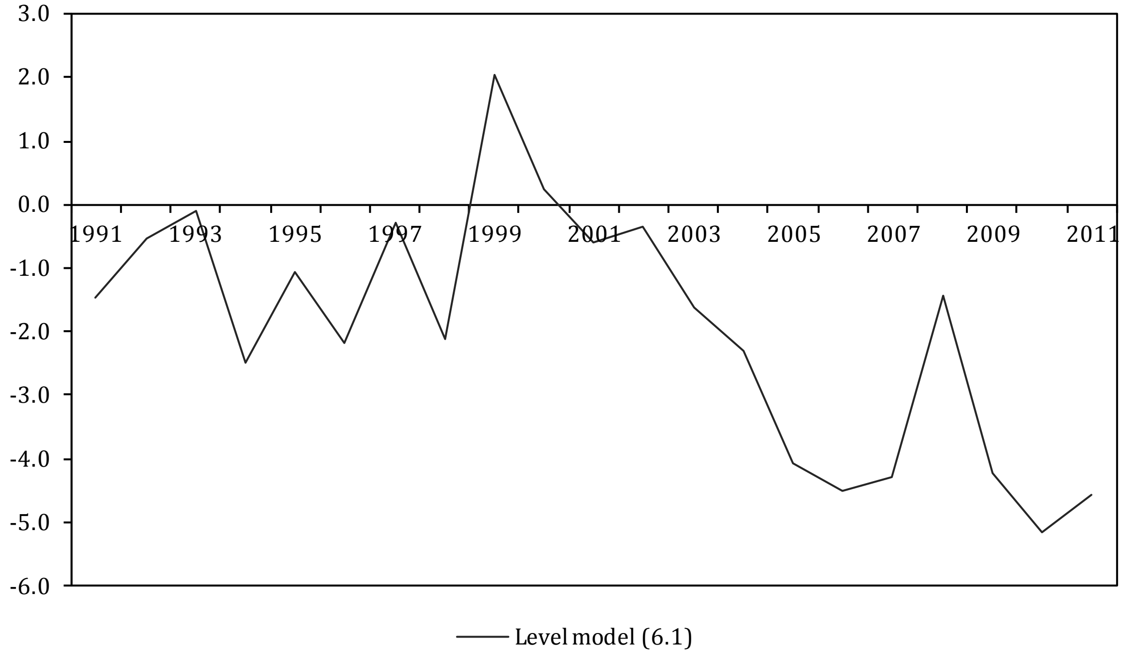
	Level model (6.1)		**Price change model (6.2)**		**Return model (6.3)**	
Scale decile	*Mean* $\lvert\hat{P}\rvert$	*Pricing error*	*Mean* $\lvert\Delta\hat{P}/P\rvert$	*Pricing error*	*Mean* $\lvert\hat{R}\rvert$	*Pricing error*
1	2.370	3.234	0.036	0.436	0.032	0.347
2	3.334	2.893	0.116	0.515	0.070	0.358
3	4.086	3.714	0.198	0.506	0.096	0.368
4	4.944	4.093	0.267	0.486	0.116	0.393
5	5.614	4.566	0.360	0.481	0.145	0.412
6	6.850	6.900	0.461	0.576	0.190	0.389
7	8.750	8.420	0.539	0.635	0.238	0.397
8	10.668	11.338	0.643	0.672	0.295	0.407
9	16.419	12.495	0.770	0.755	0.377	0.415
10	39.014	15.157	1.246	0.885	0.689	0.646

Note: Pricing errors are calculated by i) estimating the different models, ii) grouping all observations across years per model and sorting these into deciles based on the absolute fitted value and iii) for each decile the pricing error is calculated by (6.4).

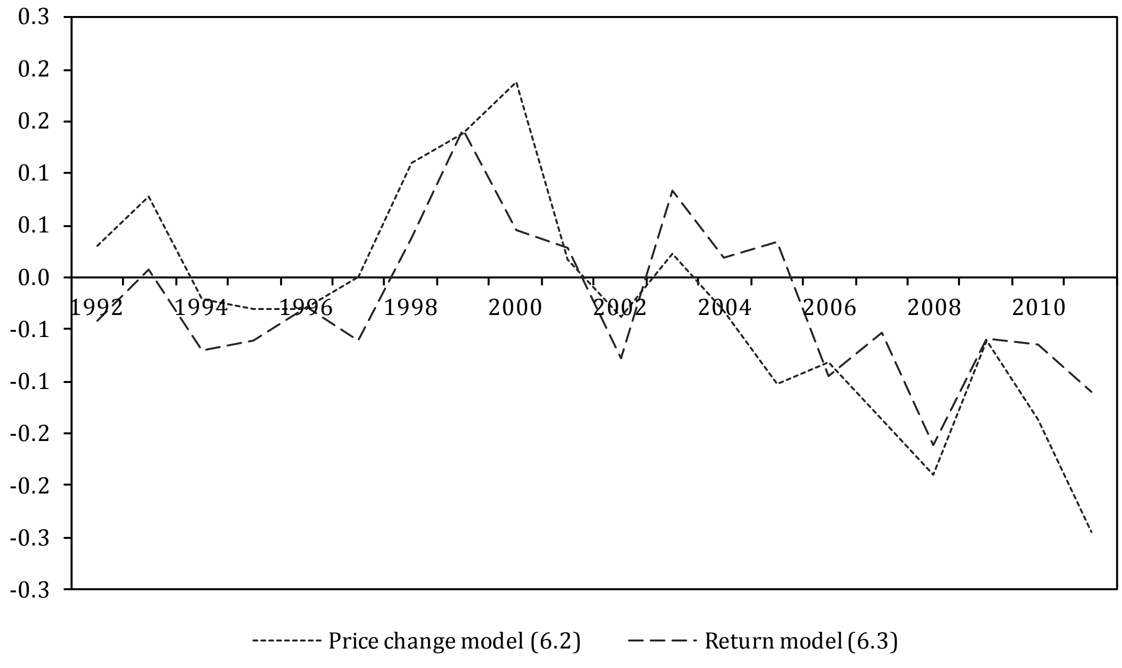
Figure 4

Temporal patterns of abnormal pricing errors

4.a



4.b



Note: These figures depict the temporal patterns of the abnormal pricing errors of the different models. Abnormal pricing errors are calculated by subtracting the pricing error of the scale decile in which a year belongs from the raw pricing error, see also section3.3.5 and 6.3.2.

Table 7

Association between abnormal pricing errors and time trend
$METRIC_t = \omega_{0,1} + \omega_{1,1}TIME_t + \varepsilon_{t,1}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\widehat{\omega}_{0,1}$	0.226	0.103	0.014
T-statistic	0.329	2.717	0.437
p-value	0.745	0.014	0.667
$\widehat{\omega}_{1,1}$	-0.199	-0.012	-0.004
T-statistic	-3.628	-3.889	-1.340
p-value	0.002	0.001	0.197

Note: The table provides coefficient estimates, T-statistics and p-values (two-sided for a null hypothesis of equal to zero) for regressions of abnormal pricing error estimates against a time variable. The level model contains abnormal pricing error estimates for the years 1991 up to 2011 and as such $TIME_t = 1, \dots, 21$. The price change and return model are estimated yearly for the period 1992 up to 2011 and for these models $TIME_t = 1, \dots, 20$. Note that a negative value for $\widehat{\omega}_{1,1}$ indicates an increase in value relevance.

The results for a formal test of the time pattern of abnormal pricing errors are provided in table seven. For both the level model and price change model there is significant evidence of increasing value relevance over time. The time trend for these models is negative with respectively T-statistics of -3.628 and -3.889 for the level model and price change model. For the return model $\widehat{\omega}_{1,1}$ is not significantly different from zero. Hence, the evidence points towards increasing value relevance in the sample period for two out of the three models, but this is not confirmed by the return model.

8.2.3 Portfolio metric

The final metric that is considered is the portfolio metric, which is used as cross-check and to test for robustness of results. Figure five depicts the yearly market adjusted return on the accounting based hedge portfolio and the perfect foresight hedge portfolio. Also, the proportion of the perfect hedge portfolio that is earned by the accounting based hedge portfolio, i.e. R_t^*/R_t^P, is depicted. From figure five it can be seen that the perfect foresight portfolio earns yearly returns of approximately 50.0% to 130.0%. The accounting based hedge portfolio earns much smaller returns, in the range of circa 10.0% to 55.0%. The average return for the accounting based hedge portfolio is 30.2%, while for the perfect foresight portfolio this is 73.5%. Finally, the accounting based hedge portfolio earns on average 40.9% of perfect foresight returns. This figure is substantially below what Francis and Schipper (1999) report for their Ratio2 portfolio, which earns on average 61.0% of the perfect foresight portfolio.

Figure 5
Returns and portfolio metric

Note: The return on the accounting based hedge portfolio is the return that could be earned by perfect foresight of earnings and book values and the use of (6.6). The return on the perfect foresight portfolio is the return that could be earned by perfect foresight of returns. Finally the % explained returns is the return on the accounting based hedge portfolio divided by the return on the perfect foresight portfolio.

Figure five also shows that the temporal patterns of the returns of the two portfolios closely mimic each other and that mostly the level of returns earned differs. This can also be seen by the relatively stable trend of the percentage explained returns. If value relevance would be increasing as indicated by the coefficient of determination and residual dispersion metric, an increasing trend in the percentage explained returns is expected.

A formal test of this hypothesis shows however no such trend. Regression of the percentage explained returns against a time variable does yield a positive sign for $\widehat{\omega}_{1,1}$, but these results are not significant with a T-statistic of 0.483. Results of this regression are shown in table eight.

8.3 H2: The rate of business change

In the previous section some evidence was presented of an increase in value relevance over time. Lev and Zarowin (1999) identified the rate of business change as a key driving factor behind declining value relevance in the U.S. That no evidence of declining value relevance was found in the previous section does not necessarily imply that value

Table 8

Association between portfolio metric and time trend
$METRIC_t = \omega_{0,1} + \omega_{1,1}TIME_t + \varepsilon_{t,1}$

$\widehat{\omega}_{0,1}$	0.386
T-statistic	7.071
p-value	0.000
$\widehat{\omega}_{1,1}$	0.002
T-statistic	0.483
p-value	0.635

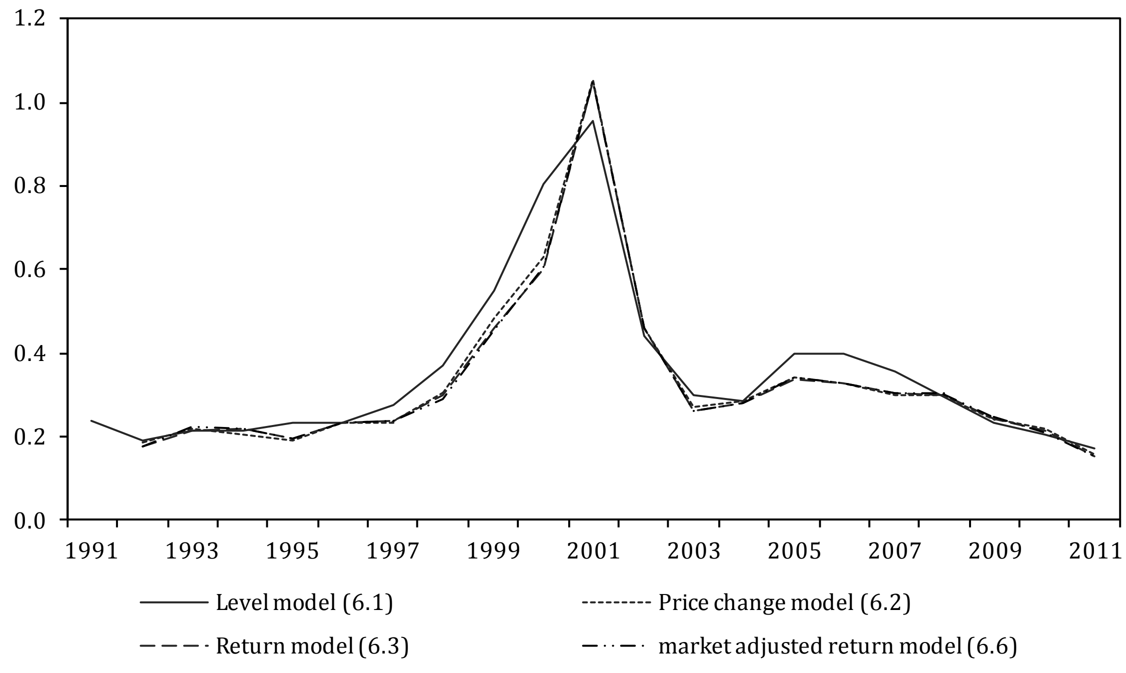
Note: The table provides coefficient estimates, T-statistics and p-values (two-sided for equal to zero) for regressions of the portfolio metric against a time variable. The time variable is defined as $TIME_t = 1, \dots, 20$.

relevance in this European sample is not influenced by the rate of business change. As such it is useful to investigate temporal patterns of this variable.

Business change is measured by the mean absolute rank change of all firms in a given year t, or equivalently by $MARC_t$. Refer to section 6.3.4 for details on the calculation of $MARC_t$. Figure six show the movements of $MARC_t$ over the sample period for the different samples. It can be seen that there are no structural differences in $MARC_t$ for the different samples. Furthermore, there does not seem to be a clear pattern in $MARC_t$ over the

Figure 6

Temporal pattern of the rate of business change



Note: This figure shows the temporal pattern of the rate of business change as measured by $MARC_t$. For details on the calculation of $MARC_t$, please refer to section 6.3.4.

entire sample period, but there does seem to be an increase in the rate of business change for 1991-2001, after which it sharply drops again.

Table nine shows the results of a regression of $MARC_t$ against a time variable. As was expected based on figure six, no substantial differences exist between the rates of business change across the different samples. That no clear time trend is present over the entire sample period is also confirmed by this analysis; none of the estimates of $\omega_{1,2}$ is significantly larger than zero. Finally, note that the reported p-values in table nine are for a one-sided test of $\omega_{1,2}$ being larger than zero, since an increase in the rate of business change was expected.

Table 9

Association between the rate of business change and time trend

$MARC_t = \omega_{0,2} + \omega_{1,2}TIME_t + \varepsilon_{t,2}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)	Market adjusted return model (6.6)
$\widehat{\omega}_{0,2}$	0.336	0.328	0.324	0.324
T-statistic	3.593	3.347	3.356	3.364
p-value	0.002	0.004	0.004	0.004
$\widehat{\omega}_{1,2}$	0.001	0.000	0.000	0.000
T-statistic	0.181	0.042	0.055	0.048
p-value	0.429	0.484	0.478	0.481

Note: The table provides coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,2}$ equal to zero and one-sided for $\widehat{\omega}_{1,2}$ smaller than zero) for regressions of $MARC_t$ against a time variable for the different samples of this research. The level model contains the years 1991 up to 2011 and as such $TIME_t = 1, \dots, 21$. The price change, return and market adjusted return models contain the years 1992 up to 2011 and for these models $TIME_t = 1, \dots, 20$.

8.4 H3 and H4: Business change and value relevance

8.4.1 Coefficient of determination metric

In this section the influence of business change on value relevance will be investigated. As was described in section 6.3.4, the firms in each of the samples have been divided into a low and high change group based on their across time absolute rank change or $ATARC_i$. The first hypothesis that will be investigated is H3; concerning differences in the level of value relevance between these two groups. For this purpose, table 10 summarises the yearly estimates of the coefficient of determination for the low and high change group for the different models.

Table 10

Coefficient of determination estimates for high and low change firms

	Level model (6.1)		**Price change model (6.2)**		**Return model (6.3)**	
Year	*High change*	*Low change*	*High change*	*Low change*	*High change*	*Low change*
1991	0.484	0.428	-	-	-	-
1992	0.688	0.702	0.064	-0.007	0.030	0.015
1993	0.663	0.675	0.073	0.042	0.061	0.033
1994	0.644	0.706	0.023	-0.004	0.011	0.011
1995	0.646	0.778	0.068	0.045	0.086	0.062
1996	0.526	0.561	0.125	0.044	0.057	0.033
1997	0.750	0.718	0.140	0.151	0.091	0.121
1998	0.561	0.635	0.077	0.049	0.027	0.031
1999	0.332	0.590	0.113	0.117	0.049	0.035
2000	0.337	0.617	0.203	0.220	0.108	0.197
2001	0.534	0.555	0.294	0.263	0.022	0.034
2002	0.632	0.747	0.343	0.305	0.225	0.217
2003	0.657	0.786	0.183	0.331	0.196	0.396
2004	0.646	0.796	0.185	0.242	0.019	0.003
2005	0.740	0.789	0.225	0.290	0.096	0.024
2006	0.795	0.852	0.223	0.281	0.140	0.267
2007	0.773	0.846	0.174	0.258	0.123	0.176
2008	0.669	0.669	0.205	0.043	0.155	0.049
2009	0.762	0.820	0.315	0.299	0.339	0.266
2010	0.815	0.829	0.322	0.238	0.299	0.135
2011	0.815	0.874	0.184	0.132	0.208	0.134
Average	**0.641**	**0.713**	**0.177**	**0.167**	**0.117**	**0.112**

Note: Coefficients of determination estimates for firms classified as high change and low change. Classification of firms in each of these groups is based on their across time absolute rank change, or $ATARC_i$. Firms for which $ATARC_i > 0.1$ are classified as high change, while firms for which $ATARC_i \leq 0.1$ are classified as low change.

It can be seen that for the levels model, the coefficient of determination is generally higher for the low change group than for the high change group. This is the case for 18 of the 21 years in this sample and is also exemplified by the higher average coefficient of determination. For the price change and return model such a difference is not directly visible; the average coefficient of determination is lower for the low change group than for the high change group, although only slightly. This finding appears contrary to the findings of Lev and Zarowin (1999).

The outcomes of a test on differences in levels of coefficients of determination for the low and high change group are presented in table 11. Table 11 shows that coefficients of determination for the high change group are lower than those of the low change group; the sign of $\widehat{\omega}_{1,3}$ is negative for all three models. However, only for the level model this finding is significant in a one-sided T-test. As such, only weak evidence is found that value relevance is lower for firms experiencing a high amount of business change. It is not directly clear what drives the differing results between the different models. Based on the discussion in chapter three, the first reason that comes to mind is the incomparability of coefficients of determination. The subsequent section shows that this actually is not the case.

Table 11

Differences in coefficients of determination for high change and low change firms

$METRIC_t^g = \omega_{0,3} + \omega_{1,3}CHANGE_{g,t} + \varepsilon_{t,3}$

	Level model (6.1)	**Price change model (6.2)**	**Return model (6.3)**
$\widehat{\omega}_{0,3}$	0.713	0.167	0.112
T-statistic	25.423	7.052	4.891
p-value	0.000	0.000	0.000
$\widehat{\omega}_{1,3}$	-0.072	-0.010	-0.005
T-statistic	-1.807	0.299	0.161
p-value	0.039	0.383	0.437

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,3}$ equal to zero and one-sided for $\widehat{\omega}_{1,3}$ smaller than zero) for a regression of coefficients of determination for the low and high change group against a dummy variable with value one for coefficients of determination of the high change group and zero for coefficients of determination of the low change group. The total number of observations in this regression is thus 42 for the level model and 40 for the change or return model. Note that a negative value of $\widehat{\omega}_{1,3}$ indicates higher value relevance for the low change group, which is the expectation.

Hypothesis H4 states that the rate of change in value relevance is smaller for firms experiencing a high amount of business change. To test this hypothesis the coefficients of determination are regressed against a time variable and a coefficient dummy variable. Results of this analysis are summarized in table 12.

The results of a test of hypothesis H4 are similar to those of H3. For the level model the change in value relevance is significantly lower for the high change group than for the low change group. This can be seen by the negative sign of $\widehat{\omega}_{2,4}$ and its p-value of 0.026 in a one-sided test. Despite this finding, note that value relevance for the level model has been increasing for both groups as $\widehat{\omega}_{1,4}$ is postitve and the sum of $\widehat{\omega}_{1,4}$ and $\widehat{\omega}_{2,4}$is also positive. As

Table 12

Differences in the rate of change in value relevance over time for high and low change firms
$METRIC_t^g = \omega_{0,4} + \omega_{1,4}TIME_t + \omega_{2,4}CHANGE_{g,t} \cdot TIME_t + \varepsilon_{t,4}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\widehat{\omega}_{0,4}$	0.546	0.045	0.011
T-statistic	15.897	1.741	0.387
p-value	0.000	0.090	0.701
$\widehat{\omega}_{1,4}$	0.015	0.012	0.009
T-statistic	4.797	4.884	3.611
p-value	0.000	0.000	0.001
$\widehat{\omega}_{2,4}$	-0.005	0.001	0.001
T-statistic	-2.007	0.235	0.489
p-value	0.026	0.592	0.686

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,4}$ and $\widehat{\omega}_{1,4}$ equal to zero and one-sided for $\widehat{\omega}_{2,4}$ larger than zero) for a regression of coefficients of determination for the low and high change group against a time variable and a coefficient dummy variable. The coefficient dummy variable takes value one for coefficients of determination of the high change group and zero for coefficients of determination of the low change group.

was indicated in chapter six, the sum of $\widehat{\omega}_{1,4}$ and $\widehat{\omega}_{2,4}$ indicates the change in value relevance over time for the high change group. For the price change and return model value relevance has been increasing for both change groups, there is however no significant difference in the rate of change. Summarising, only weak evidence is found in support of hypothesis H4 and in line with the results of hypothesis H3, this evidence is based on the level model.

8.4.2 Residual dispersion metric

Table 13 shows the yearly abnormal pricing errors for the high and low change groups based on the three different regression models. For the level model it appears that high change firms have substantially higher abnormal pricing errors. Only in the years 1997, 1999 and 2000 are the abnormal pricing errors for firms experiencing little business change higher than for firms experiencing high amounts of business change. Finally, the average abnormal pricing errors for high change firms are larger than those of low change firms. All of this is in line with expectations. As was the case for the coefficient of determination metric, there does not seem to be a difference between abnormal pricing errors for high and low change firms for the price change or return model. What is interesting to note though, is that for the return model in the period 1992-2002 the abnormal pricing error is larger for

Table 13

Abnormal pricing errors for high and low change firms

	Level model (6.1)		Price change model (6.2)		Return model (6.3)	
Year	*High change*	*Low change*	*High change*	*Low change*	*High change*	*Low change*
1991	0.921	-1.925	-	-	-	-
1992	-0.550	-1.029	0.043	-0.003	0.043	-0.115
1993	-0.180	-4.851	0.078	0.071	0.078	-0.067
1994	-0.232	-4.931	-0.029	-0.007	-0.029	-0.074
1995	-1.470	-4.606	-0.047	0.027	-0.047	-0.061
1996	-0.103	-1.757	-0.044	0.019	-0.044	-0.062
1997	-0.528	-0.256	0.020	0.019	0.020	-0.044
1998	0.641	-2.307	0.100	0.131	0.100	-0.026
1999	2.669	2.687	0.104	0.071	0.104	0.109
2000	0.182	1.069	0.199	0.047	0.199	0.030
2001	-0.536	-1.759	0.020	0.049	0.020	-0.007
2002	-0.420	-0.804	-0.044	-0.052	-0.044	-0.052
2003	-1.336	-4.652	0.030	0.146	0.030	0.107
2004	-1.891	-3.741	-0.029	-0.086	-0.029	0.038
2005	-3.177	-3.996	-0.080	-0.116	-0.080	0.015
2006	-3.543	-4.671	-0.088	-0.078	-0.088	-0.044
2007	-4.670	-5.132	-0.136	-0.093	-0.136	-0.029
2008	-1.569	-5.007	-0.191	-0.152	-0.191	-0.080
2009	-3.999	-5.052	-0.083	0.001	-0.083	0.007
2010	-4.493	-5.082	-0.134	-0.092	-0.134	-0.065
2011	-4.005	-4.496	-0.248	-0.118	-0.248	-0.118
Average	**-1.347**	**-2.967**	**-0.028**	**-0.011**	**-0.028**	**-0.027**

Note: Abnormal pricing errors for firms classified as high change or low change. Classification of firms in each of these groups is based on their across time absolute rank change, or $ATARC_i$. Firms for which $ATARC_i > 0.1$ are classified as high change, while firms for which $ATARC_i \leq 0.1$ are classified as low change.

high change firms than for low change firms, but this finding is reversed in the period 2003-2011.

Results of a test of hypothesis H3 based on abnormal pricing errors are provided in table 14. These results closely resemble the results based on the coefficient of determination metric. For the level model, value relevance is significantly higher for the low change group than for the high change group. Also, the difference in value relevance between the groups is

Table 14

Differences in abnormal pricing errors of high change and low change firms

$METRIC_t^g = \omega_{0,3} + \omega_{1,3}CHANGE_{g,t} + \varepsilon_{t,3}$

	Level model (6.1)	**Price change model (6.2)**	**Return model (6.3)**
$\widehat{\omega}_{0,3}$	-2.970	-0.011	-0.027
T-statistic	-6.339	-0.501	-1.762
p-value	0.000	0.619	0.086
$\widehat{\omega}_{1,3}$	1.623	-0.017	0.005
T-statistic	2.449	-0.558	0.211
p-value	0.009	0.710	0.417

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,3}$ equal to zero and one-sided for $\widehat{\omega}_{1,3}$ smaller than zero) for a regression of abnormal pricing errors of the low and high change group against a dummy variable with value one for abnormal pricing errors of the high change group and zero for abnormal pricing errors of the low change group. The total number of observations in this regression is thus 42 for the level model and 40 for the change or return model. Note that a positive value of $\widehat{\omega}_{1,3}$ indicates higher value relevance for the low change group, which is the expectation.

more pronounced for the residual dispersion metric; the absolute T-statistic is 2.449 compared to an absolute T-statistic for the coefficient of determination metric of 2.007.

Finally, $\widehat{\omega}_{1,3}$ is not significantly larger than zero for the level and return model. This finding corresponds to the findings based on the coefficient of determination. It also supports the hypothesis that the differing results for the coefficient of determination metric on hypothesis H3 are not caused only by scale effects or across sample incomparability of this metric. In chapter nine a more thorough analysis of this inconsistency between the level model and the price change and return model is provided.

Differences in the rate of change in abnormal pricing errors are shown in table 15. Again these results closely mimic the results of the coefficient of determination metric. The results in table 15 should be interpreted carefully; a negative signs for $\widehat{\omega}_{1,4}$ indicates an increase in value relevance for low change firms, while the sum of $\widehat{\omega}_{1,4}$ and $\widehat{\omega}_{2,4}$ indicates the change in value relevance for high change firms. Hence if $\widehat{\omega}_{1,4} + \widehat{\omega}_{2,4} > \widehat{\omega}_{1,4}$, or equivalently if $\widehat{\omega}_{2,4} > 0$, value relevance has increased more for low change firms than for high change firms, which is line with hypothesis H4. Based on the level model, $\widehat{\omega}_{2,4}$ is significantly larger than zero with a T-statistic of 2.001 and a one-sided p-value of 0.026. Results for $\widehat{\omega}_{2,4}$ for the price change and return model do have the correct sign, but are not statistically significant. It is concluded that only weak evidence is available to support differing rates of change in value relevance across high and low change firms. It also appears that the inconsistent conclusions for hypothesis H4 based on the coefficient of determination metric, are not driven merely by the choice of metric.

Table 15

Differences in the rate of change in value relevance over time for high and low change firms

$METRIC_t^g = \omega_{0,4} + \omega_{1,4}TIME_t + \omega_{2,4}CHANGE_{g,t} \cdot TIME_t + \varepsilon_{t,4}$

	Level model (6.1)	**Price change model (6.2)**	**Return model (6.3)**
$\widehat{\omega}_{0,4}$	0.029	0.045	0.011
T-statistic	0.048	1.741	0.387
p-value	0.962	0.090	0.701
$\widehat{\omega}_{1,4}$	-0.245	0.012	0.009
T-statistic	-4.618	4.884	3.611
p-value	0.000	0.000	0.001
$\widehat{\omega}_{2,4}$	0.092	0.001	0.001
T-statistic	2.001	0.235	0.489
p-value	0.026	0.408	0.314

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,4}$ and $\widehat{\omega}_{1,4}$ equal to zero and one-sided for $\widehat{\omega}_{2,4}$ smaller than zero) for a regression of abnormal pricing errors for the low and high change group against a time variable and a coefficient dummy variable. The coefficient dummy variable takes value one for abnormal pricing errors of the high change group and zero for abnormal pricing errors of the low change group.

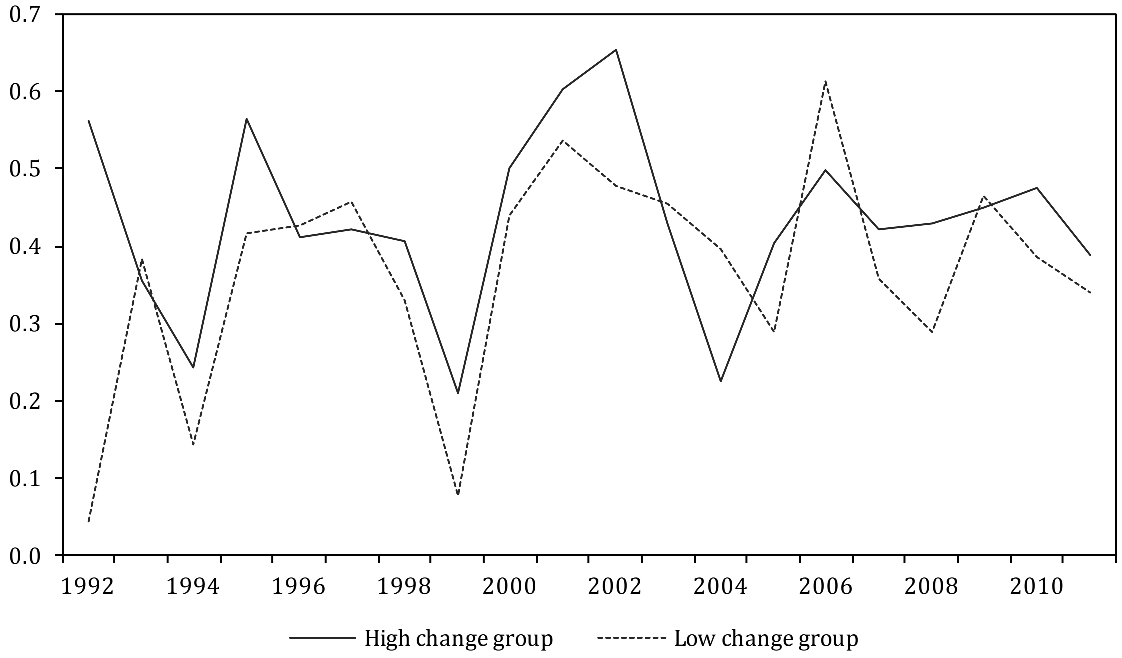
8.4.3 Portfolio metric

Figure seven depicts the temporal pattern of the portfolio metric for both the high and low change groups. It is the return that the accounting based hedge portfolio per group would earn, divided by the perfect foresight portfolio return of that group. The reason that the metrics of the two groups are completely separated is to correct for across sample differences. For example, one might expect that the high change firms on average earn higher returns than the low change firms. If the return on the combined perfect foresight portfolio is then used as a denominator for the portfolio metric, it is likely that the high change group will show higher value relevance. By completely separating the two groups and performing independent calculations, this incomparability between the groups is mitigated.

Figure seven shows that the percentage explained return is lower for the low change group than for the high change group in most of the sample years. This is contrary to what is expected. The average level of percentage explained returns is 43.3% for the high change group and 36.7% for the low change group. A one-sided test of value relevance being higher for high change firms than for low change firms yields a T-statistic of 1.595 and a p-value of 0.059. Concluding, the portfolio metric does not support the expectations based on hypothesis H3, but points weakly to the contrary.

Figure 7

Temporal pattern of the portfolio metric for high and low change firms



Note: This figure shows the temporal pattern of the portfolio metric for firms classified as high and low change. Firms are sorted in one of these groups based on the value of their across time absolute rank change, or $ATARC_i$. Firms for which $ATARC_i > 0.1$ are classified as high change, while firms for which $ATARC_i \leq 0.1$ are classified as low change.

Table 16

Differences in portfolio metric for high and low change firms

$METRIC_t^g = \omega_{0,3} + \omega_{1,3} CHANGE_{g,t} + \varepsilon_{t,3}$

$\widehat{\omega}_{0,3}$	0.367
T-statistic	12.481
p-value	0.000
$\widehat{\omega}_{1,3}$	0.067
T-statistic	1.595
p-value	0.941

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,3}$ equal to zero and one-sided for $\widehat{\omega}_{1,3}$ larger than zero) for a regression of the portfolio metric for the low and high change group against a dummy variable with value one for the portfolio metric of the high change group and zero for portfolio metric of the low change group. Note that a negative value of $\widehat{\omega}_{1,3}$ indicates higher value relevance for the low change group, which is the expectation.

Table 17 reports the results of a test of hypothesis H4 on differing rates of changes in value relevance. It can be seen that the sign of $\widehat{\omega}_{2,4}$ is the opposite of what is expected. A one-sided test if the rate of change in value relevance is larger for high change firms, i.e. testing the opposite of H4, also is not significant. It is concluded that there are no significant

Table 17

Differences in the rate of change in value relevance over time for high and low change firms
$METRIC_t^g = \omega_{0,4} + \omega_{1,4}TIME_t + \omega_{2,4}CHANGE_{g,t} \cdot TIME_t + \varepsilon_{t,4}$

$\widehat{\omega}_{0,4}$	0.354
T-statistic	8.069
p-value	0.000
$\widehat{\omega}_{1,4}$	0.003
T-statistic	0.684
p-value	0.498
$\widehat{\omega}_{2,4}$	0.003
T-statistic	0.884
p-value	0.809

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,4}$ and $\widehat{\omega}_{1,4}$ equal to zero and one-sided for $\widehat{\omega}_{2,4}$ larger than zero) for a regression of the portfolio metric for the low and high change group against a time variable and a coefficient dummy variable. The coefficient dummy variable takes value one for portfolio metrics of the high change group and zero for portfolio metrics of the low change group.

differences in value relevance based on the portfolio metric between firms experiencing high amounts of business change or low amounts of business change.

8.5 H5: The influence of IFRS

In this section hypothesis H5 is tested; whether value relevance has increased after the introduction of IFRS. To do so the different metrics for value relevance are regressed against an IFRS dummy variable. This dummy variable has a value one for the years 2005-2011 and zero otherwise. Table 18 presents the results for the coefficient of determination metric. Based on the results in table 18 it appears that the introduction of IFRS is connected to a significant increase in value relevance. For all three models $\widehat{\omega}_{1,5}$ is positive and has T-statistics of respectively 3.900, 2.133 and 2.435 for the level, price change and return model.

Increasing value relevance after the introduction of IFRS is also found for the residual dispersion metric. As can be seen in table 19, all estimates of $\widehat{\omega}_{1,5}$ are negative, indicating lower abnormal pricing errors after the introduction of IFRS and thus higher value relevance. These results are also statistically significant, as is evidenced by the T-statistics of -5.492, -5.125 and -2.483 for the level, price change and return model respectively.

Somewhat contrary to the findings above are those for the portfolio metric. It appears that no significant change has occurred in value relevance after the introduction of IFRS based on the portfolio metric. Results are shown in table 20. The coefficient estimate of

Table 18

IFRS and value relevance measured by coefficients of determination
$METRIC_t = \omega_{0,5} + \omega_{1,5} IFRS_t + \varepsilon_{t,5}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\hat{\omega}_{0,5}$	0.623	0.142	0.074
T-statistic	25.093	5.863	3.136
p-value	0.000	0.000	0.006
$\hat{\omega}_{1,5}$	0.168	0.087	0.097
T-statistic	3.900	2.133	2.435
p-value	0.001	0.024	0.013

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\hat{\omega}_{0,5}$ equal to zero and one-sided for $\hat{\omega}_{1,5}$ smaller than zero) for a regression of coefficients of determination against an IFRS dummy variable. Note that a positive value of $\hat{\omega}_{1,5}$ indicates higher value relevance after the introduction of IFRS.

Table 19

IFRS and value relevance measured by the residual dispersion metric
$METRIC_t = \omega_{0,5} + \omega_{1,5} IFRS_t + \varepsilon_{t,5}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\hat{\omega}_{0,5}$	-0.917	0.033	0.002
T-statistic	-2.793	1.692	0.104
p-value	0.012	0.108	0.918
$\hat{\omega}_{1,5}$	-3.122	-0.169	-0.075
T-statistic	-5.492	-5.125	-2.483
p-value	0.000	0.001	0.012

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\hat{\omega}_{0,5}$ equal to zero and one-sided for $\hat{\omega}_{1,5}$ larger than zero) for a regression of abnormal pricing errors against an IFRS dummy variable. Note that a negative value of $\hat{\omega}_{1,5}$ indicates higher value relevance after the introduction of IFRS.

the dummy variable, i.e. $\hat{\omega}_{1,5}$, is negative, but also not significantly different from zero, with a T-statistic of -0.0812. Nevertheless it is concluded that, based on an overall view of the results presented in this section, a significant increase in value relevance has occurred in the period 2005-2011. The most obvious explanation for this is the effect of the introduction of IFRS on value relevance.

Table 20

IFRS and value relevance measured by the portfolio metric
$METRIC_t = \omega_{0,5} + \omega_{1,5} IFRS_t + \varepsilon_{t,5}$

$\widehat{\omega}_{0,5}$	0.411
T-statistic	12.521
p-value	0.000
$\widehat{\omega}_{1,5}$	-0.005
T-statistic	-0.081
p-value	0.532

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,5}$ equal to zero and one-sided for $\widehat{\omega}_{1,5}$ smaller than zero) for a regression of the portfolio metric against an IFRS dummy variable. Note that a positive value of $\widehat{\omega}_{1,5}$ indicates higher value relevance after the introduction of IFRS.

8.6 H6: International differences in value relevance

8.6.1 Coefficient of determination metric

The cross-country sample of this research provides the opportunity to test for international differences in value relevance. Similarly to the procedure to test for hypothesis H3, the sample is split into two groups, only based on a different criterion. One group consist of firms in the Netherlands and the United Kingdom, the countries with a market-oriented financial system, while the other group consisting of France and Germany represent countries with a bank-oriented system. Without attempting to prove full causal relations, it is conjectured in this research that the countries with a market-oriented financial system will have higher value relevance than the countries with a bank-oriented financial system.

Table 21 summarises estimates of coefficients of determination over the sample period for the two groups and different models. Contrary to expectations, for all three models the average coefficient of determination is higher for bank-oriented financial systems than for market-oriented financial systems. These differences are also significant for the price change and return model as indicated by the T-statistics in table 22. If the opposite of hypothesis H6 is tested, namely that bank-oriented financial systems have higher value relevance than market-oriented financial systems, p-values of 0.000 and 0.037 for respectively the price change and return model would result. For the level model there is no significant difference between the estimated coefficients of determination for bank and market-oriented financial systems. Based on these results is seems that market-oriented countries have significantly higher value relevance measured by the coefficient of determination metric in the sample period.

Table 21

Coefficient of determination estimates for bank and market-oriented financial systems

	Level model (6.1)		**Price change model (6.2)**		**Return model (6.3)**	
Year	*Bank*	*Market*	*Bank*	*Market*	*Bank*	*Market*
1991	0.176	0.368	-	-	-	-
1992	0.469	0.283	0.252	0.038	0.011	0.042
1993	0.356	0.369	0.388	0.049	0.201	0.036
1994	0.384	0.375	0.399	0.040	0.085	0.004
1995	0.683	0.354	0.336	0.091	0.215	0.086
1996	0.366	0.390	0.404	0.089	0.104	0.053
1997	0.718	0.513	0.316	0.101	0.180	0.045
1998	0.498	0.474	0.206	0.063	0.069	0.024
1999	0.359	0.321	0.229	0.066	0.099	0.010
2000	0.456	0.353	0.384	0.111	0.241	0.054
2001	0.452	0.519	0.349	0.248	0.015	0.025
2002	0.628	0.620	0.415	0.286	0.267	0.197
2003	0.656	0.623	0.233	0.174	0.237	0.216
2004	0.661	0.632	0.205	0.192	0.018	0.001
2005	0.734	0.685	0.184	0.283	0.042	0.009
2006	0.803	0.738	0.248	0.225	0.190	0.148
2007	0.792	0.783	0.211	0.180	0.140	0.129
2008	0.660	0.634	0.187	0.150	0.144	0.128
2009	0.781	0.771	0.304	0.318	0.322	0.306
2010	0.809	0.796	0.277	0.324	0.239	0.305
2011	0.846	0.753	0.244	0.102	0.242	0.139
Average	**0.585**	**0.541**	**0.289**	**0.156**	**0.153**	**0.098**

Note: Coefficients of determination estimates for bank and market-oriented financial systems. Germany and France are classified as having a bank-oriented financial system, while the Netherlands and the United Kingdom are classified as having a market-oriented financial system.

8.6.2 Residual dispersion metric

In this section international differences in value relevance are compared based on abnormal pricing errors, or equivalently the residual dispersion metric. First of all, table 23 summarises yearly abnormal pricing errors for the two subgroups of countries.

Table 22

Differences in coefficients of determination for bank and market-oriented financial systems
$METRIC_t^c = \omega_{0,6} + \omega_{1,6} COUNTRY_{c,t} + \varepsilon_{t,6}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\widehat{\omega}_{0,6}$	0.585	0.289	0.153
T-statistic	14.648	14.630	7.227
p-value	0.000	0.000	0.000
$\widehat{\omega}_{1,6}$	-0.0045	-0.132	-0.055
T-statistic	-0.787	-4.736	-1.842
p-value	0.782	1.000	0.963

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,6}$ equal to zero and one-sided for $\widehat{\omega}_{1,6}$ smaller than zero) for a regression of coefficients of determination for bank and market-oriented financial systems against a dummy variable with value one for market-oriented financial systems and zero for coefficients of determination of bank-oriented financial systems. Note that a positive value of $\widehat{\omega}_{1,6}$ indicates higher value relevance for market-oriented financial systems, which is the expectation.

It can be seen from table 23 that the differences in value relevance between bank and market-oriented financial systems observed for the coefficient of determination metric are not present for the residual dispersion metric. For the level model the average abnormal pricing error is smaller over the entire period for the countries with a bank-oriented financial system, but on a year-by-year comparison only for 12 out of the 21 years value relevance is actually higher in these countries. This is in sharp contrast with the coefficient of determination metric, for which value relevance is higher in bank-oriented countries for 17 out of the 21 years. For the price change model, value relevance is higher in market-oriented countries for 11 out of the 20 sample years. This is an even more striking finding, as value relevance measured by the coefficient of determination metric was found to be significantly higher for bank-oriented countries than for market-oriented countries.

Table 23

Abnormal pricing errors for bank and market-oriented financial systems

	Level model (6.1)		Price change model (6.2)		Return model (6.3)	
Year	*Bank*	*Market*	*Bank*	*Market*	*Bank*	*Market*
1991	0.685	0.798	-	-	-	-
1992	-2.554	0.554	-0.009	0.075	-0.052	-0.073
1993	-4.672	0.212	-0.097	0.026	0.031	-0.026
1994	-5.029	0.286	-0.095	-0.035	-0.154	-0.078
1995	-3.208	0.016	-0.069	-0.044	-0.073	-0.064

Table 23 (cont'd)

Abnormal pricing errors for bank and market-oriented financial systems

	Level model (6.1)		Price change model (6.2)		Return model (6.3)	
Year	*Bank*	*Market*	*Bank*	*Market*	*Bank*	*Market*
1996	-0.762	-0.744	-0.049	-0.038	-0.048	-0.017
1997	2.849	-0.654	-0.028	0.008	-0.059	-0.007
1998	-0.292	0.986	0.127	0.068	0.022	0.034
1999	3.615	1.264	0.203	0.106	0.162	0.102
2000	1.718	-0.269	0.213	0.080	0.045	0.057
2001	-0.001	0.490	0.092	0.010	0.038	0.025
2002	-0.624	0.242	-0.033	-0.039	-0.058	-0.058
2003	-1.387	-0.363	0.036	0.122	0.052	0.106
2004	-0.333	-0.503	0.029	-0.098	0.052	0.016
2005	-1.396	-1.873	-0.022	-0.118	0.013	0.041
2006	-1.980	-2.243	-0.014	-0.118	-0.050	-0.060
2007	-3.458	-2.563	-0.120	-0.083	-0.047	0.005
2008	-0.651	-0.380	-0.153	-0.231	-0.111	-0.356
2009	-1.850	-2.660	-0.017	0.041	-0.114	0.025
2010	-3.265	-2.550	-0.093	-0.173	-0.054	-0.055
2011	-2.861	-5.690	-0.174	-0.182	-0.100	-0.106
Average	**-1.212**	**-0.745**	**-0.013**	**-0.031**	**-0.025**	**-0.024**

Note: In this table abnormal pricing errors for bank and market-oriented financial systems are shown. Germany and France are classified as having a bank-oriented financial system, while the Netherlands and the United Kingdom are classified as having a market-oriented financial system.

In table 24 results from a regression of the residual dispersion metric against a country dummy variable are presented. These results confirm the initial findings based on an inspection of table 23. Although $\widehat{\omega}_{1,6}$ has a different sign for the level and return model than expected based on hypothesis H6 they are not significantly different from zero with T-statistics of 0.766 and 0.025. For the price change model $\widehat{\omega}_{1,6}$ does have the correct sign, unfortunately however it is also not significantly different from zero. For the level model $\widehat{\omega}_{1,6}$ has a T-statistic of -0.533 and a p-value of 0.299 for a one-sided test under the null hypothesis that $\omega_{1,6}$ is equal or larger than zero. As such, H6 is rejected based on the data. In the subsequent chapter the differing results between the coefficient of determination and residual dispersion metric will be further investigated.

Table 24

Differences in abnormal pricing errors for bank and market-oriented financial systems
$METRIC_t^c = \omega_{0,6} + \omega_{1,6} COUNTRY_{c,t} + \varepsilon_{t,6}$

	Level model (6.1)	Price change model (6.2)	Return model (6.3)
$\widehat{\omega}_{0,6}$	-1.212	-0.014	-0.025
T-statistic	-2.811	-0.592	-1.293
p-value	0.008	0.558	0.204
$\widehat{\omega}_{1,6}$	0.467	-0.017	0.001
T-statistic	0.766	-0.533	0.025
p-value	0.776	0.299	0.510

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,6}$ equal to zero and one-sided for $\widehat{\omega}_{1,6}$ larger than zero) for a regression of abnormal pricing errors for bank and market-oriented financial systems against a dummy variable with value one for market-oriented financial systems and zero for coefficients of determination of bank-oriented financial systems. Note that a negative value of $\widehat{\omega}_{1,6}$ indicates higher value relevance for market-oriented financial systems, which is the expectation.

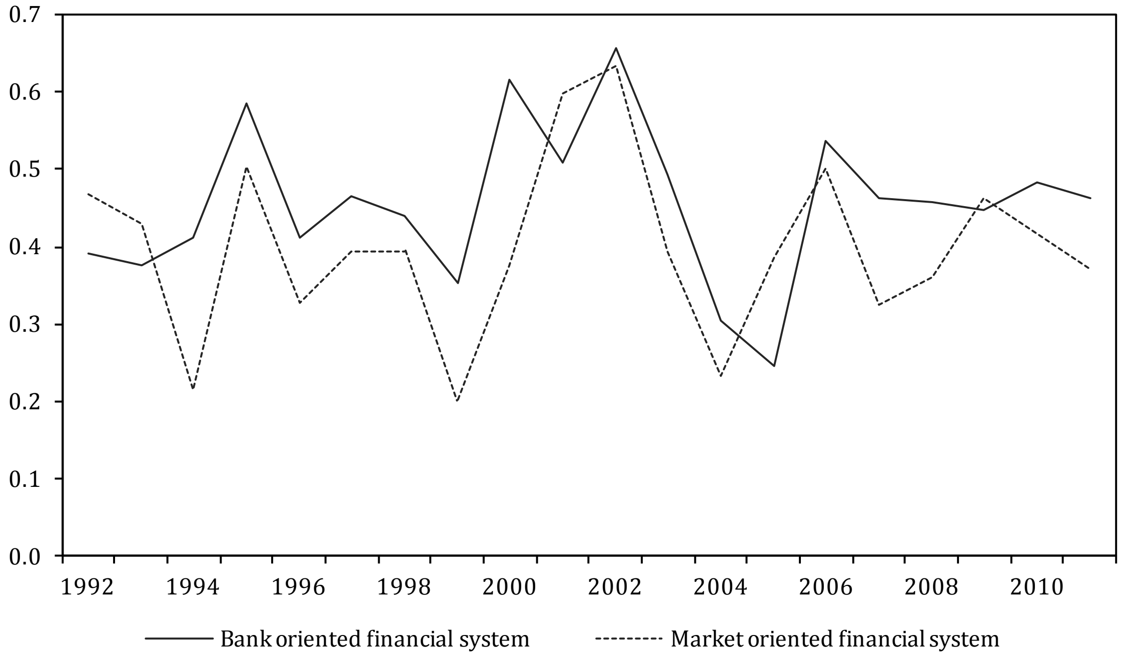
8.6.3 Portfolio metric

Figure eight shows the percentage return that could be earned based on perfect foreknowledge of accounting information for the bank and market-oriented financial systems over time. It can be seen from this figure that the temporal pattern of value relevance is comparable across these two groups of countries. Furthermore, on a first glance it appears that bank-oriented financial systems have higher value relevance measured by the portfolio metric. A further investigation of the data confirms this suspicion; value relevance is higher in the countries with a bank-oriented financial system for 15 of the 20 years in this sample. The portfolio metric is on average 45.6% for France and Germany compared to 40.0% for the Netherlands and the United Kingdom.

This difference in averages is not significant in a one-sided test under the null hypothesis that $\omega_{1,6}$ is equal or larger than zero, but only barely with a p-value of 0.053. The converse, a one-sided test of hypothesis H6, i.e. under the null hypothesis that $\omega_{1,6}$ is smaller or equal than zero, yields a p-value of 0.947 as shown in table 25. Based on the portfolio metric it appears that value relevance is higher in the market-oriented countries, a finding that corresponds with the results based on the coefficient of determination.

Figure 8

Temporal pattern of the portfolio metric for bank and market-oriented financial systems



Note: This figure shows the temporal pattern of the portfolio metric for firms operating in a bank-oriented or market-oriented financial system. Of the total sample, the Netherlands and the United Kingdom were classified as having market-oriented financial systems, while France and Germany were classified as having bank-oriented financial systems.

Table 25

Differences in the portfolio metric for bank and market-oriented financial systems

$METRIC_t^c = \omega_{0,6} + \omega_{1,6}COUNTRY_{c,t} + \varepsilon_{t,6}$

$\widehat{\omega}_{0,5}$	0.456
T-statistic	19.220
p-value	0.000
$\widehat{\omega}_{1,5}$	-0.055
T-statistic	-1.654
p-value	0.947

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided for $\widehat{\omega}_{0,6}$ equal to zero and one-sided for $\widehat{\omega}_{1,6}$ smaller than zero) for a regression of the portfolio metric for bank and market-oriented financial systems against a dummy variable with value one for market-oriented financial systems and zero for bank-oriented financial systems. Note that a positive value of $\widehat{\omega}_{1,6}$ indicates higher value relevance for market-oriented financial systems, which is the expectation.

8.7 Summary

In this chapter the results of this research were presented. As was discussed in chapter six, a number of metrics for value relevance has been used across a number of relationships to test the results for robustness and for possible econometrical issues errors in e.g. the coefficient of determination metric. Table 26 provides a short overview of all results presented in this chapter, by indicating the sign of the estimated variables of interest, as well as the expected signs of these variables.

Table 26

Summary of results

	Coefficient of determination metric				**Residual dispersion metric**				**Portfolio metric**	
Model	*(6.1)*	*(6.2)*	*(6.3)*	𝔼[…]	*(6.1)*	*(6.2)*	*(6.3)*	𝔼[…]	*(6.6)*	𝔼[…]
Hypothesis H1	+***	+***	+***	+/-	-***	-***	-	+/-	+	+/-
Hypothesis H2		+		+		+		+	+	+
Hypothesis H3	-**	-	-	-	+***	-	+	+	+	-
Hypothesis H4	-**	+	+	-	+**	+	+	+	+	-
Hypothesis H5	+***	+**	+**	+	-***	-***	-**	-	-	+
Hypothesis H6	-	-[a]	-[b]	+	+	-	+	-	-[c]	+

Notes: * indicates significance at 10% level, ** at a 5% level and *** at a 1% level. Where applicable results are based on one-sided tests.
[a] This finding is significant at a 1% level, but is contrary to expectations.
[b] This finding is significant at a 5% level, but is contrary to expectations.
[c] This finding is significant at a 10% level, but is contrary to expectations.

Hypothesis H1 questions whether any significant change has occurred over time in value relevance. Table 26 indicates that there is substantial evidence indicating an increase in value relevance. Contrary to expectation an increase in the rate of business change is not established for this sample. Hypothesis H3 states that value relevance should be higher for firms experiencing low amounts of business change. This result is established for the level model, but not for the price change and return model. Similar results are obtained for hypothesis H4, which states that the change in value relevance should be lower for firms experiencing high amounts of business change. Table 26 also shows that substantial evidence is found for an increase in value relevance over the period 2005-2011. This period coincides with the mandatory introduction of IFRS in the sample countries. Finally, it was expected that value relevance would be higher in the Netherlands and the United Kingdom than in France and Germany. This result is not established, and if anything, results based on the coefficient of determination metric and portfolio metric indicate the opposite.

One might wonder what drives the differing results for the model specifications for hypotheses H3 and H4, and the differing results for the metrics of value relevance for hypothesis H6. These issues will be explored more thoroughly in the subsequent chapter.

9. Robustness and interpretation

9.1 Introduction

In this chapter the result established in the previous chapter will be analysed further and compared to previous literature. The objective of this analysis is to scrutinise the results of chapter eight so that stronger inferences can be drawn. Also, as a robustness check, an additional variable will be introduced, namely the percentage of loss making firms in a sample year. Collins et al. (1997) found that this variable has a significant impact on value relevance and it is worthwhile to consider if it also influences value relevance in this European sample. Furthermore, for hypothesis H3, H4 and H6 the sample was split into groups. It might be that the results for these hypotheses are not attributable to the factor considered, i.e. business change or orientation of financial system, but by differences in the two samples, such as the percentage of loss making firms. Finally, it should be noted that only the results of the coefficient of determination metric and residual dispersion metric are tested for robustness. The results for the portfolio metric are shortly discussed separately.

The remainder of this chapter is organised as follows. Section 9.2 briefly discusses the control variable 'percentage losses'. In section 9.3 hypothesis H1 and H5 are further discussed, because of their interrelatedness. In section 9.4 the results for hypothesis H3 and H4 concerning business change are analysed. Section 9.5 discusses the results for hypothesis H6 on international differences in value relevance. In section 9.6 the results for the portfolio metric are shortly analysed and in section 9.7 a simple robustness test is performed on the outlier removal procedure. Section 9.7 summarises.

9.2 Percentage losses

Collins et al. (1997) investigate changes in value relevance over time and propose four different factors that might have driven these changes, which are changes in the amount of intangible and technology intensive firms in the sample, nonrecurring items, negative earnings and size. Following Lev and Zarowin (1999), of these four factors, only negative earnings are considered as a control variable for the results of this research. As was previously discussed, Lev and Zarowin (1999) argue that both nonrecurring items and negative earnings are a symptom of business change. However, it is possible that this is not the case for this European sample, especially since no increase in the rate of business change is found.

Furthermore, Collins et al. (1997) find that negative earnings increase the value relevance of earnings and book values combined. Given the increasing value relevance found in this research combined with the economic crises in the sample period, i.e. the collapse of the dot-com bubble in 2000-2001, the 9/11 events in 2001 and the global financial crisis of 2007-2012, it might be possible that negative earnings drive the results of chapter eight. Finally, the results of hypotheses H3, H4 and H6 might be caused by differences in the amount of negative earnings reported for the different groups. For all these reasons it is important to test the results for robustness against the frequency of negative earnings reported.

The first step is to define the control variable for negative earnings, which is rather straightforward. For each year and each group, the percentage of firms reporting negative earnings is calculated and this is the control variable $\%LOSSES_t$ or for the change and country groups respectively: $\%LOSSES_t^g$ and $\%LOSSES_t^c$ with $g \in \{H, L\}$ and $c \in \{bank, market\}$ indicating group membership.

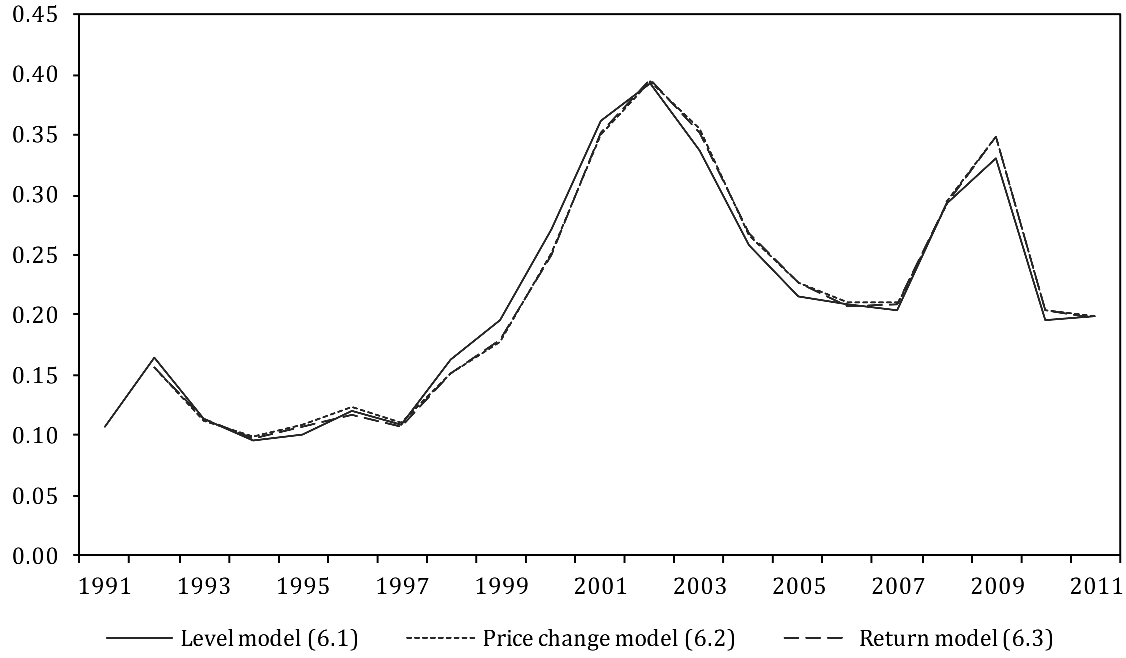
The temporal pattern of the percentage of firms reporting negative earnings for the different samples of the level, price change and return model are depicted in figure eight. From this figure it appears that the frequency with which firms report negative earnings has increased and that this increase coincides partly with the increase in value relevance found in the post-IFRS period. Finally, note that there patterns for the different samples, i.e. the level model, price change model and return model samples, are very similar.

9.3 Changes in value relevance over time

In this section the results of hypothesis H1 and H5 are further investigated. The first interesting point to make is that the results of hypothesis H5 do not necessarily indicate that IFRS was the causing factor for increasing value relevance. If for example, there has been an increasing trend in value relevance over the entire sample period, to which the results of hypothesis H1 hint, then it is very likely that an increase in value relevance will be found for the final years of the sample period. Furthermore, as was discussed in the previous section, it appears that the final sample years were characterised by a high number of firms reporting losses, also a factor found to increase value relevance.

Table 27 reports regression results for a regression of coefficients of determination and abnormal pricing errors against a time variable, but this time for the sub-period 1991-2004. As can be seen from this table, only for the coefficient of determination metric for the price change model a significant change in value relevance is found. This indicates that the

Figure 8
Temporal pattern of percentage of firms reporting negative earnings ($\%LOSSES_t$)



Note: This figure shows the temporal pattern of the percentage of firms reporting losses for the different samples. It can be seen that relatively a lot of firms report losses following the collapse of the dot-com bubble and during the initial years of the global financial crisis.

results for hypothesis H5 do not seem to be driven by an underlying increase in value relevance over the entire sample period.

Table 28 reports results of the following regression:

$$METRIC_t = \omega_{0,7} + \omega_{1,7}TIME_t + \omega_{2,7}IFRS_t + \omega_{3,7}\%LOSSES_t + \varepsilon_{t,7} \qquad (9.1)$$

By testing this hypothesis, both changes in the tendency of firms to report negative earnings as well as underlying changes in value relevance over time are controlled for. A significantly positive estimate of $\omega_{2,7}$, would indicate that value relevance has increased after the mandatory introduction of IFRS in 2005, despite changes in these other variables. As can be seen from table 28, results of the coefficient of determination metric for the IFRS dummy become insignificant. For the residual dispersion metric, results are also not significant, but only slightly so with T-statistics of -1.692, -1.622 and -1.761 for respectively the level, price change and return model. Also, based on the absolute value of the coefficient estimates and their p-values, it appears that the introduction of IFRS is the variable which most likely has influenced value relevance for the residual dispersion metric. Finally, note that for table 28, in contrast to some of the previous tables, p-values are reported for a two-sided test against a null-hypothesis of equality to zero.

Table 27

Changes in value relevance over 1991-2004

$METRIC_t = \omega_{0,1} + \omega_{1,1}TIME_t + \varepsilon_{t,1}$

Model	Coefficient of determination metric			Residual dispersion metric		
	Level model	*Price change model*	*Return model*	*Level model*	*Price change model*	*Return model*
$\widehat{\omega}_{0,1}$	0.622	-0.003	0.011	-1.011	0.034	-0.049
T-statistic	9.983	-0.010	0.266	-1.386	0.741	-1.340
p-value	0.000	0.924	0.796	0.191	0.475	0.207
$\widehat{\omega}_{1,1}$	0.000	0.021	0.009	0.013	0.000	0.007
T-statistic	0.016	4.746	1.735	0.147	-0.011	1.577
p-value	0.987	0.001	0.111	0.885	0.992	0.143

Notes: The table provides coefficient estimates, T-statistics and p-values (two-sided for a null hypothesis of equal to zero) for regressions of coefficients of determination and abnormal pricing error estimates against a time variable. The level model contains abnormal pricing error estimates for the years 1991 up to 2004 and as such $TIME_t = 1, \ldots, 24$. The price change and return model are estimated yearly for the period 1992 up to 2004 and for these models $TIME_t = 1, \ldots, 13$. Note that a positive (negative) value for $\widehat{\omega}_{1,1}$ indicates an increase in value relevance for the coefficient of determination (residual dispersion) metric.

The results presented in this section, as well as in chapter eight, indicate an increasing trend in value relevance for the sample countries in this research. The most likely cause for this appears to be the introduction of IFRS. These results are in contrast to those found for the U.S. in previous research as discussed in chapter four. For the U.S. a decline in value relevance under Interpretation 4 is predominantly established, see e.g. Chang (1999), Brown et al. (1999) and Gu (2007). As was previously noted, increasing value relevance is also found in the U.S. by some researchers, but this appears to be driven by methodological issues, see Brown et al. (1999) and Gu (2007).

There are a number of factors that might explain these differing results. First of all, the sample period used in this research differs from those used for the U.S. researches; none of the U.S. researches discussed has a sample period beyond 1996. Also, where in the U.S. declining value relevance appears to be caused by an increase in the rate of business change over the sample period (Lev and Zarowin, 1999), such an increase is not found for this European sample over the period 1991-2011. Without an increase in the rate of business change it might also be that there has not been a decrease in value relevance. Finally, the increase in value relevance for this research seems to be closely related to the introduction of IFRS. Although this is only the case for the residual dispersion metric, based on the econometrical problems associated with the coefficient of determination metric, these results are considered most reliable.

Table 28

Changes in value relevance after the mandatory introduction of IFRS

$METRIC_t = \omega_{0,7} + \omega_{1,7}TIME_t + \omega_{2,7}IFRS_t + \omega_{3,7}\%LOSSES_t + \varepsilon_{t,7}$

	Coefficient of determination metric			**Residual dispersion metric**		
Model	*Level model*	*Price change model*	*Return model*	*Level model*	*Price change model*	*Return model*
$\widehat{\omega}_{0,7}$	0.638	-0.028	-0.028	-1.326	0.038	-0.030
T-statistic	11.596	-0.931	-0.594	-1.865	0.855	-0.758
p-value	0.000	0.366	0.561	0.080	0.405	0.458
$\widehat{\omega}_{1,7}$	0.009	0.006	0.008	-0.125	-0.005	0.004
T-statistic	0.983	1.179	1.009	-1.108	-0.722	0.662
p-value	0.339	0.256	0.328	0.283	0.481	0.517
$\widehat{\omega}_{2,7}$	0.092	0.006	0.011	-2.050	-0.122	-0.119
T-statistic	0.978	0.127	0.142	-1.692	-1.622	-1.761
p-value	0.342	0.901	0.889	0.109	0.124	0.097
$\widehat{\omega}_{3,7}$	-0.401	0.631	0.233	6.751	0.157	0.008
T-statistic	-1.160	3.533	0.833	1.510	0.580	0.031
p-value	0.262	0.003	0.417	0.149	0.570	0.976

Notes: This table provides coefficient estimates, T-statistics and p-values (two-sided for all coefficient estimates against a null hypothesis of equal to zero) for regressions of coefficients of determination and abnormal pricing error estimates against a time variable, an IFRS dummy variable and a control variable for the percentage of firms reporting negative earnings.

9.4 Business change and value relevance

An interesting finding concerning hypothesis H3 and H4 are the contrary results for the level model compared to the price change and return model. As was shown in chapter eight, for the latter two models there does not appear to be a significant difference between the value relevance of low and high change firms. This is especially puzzling since it is the case for both the coefficient of determination metric and residual dispersion metric. One might initially think that the difference could be caused by scale effects, and the fact that the level model does not properly correct for these. However, for the residual dispersion metric, the level model does not suffer from scale effects and thus this does not appear to be the case.

Gu (2007) finds differing results between a level model and a return model for changes in value relevance over time. He states that this difference might be due to the fact that the return models overcompensate for scale effects. Note that the price change and return model are basically similar to the level model only the variables are proportionally scaled down. While it is true that firms with a large scale, c.q. price per share, have higher pricing errors they are not proportional. If the proportionally small pricing errors of large scale firms are then deflated by the high share prices of these firms, relatively small pricing errors result. Hence, if the low change firms would have a relatively small scale, this could explain the difference in results across model specifications.

A test of differences in scale does not indicate that this is the case. In a manner similar to testing hypothesis H3 and H6, the average scale per year in the high and low change group was regressed against a dummy variable with value one for group membership to the high change firms. Results (not shown) indicate a coefficient estimate of -6.640 and a T-statistic of -5.619 for the dummy variable. Average scale per year was measured by taking the mean absolute fitted value of the level model. As was shown in chapter eight, the samples across the models are very similar and it leads to the conclusion that firms in the low change groups have on average a higher scale. This is also in line with what one might expect a priori.

An additional possible explanation for not finding higher value relevance for the low change firms is that the high change firms report on average more losses. As this might increase the combined value relevance of earnings and book values, not correcting for this factor could lead to insignificant results for hypothesis H3. Note that this finding would indicate a model misspecification for all models, which could, but not necessarily, explains the difference in results across models.[42]

The results of a test in differences in the percentage of firms reporting losses between the low and high change group are depicted in table 29. This table indicates that for all three samples, the a firm in the high change group is more likely to report a loss, but this difference is not significant for the return model, with a T-statistics of 1.297, and also not for the price change model with a T-statistic of 1.579, but only slightly so.

In table 30 the results of a regression of metrics of value relevance against the yearly percentage of firms reporting losses and the change dummy variable are presented. First of all, it should be noted that except for the residual dispersion metric of the return model, all estimates have the expected sign. For the coefficient of determination metric the results indicate that for two out of the three models the difference in value relevance between high and low change firms appears to be caused by the number of firms reporting losses. For the residual dispersion metric the level and price change model indicate that different levels of business change experienced by firms leads to different levels of value relevance after controlling for differences in number of negative earnings reported. Firms experiencing high

[42] Models misspecification leads to biased and inconsistent parameter estimates (Greene, 2008), and this might explain the differing results if the price change and return models suffer more from this misspecification.

amounts of business change have lower value relevance and these results are statistically significant with T-statistics of 2.674 and 2.378 for the level and price change model respectively. Note that these results are stronger than those reported in chapter eight. Based on the results presented here, hypothesis H3 is not rejected. Finally, these findings are in line with the findings of Lev and Zarowin (1999) for a U.S. sample.

Table 29

Percentage of high and low change firms reporting losses

$\%LOSSES_t^g = \omega_{0,8} + \omega_{1,8}CHANGE_{g,t} + \varepsilon_{t,8}$

	Level model (6.1)	**Price change model (6.2)**	**Return model (6.3)**
$\widehat{\omega}_{0,8}$	0.162	0.181	0.186
T-statistic	8.009	8.403	8.566
p-value	0.000	0.000	0.000
$\widehat{\omega}_{1,8}$	0.063	0.048	0.040
T-statistic	2.219	1.579	1.297
p-value	0.032	0.123	0.203

Note: This table presents coefficient estimates, T-statistics and p-values (two-sided under the null hypothesis of equal to zero) for a regression of the percentage firms in the low and high group that report losses against a dummy variable for group membership to the high or low change group.

Table 30

Differences in value relevance for high and low change firms

$METRIC_t^g = \omega_{0,9} + \omega_{1,9}CHANGE_{g,t} + \omega_{2,9}\%LOSSES_{g,t} + \varepsilon_{t,9}$

	Coefficient of determination metric			**Residual dispersion metric**		
Model	*Level model*	*Price change model*	*Return model*	*Level model*	*Price change model*	*Return model*
$\widehat{\omega}_{0,9}$	0.682	0.007	0.013	2.323	-2.531	-0.043
T-statistic	15.029	0.280	0.370	-3.082	-3.078	-1.280
p-value	0.000	0.781	0.714	0.004	0.004	0.208
$\widehat{\omega}_{1,9}$	-0.084	-0.033	-0.016	1.872	1.688	-0.004
T-statistic	-1.986	-1.570	-0.556	2.674	2.378	-0.154
p-value	0.054	0.125	0.582	0.011	0.023	0.879
$\widehat{\omega}_{2,9}$	0.190	0.884	0.534	-3.983	-2.690	0.088
T-statistic	0.863	8.266	3.573	-1.088	-0.735	0.596
p-value	0.393	0.000	0.001	0.283	0.467	0.554

Notes: This table provides coefficient estimates, T-statistics and p-values (two-sided for all coefficient estimates against a null hypothesis of equal to zero) for regressions of coefficients of determination and abnormal pricing error estimates against a change dummy variable and a control variable for the percentage of firms reporting negative earnings.

9.5 International differences in value relevance

At first sight, the most surprising finding of this research is perhaps that firms in market-oriented countries do not have higher value relevance than firms in bank-oriented countries. Moreover, if anything, the coefficient of determination metric indicates that value relevance is higher in bank-oriented countries. This research is not the first to obtain this sort of results. As was discussed in chapter five, Joos and Lang (1994) find that value relevance is higher in France and Germany than in the United Kingdom if value relevance is measured by the coefficients of determination of the level model. With a similar methodology this result is also found by Joos (1997) and Arce and Mora (2002). It should be noted that Arce and Mora (2002) even extend the number of sample countries in their research to eight. Hence, the findings of this research are not far off from the findings of previous researchers.

The question arises why other researcher such as Alford et al. (1993) and Ali and Hwang (2000) are able to establish differences in value relevance between market and bank-oriented financial systems. Partly this might be explained by the research methodology; Alford et al. (1993) use the portfolio metric only and Ali and Hwang (2000) use a matched sample for their coefficients of determination. Also, both researchers examine a more extended set of countries, which make inferences more reliable.

Finally, it is possible that in this dual country system other factors are overlooked. As was shown in chapter eight, no significant difference in value relevance of the Netherlands and the United Kingdom is found for the residual dispersion metric. This automatically raises the question of scale effects. Regression analysis (not shown) indicates that there are significant differences in average scale between firms in market-oriented and bank-oriented financial systems. For a regression of mean absolute fitted prices based on the level model against a dummy variable indicating country group membership a coefficient estimate of -15.145 results. This result is highly significant with a T-statistic of -11.639 and it is concluded that firms in market-oriented financial systems have a substantially lower scale. Based on the discussion in the previous section, it might be the case that the results of the price change and return model are caused by overcompensating non-proportional scale by deflating variables. This hypothesis is however not investigated further.

Another influencing factor might be the percentage of firms reporting losses. Results of regression analysis (not shown) on differences in the percentage of firms reporting losses indicate however that there is no significant difference in this variable for the countries with a market or bank-oriented financial system. Including a control variable for the percentage of firms reporting losses in regression (6.12) does not qualitatively alter the results of hypothesis H6 as is shown in table 31. Since the residual dispersion metric is leading, it is concluded that no significant differences exist in the value relevance of the Netherlands and the United Kingdom compared to that in France and Germany.

Table 31

Differences in value relevance for firms in market and bank-oriented financial systems

$METRIC_t^c = \omega_{0,10} + \omega_{1,10}COUNTRY_{c,t} + \omega_{2,10}\%LOSSES_{c,t} + \varepsilon_{t,10}$

	Coefficient of determination metric			**Residual dispersion metric**		
Model	*Level model*	*Price change model*	*Return model*	*Level model*	*Price change model*	*Return model*
$\widehat{\omega}_{0,10}$	0.433	0.207	0.067	-1.298	-0.019	-0.037
T-statistic	6.577	6.081	1.832	-1.668	-0.441	-0.998
p-value	0.000	0.000	0.075	0.103	0.662	0.325
$\widehat{\omega}_{1,10}$	-0.062	-0.141	-0.064	0.457	-0.018	-0.004
T-statistic	-1.169	-5.472	-2.287	0.736	-0.541	-0.016
p-value	0.249	0.000	0.028	0.466	0.592	0.988
$\widehat{\omega}_{2,10}$	0.770	0.395	0.413	0.434	0.027	0.056
T-statistic	2.794	2.850	2.773	0.133	0.153	0.373
p-value	0.008	0.007	0.009	0.895	0.879	0.711

Notes: This table provides coefficient estimates, T-statistics and p-values (two-sided for all coefficient estimates against a null hypothesis of equal to zero) for regressions of coefficients of determination and abnormal pricing error estimates against a country dummy variable and a control variable for the percentage of firms reporting negative earnings.

9.6 Portfolio metric

From the summary table in section 8.7 it can be seen that none of the results of the portfolio metric is very significant. This result is somewhat puzzling if compared to previous research. In this section a number of possible explanations for the insignificant results of the portfolio metric are considered.

First of all, there are some differences in which the portfolio metric is calculated compared to Francis and Schipper (1999). Francis and Schipper use a different estimate for the market return, which they define as the equally weighted return on alls assets in their sample. In this research an ‘external’ market return is used in the form of the return on the Dow Jones STOXX Europe 600 Index. Other differences relate to the return window and the calculation of the return on the perfect foresight hedge portfolio.

In this research a return window of 12 months has been applied, i.e. from three months after fiscal year start, up to three months after fiscal year end. This contrast to the 15 month return window used by Francis and Schipper (1999). Note however that the choice of a return window appears somewhat arbitrary in the literature; although it is common to incorporate returns after fiscal year end, the number of months that are incorporated varies across researches. Ely and Waymire (1999) use for example 16 month return windows. The difference in this research is that the first months of returns are not included, similar to the approach followed by Gu (2007).

Finally, the calculation of the return on the perfect foresight hedge portfolio is different; Francis and Schipper's (1999) perfect hedge portfolio takes a long (short) position in all assets that have a positive (negative) return. In this research a long (short) position is taken in the assets with the top (bottom) 40% of returns. These differences individually do not appear to be very substantial, but together they might explain the insignificant result.

A final possible explanation might be that the portfolio metric, like the coefficient of determination metric, is sensitive to sample properties (Gu, 2007). This topic was briefly touched upon in chapter three and an illustration of the problem is found in Gu (2007). Gu (2007) notes that if the range of returns increases over the years, the return on the perfect foresight portfolio and the accounting based hedge portfolio increases. This is because of the long/short set-up of these portfolios. Furthermore, this effect is not cancelled out by deflating the return on the accounting based hedge portfolio by the return on the perfect foresight hedge portfolio and it might lead to wrong inferences. This hypothesis is not investigated further in this research, but it has intuitive appeal given the financial turmoil and high volatility of returns in recent years.

9.7 Outliers

In chapter 7 the procedures for outlier removal were discussed. In this section a short look will be taken at the consequences of using an alternative procedure. In chapter seven, it was explained that the outliers with absolute studentised residual larger than four for a number of regressions are removed. This procedure is repeated until no such outliers exist. Gu (2007) notes that if this procedure is performed only once, outliers might still be present and inferences and conclusions might alter.

Table 32 shows the result of performing the outlier removal procedure for only one iteration. A comparison with table 26 shows that the results of the level model are mostly unaltered. This is also the case for the return model. The result of the price change model do differ significantly, i.e. for the coefficient of determination metric no increase in value relevance is found any more and for the residual dispersion metric most results become less significant. If this outlier removal procedure is not performed altogether, most results become insignificant (outcomes not tabulated). Hence it is concluded that the procedures for outlier removal might have a significant effect on value relevance and the conclusions that result from a research. This is in line with findings by Gu (2007). For this research, the 'converging' approach is preferred as outliers appear to be a problem.

Table 32

Summary of results based on one iteration of outlier removal

	Coefficient of determination metric				**Residual dispersion metric**				**Portfolio metric**	
Model	*(6.1)*	*(6.2)*	*(6.3)*	$\mathbb{E}[\dots]$	*(6.1)*	*(6.2)*	*(6.3)*	$\mathbb{E}[\dots]$	*(6.6)*	$\mathbb{E}[\dots]$
Hypothesis H1	+***	-	+***	+/-	-***	-*	-	+/-	+	+/-
Hypothesis H2		+		+		+		+	+	+
Hypothesis H3	-**	+***	+	-	+*	-	-	+	+[c]	-
Hypothesis H4	-***	+	+	-	+***	-	-	+	+	-
Hypothesis H5	+***	-	+***	+	-***	-**	-**	-	-	+
Hypothesis H6	-[c]	-[a]	-[b]	+	-	-	-	-	-[c]	+

Notes: * indicates significance at 10% level, ** at a 5% level and *** at a 1% level. Where applicable results are based on one-sided tests.
[a] This finding is significant at a 1% level, but is contrary to expectations.
[b] This finding is significant at a 5% level, but is contrary to expectations.
[c] This finding is significant at a 10% level, but is contrary to expectations.

9.8 Summary

This chapter provides robustness tests and further analysis for the results of chapter eight. An additional control variable, in the form of the percentage of firms reporting losses was introduced. It was graphically shown that the percentage of firms reporting losses increases over time. If this finding is combined with higher value relevance for negative earnings, as reported by Collins et al. (1997), this could have driven the increasing value relevance over time found in chapter eight. The analysis presented in section 9.3 indicates that this does not appear likely based on the residual dispersion metric. Furthermore, evidence is found that confirms the hypothesis that value relevance has increased after the mandatory introduction of IFRS.

It is also established that based on the residual dispersion metric there does appear to be a significant difference in value relevance between firms experiencing high or low amounts of business change. The latter type generally has higher value relevance. Finally, this finding appears robust to in the percentage of firms reporting losses, a conclusion similar to that of Lev and Zarowin (1999). The results do not indicate a substantial difference in value relevance between the countries classified as having a market-oriented financial system compared to those having a bank-oriented financial system. This result has also been established in some previous researches, such as Joos and Lang (1994), Joos (1997) and Arce and Mora (2002). As a possible explanation for the differing results of the coefficient of determination metric, scale effects were proposed, but a causal relationship has not been investigated.

This chapter also highlighted aspects in the calculation of the portfolio metric that differ from the methodology used by Francis and Schipper (1999). It is conjectured that these differences might drive the insignificant results for this metric. Finally, the impact of outliers was considered and it is shown that different procedures for outlier removal might lead to different inferences about value relevance.

10. Conclusions

10.1 Conclusions

In this study, temporal changes and differences in value relevance for a European sample were investigated. The main research question of this research is *"Have there been any changes in value relevance in France, Germany, the Netherlands and the United Kingdom over time and what are plausible explanations for these possible changes?"* This research is inspired by reported declines in value relevance over time for the U.S. over the period 1953-1996, see e.g. Brown et al. (1999), Chang (1999), Lev and Zarowin (1999) and Gu (2007). To the authors best knowledge no substantive research has performed on changes in value relevance over time in Europe. Also, a unique event has taken place in the sample countries of this research in 2005, namely the mandatory introduction of IFRS. Earlier research establishes an increase in value relevance after the introduction of IFRS, however these researches do not investigate changes in value relevance over time or use a different research methodology compared to this research, see e.g. Barth et al. (2008), Capkun et al. (2008) and Daske et al. (2008).

Value relevance in this research is defined under Interpretation 4, namely accounting information is deemed relevant if there exists a statistical association between share prices or returns and financial statement information. Under this interpretation, value relevance is measured by three different metrics; the coefficient of determination metric, the residual dispersion metric and the portfolio metric. The reason for using multiple metrics lies in the econometrical problems associated to the coefficient of determination metric and portfolio metric, such as scale effects. The residual dispersion metric for value relevance is a relatively advanced metric, proposed by Gu (2007) that can handle most of these econometrical issues.

The sample used in this research consists of all listed industrial firms in France, Germany, the Netherlands and the United Kingdom on which data was available over the period 1991-2011.

Some interesting conclusions are found in this research. First of all, it appears that there has been little change in value relevance in the sample countries of this research over the period 1991-2004. From 2005 and onwards, the introduction of IFRS appears to have significantly positively impacted value relevance. The sample years for which IFRS was the mandatory accounting standard in the sample countries coincides with some economically extremely bad years, i.e. of the global financial crisis. Since in previous researches it is found that negative earnings are more value relevant, this could also have driven the increase in value relevance. Robustness test incorporating the percentage of firms reporting losses indicates that this does not appear to be the case.

Similar to the results of Lev and Zarowin (1999), evidence is found that firms that experience large amounts of business change have on average lower value relevance. Lev and Zarowin (1999) also find an increase in the rate of business change over the years in their U.S. sample and relate this to a decline in average value relevance. In this research no increase in the rate of business change is found and this is in line with the finding of non-decreasing value relevance. It is established, although only weakly, that the rate of increase in value relevance is lower for firms experiencing a lot of business change.

Finally, because of the applied cross-country research design, it is possible to test for international differences in value relevance. Results indicate that there does not appear to be a significant difference in the combined value relevance of earnings and book values for countries with a market-oriented financial system compared to that of countries with a bank-oriented financial system. This finding is also established in previous researches, see e.g. Joos and Lang (1994), Joos (1997) and Arce and Mora (2002). Firms in these countries do appear to have a significant difference in average scale, with firms in bank-oriented countries having a higher average scale. This indicates that the coefficient of determination might not be appropriate to establish international differences in value relevance among these countries.

In addition to the results established in this research, there are two other important conclusions that can be drawn. First of all, as discussed in chapter nine, the procedures used to eliminate outliers can have a significant effect on the outcomes of value relevance research. Secondly, the sometimes contrary results established across methodologies, both in previous researches as well as in this research,, indicate that researchers should be wary about inferences based on only one methodology. Although some methodologies can be preferred over others, such as the residual dispersion metric over the coefficient of determination metric, it is advised to test for robustness based on different specifications. At the least, this research confirms the necessity of solid econometrics in accounting research.

10.2 Limitations and recommendations for further research

Despite the fact that a holistic approach to measure value relevance is taken in this research, it is bound to have some limitations. A first limitation is the variable definitions; some variables such as business change are defined in a single way. Only the book value of equity is considered to measure the mean absolute rank change. It might be interesting to test if the results are robust for other variable definitions, such as using the market capitalisation of firms to measure business change. This comment can also be considered more broadly to the other variable definitions.

A great deal of attention in this research was given to the different interpretations, models and metrics to measure value relevance. However, some metrics, such as the coefficient metric were not tested or only limited, such as the portfolio metric. Future researchers might investigate temporal changes and differences in value relevance measured by these metrics or measured under different interpretations of value relevance, such as Interpretation 2 applied by Kim and Kross (2005).

This research considers only the combined value relevance of earnings and book values. Previous research indicates that temporal changes in the value relevance of earnings or book values on a stand-alone basis are not similar (Collins et al., 1997). The value relevance of these items individually might also differ across the sample countries (Arce and Mora, 2002). A further investigation can be performed on these individual items.

Another limitation of this research is the limited number of sample countries used. Since only four countries are considered, this does not really constitute Europe as a whole. It might be interesting to see if the results of this research are valid for other European countries. Also, for most inferences in this study the countries were grouped. As was shown in chapter nine and discussed in chapter five, differences in accounting systems exist across countries. It therefore might not be appropriate to group the countries, but instead treat these countries on an individual basis. Based on the research methodology of this study no inferences can be made for e.g. changes in value relevance over time in the Netherlands; only for the four countries grouped. Using a country-by-country analysis might be useful to draw inferences on for example the influence of the introduction of IFRS on value relevance in France, as this does not need to be the same as in the United Kingdom.

Due to limited data availability, this research only considers changes in value relevance over the period 1991-2011. This period does not coincide with the researches applying U.S. samples, which are mostly for the period 1953-1996. It is interesting to see if different patterns in value relevance are present in Europe for this earlier sample period.

In closing, the discussion above and the research described throughout this thesis highlights a significant gap in the literature as to value relevance in Europe. This in turn presents many exciting opportunities for further study and analysis.

References

Alford, A., Jones, J., Leftwich, R. & Zmijewski, M. (1993). The relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research, 31 (supplement)*, 183-233.

Ali, A. & Hwang, L. (2000). Country specific factors related to financial reporting and the value relevance of accounting data. *Journal of Accounting Research, 38 (1)*, 1-21.

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

Arce, M. & Mora, A. (2002). Empirical evidence of the effect of European accounting differences on the stock market valuation of earnings and book value. *The European Accounting Review, 11 (3),* 573-599.

Ball, R., Kothari, S.P. & Robin, A. (2000). The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics, 29,* 1-51.

Ball, R., Robin, A. & Wu, J.S. (2003). Incentive versus standards: properties of accounting income in four East Asian countries. *Journal of Accounting and Economics, 36,* 235-270

Barth, M.E. & Kallapur, S. (1996). The effects of cross-sectional scale differences on regression results in empirical accounting research. *Contemporary Accounting Research, 13 (2),* 527-567.

Barth, M.E., Landsman, W.R. & Lang, M.H. (2008). International accounting standards and accounting quality. *Journal of Accounting Research, 46 (3),* 467-498.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics, 24,* 3-37.

Beaver, W. (1968). The information content of annual earnings announcements. *Journal of Accounting Research, 6 (supplement),* 67-92.

Brown, S., Lo, K. & Lys, T. (1999). Use of R^2 in accounting research: measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics, 28,* 83-115.

________ (2002). Erratum to "Use of R^2 in accounting research: measuring changes in value relevance over the last four decades." *Journal of Accounting and Economics, 33,* 141.

Campbell, J.Y., Lo, A.W. & MacKinlay, C. (1997). *The econometrics of financial markets* (2nd edition). Princeton, N.J.: Princeton University Press.

Capkun, V., Casavan-Jeny, A., Jeanjean, T. & Weiss, L.A. (2008). Earnings management and value relevance during the mandatory transition from local GAAPs to IFRS in Europe. *Working paper.*

Chang, J. (1999). The decline in value relevance of earnings and book values. *Working paper,* The Wharton School, University of Pennsylvania.

Choi, F.D.S. & Mueller, G.G. Mueller (1992). *International accounting.* Englewood Cliffs, N.J.: Prentice-Hall.

Christie, A.A. (1987). On cross-sectional analysis in accounting research. *Journal of Accounting and Economics, 9,* 231-258.

Clarkson, P., Hanna, J.D., Richardson, J.D. & Thompson, R. (2011). The impact of IFRS adoption on the value relevance of book value and earnings. *Journal of Contemporary Accounting and Economics, Forthcoming.*

Collins, D.W., Maydew, E.L. & Weiss, I.S. (1997). Changes in value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics, 24 (1),* 39-67.

Collins, D.W., Zhen Li, O. & Xie, H. (2009). What drives the increased informativeness of earnings announcements over time? *Review of Accounting Studies, 14 (1),* 1-30.

Daske, H., Hail, L., Leuz, C. & Verdi, R. (2008). Mandatory IFRS reporting around the world: early evidence on the economic consequences. *Journal of Accounting Research, 46 (5),* 1085-1142.

Dechow, P.M., Kothari, S.P. & Watts, R.L. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics, 25 (2),* 133-168.

Dontoh, A., Radhakrishnan, S. & Ronen, J. (2007). Is stock price a good measure for assessing value-relevance of earnings? An empirical test. *Review of Managerial Science, 1 (1),* 3-45.

Elliot, J. & Hanna, J. (1996). Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research, 34 (supplement),* 135-155.

Ely, K. & Waymire, G. (1999). Accounting standard-setting organizations and earnings relevance: longitudinal evidence from NYSE common stocks, 1927-93. *Journal of Accounting Research, 37 (2),* 293-317.

Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance, 25 (2),* 383-417.

Feltham, G.A. & Ohlson, J.A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research, 11 (2),* 689-731.

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). Expanded disclosure and the increased usefulness of earnings announcements. *The Accounting Review, 77 (3),* 515-546.

Givoly, D. & Hayn, C. (2000). The changing time-series properties of earnings, cash flows and accruals: has financial reporting become more conservative? *Journal of Accounting and Economics, 29 (3),* 287-320.

Greene, W.H. (2008). *Econometric analysis (6th edition).* Englewood Cliffs, N.J.: Prentice-Hall.

Gu, Z, (2005). Scale factor, R^2 and choice of levels versus returns models. *Journal of Accounting, Auditing and Finance, 20 (1),* 71-91.

________ (2007). Across-sample incomparability of R^2s and additional evidence on value relevance changes over time. *Journal of Business Finance & Accounting, 34 (7) & (8),* 1073-1098.

Hayn, C. (1995). The information content of losses. *Journal of Accounting and Economics, 20,* 125-153.

Hope, O.K. (2003). Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: an international study. *Journal of Accounting Research, 41 (2),* 235-272.

Joos, P. (1997). The stock market valuation of earnings and book values across international accounting systems. *Working paper,* INSEAD.

Joos, P. and Lang, M. (1994). The effects of accounting diversity: evidence from the European Union. *Journal of Accounting Research, 32 (supplement),* 141-168.

Kim, M. & Kross, W. (2005). The ability of earnings to predict future cash flows has been increasing – not decreasing. *Journal of Accounting Research, 43 (5),* 753-780.

King, R.D. & Langli, J.C. (1998). Accounting diversity and firm valuation. *The International Journal of Accounting, 33 (5),* 529-567.

Koller, T., Goedhart, M. & Wessels, D. (2005). *Valuation, measuring and managing the value of companies (4th edition – University edition).* Hoboken, N.J.: John Wiley & Sons.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A. & Vishny, R.W. (1997). Legal determinants of external finance. *The Journal of Finance, 52 (3),* 1131-1150.

________ (1998). Law and Finance. *Journal of Political Economy, 106 (6),* 1113-1155.

Landsman, W.R. & Maydew, E.L. (2002). Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research, 40 (3),* 797-808.

Leuz, C., Nanda, D. & Wysocki, P.D. (2003). Earnings management and investor protection: an international comparison. *Journal of Financial Economics, 69,* 505-527.

Lev, B. (1989). On the usefulness of earnings and earnings research: lessons and directions from two decades of international research. *Journal of Accounting Research, 27,* 153-192.

Lev, B. & Thiagarajan, S.R. (1993). Fundamental information analysis. *Journal of Accounting Research, 31 (2),* 190-215.

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

Maydew, E. (1997). Tax-induced earnings management by firms with net operating losses. *Journal of Accounting Research, 35 (1),* 83-96.

Mueller, G.G., Gernon, H. & Meek, G. (1994). *Accounting an international perspective.* New York, N.Y.: Business One Irwin.

Nobes, C. (1998). Towards a general model of the reasons for international differences in financial reporting. *ABACUS, 34 (2),* 162-187.

Nobes, C. & Parker, R. (2006). *Comparative International Accounting.* Essex, England: Pearson Education.

Ohlson, J.A. (1995). Earnings, book values and dividends in equity valuation. *Contemporary Accounting Research, 11 (2),* 661-687.

Pope, P. & Walker, M. (1999). International differences in timeliness, conservatism, and classification of earnings. *Journal of Accounting Research, 37 (supplement),* 53-99.

Ruland, W., Shon, J. & Zhou, P. (2007). Effective controls for research in international accounting. *Journal of Accounting and Public Policy, 26,* 96-116.

Ryan, S.G. (1995). A model of accrual measurement with implications for the evolution of the book-to-market ratio. *Journal of Accounting Research, 33 (1),* 95-112.

Ryan, S.G. & Zarowin, P.A. (2003). Why has the contemporaneous linear returns-earnings relation declined? *The Accounting Review, 78 (2),* 523-553.

Scott, W.R. (2006). *Financial accounting theory (4th revised edition).* Toronto, Canada: Pearsons Education.

Theil, H. (1971). *Principles of econometrics.* New York, N.Y.: Wiley.

Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57 (2),* 307-333.

Watts, R. & Zimmerman, J. (1978). Towards a positive theory of the determinants of accounting standards. *The Accounting Review*, *53 (1),* 112-134.

________ (1986). *Positive Accounting Theory.* Edgewood Cliffs, N.J.: Prentice-Hall.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48 (4),* 817-838.

Ye, J. (2005). Price models and the value relevance of accounting information. *Working paper,* Baruch College.