Bachelor of Science Thesis

Business and Economics
Department of Accounting and Financial Management
Stockholm School of Economics
15th May 2023

Predicting Corporate Bankruptcy With Personality Traits

A Quantitative Study of the Moderating Effects of the Big Five Personality Traits on Financial Variables Used for Corporate Bankruptcy Prediction

Marko Ivanovic **Tom Schreiber**

Abstract

This study aims to research whether incumbent corporate bankruptcy prediction models that are based on solely financial variables can be augmented with personality traits variables to increase their performance. For this purpose, a data set consisting of active and bankrupt small and medium enterprises in 13 European countries is studied. A quasi-Altman-Sabato (2007) model consisting of financial variables leverage, liquidity, profitability, coverage, and activity is developed and used as a control model. Then, Big Five personality traits are used to moderate the financial variables in the control model, by creating five new models, one for each personality trait. The findings show that all financial variables are statistically significantly moderated by conscientiousness and emotional stability. A majority of the financial variables are statistically significantly moderated by agreeableness, extraversion, and openness. Furthermore, moderating the financial variables with conscientiousness showed the best performance, thus decreasing information asymmetry. On the other hand, moderating the financial variables with extraversion reduced the model performance (sensitivity rate) vis-à-vis the control model, thus increasing information asymmetry.

Keywords

Corporate Bankruptcy Prediction, Big Five Personality Traits, Behavioural Finance

Acknowledgement

The authors wish to thank Henrik Andersson for the guidance throughout this study.

Table of Contents

1 Introduction

This introductory section is divided into four parts. First, the background of the study is presented, which is followed by the research question, delimitations and finally the structure of the thesis.

1.1 Background

Corporate bankruptcy prediction models have been an important topic in research over the past decades. Much of the pioneering work on bankruptcy prediction models was undertaken by Altman (1968). Prediction models are continuously updated to improve prediction accuracy, with models such as Ohlson (1980), Zmijewski (1984) and Shumway (2001) being commonly cited (Wu et al., 2010). Uncertainties regarding a company's financial soundness and asymmetric information impose a significant risk to many different types of stakeholders (Altman et al., 2017). Therefore, bankruptcy prediction models are a powerful tool to reduce risks and uncertainty, which adds value to decision makers (Altman et al., 2017). Bankruptcy prediction is important to both internal and external stakeholders of almost every company. It impacts CEO decision-making, investors and shareholders, bankers, financial analysts, auditors, and bankruptcy judges (Altman & Hotchkiss, 2010). The interest of using these established models or proprietary versions of these can be assumed to be of especially high importance today, with data showing that in the third quarter of 2022 bankruptcy filings increased by 19% percent in the EU-area (Eurostat, 2022).

Incumbent models, such as those proposed by Altman (1968), Ohlson (1980), Zmijewski (1984) and Shumway (2001) exclusively use financial variables. However, more recent studies have widened this perspective by including non-financial variables to improve the performance of corporate bankruptcy prediction models (Grunert et al., 2005; Lehmann, 2003). These additions include for example country-specific and industry-specific differences (Altman et al., 2017; Laitinen & Suvas, 2016). This trend of model extensions with non-financial data is important because studies have found that some incumbent country-specific corporate bankruptcy models are generally not usable across countries (Ooghe & Balcaen, 2007), but also because there may be performance improvements with adding non-financial data (Altman et al., 2017; Laitinen & Suvas, 2016). The reason for this can be due to large differences in accounting practices, macroeconomic situation, consumer behaviour, laws and judicial systems, risk-taking propensity, and culture (Altman et al., 2017). In practice, this trend can be considered as especially important for banks acting in an international context, who often use a single prediction model across countries due to regulatory requirements (Altman & Sabato, 2007). The currently applicable Basel III framework necessitates that banks must validate their corporate bankruptcy prediction models and document their performance (Altman & Sabato, 2007). Recently, a study followed this trend by studying the moderating effects of Hofstede's four cultural dimensions on different financial variables used in corporate bankruptcy prediction in an international context (Laitinen & Suvas, 2016). It found that Hofstede's four cultural dimensions have moderating effects on most of the financial-based variables used for corporate bankruptcy prediction (Laitinen & Suvas, 2016).

To further study the effects of corporate bankruptcy prediction on an international level, this thesis aims to assess the moderating effects of country-specific personality traits on corporate bankruptcy prediction and potential performance improvements that these variable additions provide to an incumbent corporate bankruptcy prediction model. There are two primary reasons for why this study is important and valuable.

First, this specific topic of studying the moderating effects of personality traits in a corporate bankruptcy prediction context is unexplored. As of writing this thesis, the authors are not aware of any other study that has researched these effects in a corporate bankruptcy prediction setting. Thus, this study makes a novel contribution to the corporate bankruptcy prediction research field and the non-financial data model augmentation trend.

Second, there is previous research that suggests that behavioural aspects have significant effect on corporate bankruptcy prediction, such as shown by Laitinen and Suvas (2016). Additionally, behavioural finance theories, such as present bias theory (Angeletos et al., 2001), mental accounting theory (Thaler, 1985) and prospect theory (Kahneman & Tversky, 1979) show that decisions making and risk-taking propensity is greatly shaped by individuals' behaviour and expectations, and thus not only a factor of complete rationality. Thus, this study contributes to the behavioural finance field as well by augmenting incumbent rational models with a behavioural perspective.

1.2 Research Question

The research question of this study is: how do country-specific personality traits moderate financial variables used in corporate bankruptcy prediction models and do their addition to a model improve the model's performance? This question will be answered by using personality traits categorized according to Big Five personality traits theory, which is widely considered as the most well-developed, influential, and validated theory on personality traits (Gurven et al., 2013). The corporate bankruptcy prediction model used as the control model is a quasi-Altman-Sabato (2007) model, which is a slightly altered Altman and Sabato (2007) model. The original Altman and Sabato (2007) model is altered by the authors of this thesis to better conform to the data set studied. The rationale for using this model as a control model is because it evaluated small to medium enterprises (SMEs) in Europe, which is also chosen as the focus geography of this study. Thus, the control model will be a quasi-Altman-Sabato (2007) model.

1.3 Delimitations

The study is delimited in terms of geography to Europe and corporate size to SMEs, in the time period 2002-2021. The rationale for this geography is that the corporate framework across Europe is relatively similar across countries. Furthermore, the choice of the time period 2002-2021 is due to that business environments change over time and keeping it relatively recent allows for the most relevant data. Additionally, studying this time period also allows for sufficient data for higher statistical significance. The choice of SMEs is again because of a greater data availability, as compared to limiting to just large companies. Most of the SMEs in this study’s data set are not publicly listed. As such, market variables such as share price and volatility measures are not used, given that the data is largely unavailable for these observations. Furthermore, macroeconomic variables are also not included, even though they might improve a corporate bankruptcy model’s performance. The rationale for this decision is that this study focuses solely on the moderating effects of personality traits, which puts macroeconomic variables outside the scope of this study.

1.4 Structure

This thesis has the following structure. First, an introduction in section 1 that describes the background and rationale of the study, as well as the research question and delimitations. Following that, section 2 continues with a literature review of key corporate bankruptcy prediction models and personality traits models. Section 3 establishes the chosen bankruptcy prediction model used as the control model and the personality traits model used, as well as corresponding hypotheses development. Section 4 describes the research design and methodology, including literature review, data collection, data pre-processing, logistic regression models, and performance and validity measurements used. After that, the empirical data, results, performance, and validity of the logistic regression models are presented in section 5. Section 6 presents the discussion, conclusions, and further research ideas. References and appendices can be found after section 6.

2 Literature Review

The literature review is divided into two sections. The first section describes different bankruptcy prediction models, which is then followed by theory on personality traits in the second section.

2.1 Bankruptcy Prediction Theory

The literature on different quantitative models for corporate bankruptcy prediction is not scarce. Altman’s (1968) Z-score was one of the earliest such bankruptcy models based on financial data. Since then, Ohlson (1980), Zmijewski (1984), Shumway (2001), Hillegeist et al. (2004) have developed the field using different statistical methods and altered accounting and financial market ratios. Additionally, some authors alter models to better predict certain industries or geographies, such as the J-UK model applied for the UK market (Almamy et al., 2016) or Wang’s (2004) model for internet firms.

Furthermore, incumbent models have been tested and adjusted by various researchers, such as by Chava and Jarrow (2004) who added industry effects and by Laitinen and Suvas (2016) who added cultural effects. Beyond these, there are models that apply other statistical methods, such as neural networks (Odom & Sharda, 1990; Wilson & Sharda, 1994), random forest (Barboza et al., 2017) and k-nearest neighbour (Chen et al., 2011). Below follows a more in-depth review of some of the key incumbent corporate bankruptcy prediction models.

2.1.1 Altman’s (1968) Multi-Discriminant Model

Alman introduced the Z-score model in 1968 to predict the probability of corporate bankruptcy (Altman, 1968). The Z-score model is a multi-discriminant model used to classify an observation into one or another pre-established group based upon observations of individual characteristics (Altman, 1968). Altman (1968) used a sample of 66 listed manufacturing firms that are divided into 33 corporate observations classified as bankrupt and 33 observations classified as non-bankrupt.

The model includes five financial variables based on balance sheet and income statement data, where an increase in all the five financial variables lowers the probability of company bankruptcy (Altman, 1968). The variables in Altman’s (1968) model are:

1. $\frac{Working\ Capital}{Total\ Assets}$
2. $\frac{Retained\ Earnings}{Total\ Assets}$

3. $\frac{EBIT}{Total\ Assets}$
4. $\frac{Market\ Value\ of\ Equity}{Book\ Value\ of\ Total\ Liabilities}$
5. $\frac{Sales}{Total\ Assets}$

The outcome is known as the Z-score, which is defined as: Z-score $= 0.012 * \frac{Working\ Capital}{Total\ Assets} + 0.014 * \frac{Retained\ Earnings}{Total\ Assets} + 0.033 * \frac{EBIT}{Total\ Assets} + 0.006 * \frac{Market\ Value\ of\ Equity}{Total\ Liabilities} + 0.999 * \frac{Sales}{Total\ Assets}$, where a Z-score above 2.99 indicates a non-bankrupt firm and a Z-score below 2.99 indicates a bankrupt firm (Altman, 1968). Altman (1968) concludes that the model is accurate in a time frame of two years prior to bankruptcy, with the accuracy of the model declining as the time prior to bankruptcy is extended.

2.1.2 Ohlson's (1980) Logit Model

Ohlson (1980) is one of the first and main critics of Altman's (1968) multi-discriminant model for predicting corporate bankruptcy. Ohlson (1980) developed a conditional logit model that is based on fewer assumptions and therefore was considered superior by Ohlson. For example, the logit model does not require a normal distribution of the sample and the logistic regression model's dependent variable is either 1 or 0 making the result more indicative and intuitive for corporate bankruptcy prediction (Ohlson, 1980).

The model was developed with a sample of 105 bankrupt and 2,058 non-bankrupt US industrial firms (Ohlson, 1980). Ohlson (1980) developed three models predicting corporate bankruptcy based on whether they were one, two, or three years prior to bankruptcy. Instead of five variables as in Altman's (1968) model, the Ohlson (1980) model includes seven ratios and two dummy variables. The variables in Ohlson's (1980) model are:

1. $\log \frac{Total\ Assets}{GNP}$
2. $\frac{Total\ Liabilities}{Total\ Assets}$
3. $\frac{Working\ Capital}{Total\ Assets}$
4. $\frac{Current\ Liabilities}{Current\ Assets}$
5. $\frac{Funds\ from\ Operations}{Total\ Liabilities}$
6. $\frac{Net\ Income}{Total\ Assets}$

7. $\frac{Net\ Income_t - Net\ Income_{t-1}}{Net\ Income_t + Net\ Income_{t-1}}$
8. $\begin{cases} Y = 1\ if\ Net\ Income_t < 0 \cap Net\ Income_{t-1} < 0 \\ Y = 0\ Otherwise \end{cases}$
9. $\begin{cases} X = 1\ if\ Total\ Assets < Total\ Liabilies \\ X = 0\ Otherwise \end{cases}$

The regression is: $T = -1.32 - 0.407 * \log\frac{Total\ Assets}{GNP} + 6.03 * \frac{Total\ Liabilities}{Total\ Assets} - 1.43 * \frac{Working\ Capital}{Total\ Assets} + 0.0757 * \frac{Current\ Liabilities}{Current\ Assets} - 1.72 * X - 2.37 * \frac{Net\ Income}{Total\ Assets} - 1.83 * \frac{Funds\ from\ Operations}{Total\ Liabilities} + 0.285 * Y - 0.521 * \frac{Net\ Income_t - Net\ Income_{t-1}}{Net\ Income_t + Net\ Income_{t-1}}$

2.1.3 Zmijewski's (1984) Probit Model

Zmijewski (1984) enhanced Ohlson's corporate bankruptcy model by introducing the probit model. Zmijewski's (1984) methods differ from the original logit model through the assumption that the distribution of error terms follows a normal distribution instead of a standard logistic distribution. The ratios included in Zmijewski's (1984) model are:

1. $\frac{Net\ Income}{Total\ Assets}$
2. $\frac{Total\ Liabilities}{Total\ Assets}$
3. $\frac{Current\ Assets}{Total\ Liabilities}$

The regression is: $Z_m = -4.336 - 4.513 * \frac{Net\ Income}{Total\ Assets} + 5.679 * \frac{Total\ Liabilities}{Total\ Assets} - 0.004 * \frac{Current\ Assets}{Total\ Liabilities}$

2.1.4 Altman and Sabato's (2007) Bankruptcy Prediction for SMEs Model

Altman and Sabato (2007) developed a logistic bankruptcy prediction model for SMEs. Previous Altman models, such as the original 1968 model, were only concerned with predicting bankruptcy of publicly traded companies, while the Altman and Sabato (2007) studied bankruptcy prediction specifically for SMEs. Using a logit regression technique on panel data consisting of over 2,000 US companies over the years 1994-2002, they developed a one-year corporate bankruptcy prediction model. The model consists of five variables, which are:

1. $Leverage;\ \frac{Short\ Term\ Debt}{Equity\ Book\ Value}$
2. $Liquidity;\ \frac{Cash}{Total\ Assets}$

3. $Profitability;\ \frac{EBITDA}{Total\ Assets}$
4. $Coverage;\ \frac{Retained\ Earnings}{Total\ Assets}$
5. $Activity;\frac{EBITDA}{Interest\ Expenses}$

The unlogged variable regression is: $4.28 * \log\left(\frac{PD}{1-PD}\right) + 0.18 * Profitability - 0.01 * Leverage + 0.08 * Coverage + 0.02 * Liquditiy + 0.19 * Activity$

Altman and Sabato (2007) found that the accuracy of the model significantly increased for the non-logged model at 75%, compared to the Altman Z-score model at 68%. Furthermore, the type 1 error was reduced from 21% to 12%. Altman and Sabato (2007) concluded that the prediction accuracy was improved when applying the logit model to the SME sector, contrary to applying the MDA technique, as was conducted in the Altman Z-score model (Altman, 1968).

2.2 Personality Traits Theory

There are different theories that model personality traits. The Big Five personality traits model is to be regarded as one of the most popular models (Feher & Vernon, 2021). Alternative models include the Supernumerary personality inventory model and the HEXACO model (Feher & Vernon, 2021). Below, an in-depth review of the models is presented.

2.2.1 Paunonen’s (2002) Supernumerary Personality Inventory Model

The Supernumerary personality inventory was developed to add to the Big Five personality traits model and debate on its comprehensiveness (Hong & Ong, 2017). The model measures individual differences in 10 traits: conventionality, seductiveness, manipulativeness, thriftiness, humorousness, integrity, femininity, religiosity, risk-taking, and egotism (Paunonen et al., 2003). Each trait is tested by 15 testing items amounting to a total of 150 items and test-takers respond to each question by answering to what extent they agree through the use of a 5-point Likert scale (Paunonen et al., 2003). The higher one scores on a trait, the stronger that trait resides in the individual. The model has been further tested on its validity and found to be a promising personality traits model (Paunonen et al., 2003).

2.2.2 Ashton and Lee’s (2007) HEXACO Model of Personality Structure

The HEXACO model of personality structure was developed by Ashton and Lee (2007) as an alternative to the Big Five personality traits model. The factors in the model are honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, as well as openness to experience (Ashton & Lee, 2008). The factors openness to experience, conscientiousness and extraversion are the same as those in the Big Five personality traits model (Ashton & Lee, 2008). Furthermore, Ashton and Lee (2008) write that the factor honesty-humility is especially interesting, because the Big Five personality traits model does not capture it. Moreover, Ashton and Lee (2008) write that a low level of honesty-humility is correlated with engagement in unethical business behaviour, materialistic tendencies, a strong need for power and dominance over others, a sense of entitlement and status-driven risk-taking (combining greed with a lack of fear), as well as with corruption.

2.2.3 McCrae and Costa’s (2008) Big Five Personality Traits Model

The initial version of the Big Five personality traits model was established by Tupes and Christal in 1961 (Tupes & Christal, 1992), but was throughout the following decades re-developed by Goldberg (1990), and McCrae and Costa (2008).

The Big Five personality traits model is the most widely accepted personality traits model by psychologists (Lim, 2023). The Big Five personality traits model has been extensively studied and validated across many different cultures and languages (McCrae et al., 1998). It has been shown to have high levels of test-retest reliability, meaning that people tend to score similarly on the different dimensions over time (Gosling et al., 2003). The model also has good predictive validity, meaning that it can predict a wide range of outcomes, such as job performance or academic achievement (Judge & Zapata, 2015).

The most recent version of the model characterizes five traits that can be used to describe an individual’s personality. These are agreeableness, conscientiousness, emotional stability, extraversion, and openness (McCrae & Costa, 2008). Following are descriptions of these traits from McCrae and Costa (2008). Agreeableness refers to high levels of trust, altruism, kindness, and affection. People with high agreeableness have particularly prosocial behaviour. Conscientiousness refers to high levels of thoughtfulness, good impulse control, and goal-directed behaviour. People with this characteristic are structured, organized, and detail-orientated. Emotional stability is characterized by happiness and non-moodiness. People that exhibit emotional stability have an even temper, low anxiety, and are not irritable. Extraversion is best described as being energized in the company of others. People with this trait are particularly talkative, assertive, and show high amounts of emotional expressiveness. Openness refers to imagination and insight, an eagerness to learn and to experience new things.

3 Model Selection and Hypotheses Development

This section is divided into two parts. First, two models are selected. One model for corporate bankruptcy prediction using financial variables to be used as the control model and one model for personality traits that will augment the control model. Thereafter, based on these chosen models, corresponding hypotheses are developed.

3.1 Model Selection

In this section, the two models chosen are described. The section starts with the corporate bankruptcy prediction model, which is followed by the personality traits model.

3.1.1 Quasi-Altman-Sabato (2007) Model

The model that closest resembles the chosen focus of this study is the Altman and Sabato (2007) model, since that model studied SMEs in Europe, which coincides with the focus of this study. Due to data unavailability in Orbis Europe (Bureau Van Dijk, 2023), the original Altman and Sabato (2007) model is slightly adjusted to conform to the data that is available. Table 1 details the changes made by the authors of this study to create the quasi-Altman-Sabato (2007) model.

Table 1
Comparison of the Original Altman and Sabato (2007) Model With the Quasi-Altman-Sabato (2007) Model

Variable Name	Original Altman and Sabato (2007) Variables	Quasi-Altman-Sabato (2007) Variables
Leverage	$\frac{Short\ Term\ Debt}{Equity\ Book\ Value}$	$\frac{Current\ Liabilities}{Shareholder\ Funds}$
Liquidity	$\frac{Cash}{Total\ Assets}$	$\frac{Cash\ and\ Cash\ Equivalents}{Total\ Assets}$
Profitability	$\frac{EBITDA}{Total\ Assets}$	$\frac{EBITDA}{Total\ Assets}$
Coverage	$\frac{Retained\ Earnings}{Total\ Assets}$	$\frac{Shareholder\ Funds}{Total\ Assets}$
Activity	$\frac{EBITDA}{Interest\ Expenses}$	$\frac{EBITDA}{Financial\ Expenses}$

Note: The (hypothesized) effects of an increase in the financial variables on the bankruptcy outcome remains the same, i.e., leverage increases bankruptcy outcome, while liquidity, profitability, coverage, and activity decreases bankruptcy outcome.

3.1.2 Big Five Personality Traits Model

The Big Five personality traits model is chosen for several reasons. First, it is one of the pioneering models in the personality traits field, as it was initially developed in 1961 and has been validated in many studies (Tupes & Christal, 1992). Additionally, the model is chosen primarily because there is data available at a country-level for the Big Five personality traits for different countries in Europe (Bartram, 2013), which allows for comparison and use in the regression models. Therefore, the Big Five personality traits model has both a theoretical and practical value for this study.

3.2 Hypotheses Development

In this section, hypotheses are developed. Due to easier readability and interpretation, hypotheses 2-6 are only concerned with if there are moderating effects (and not the direction of these effects). Additionally, several hypothesized effects are combined into singular hypotheses to increase clarity and readability.

3.2.1 Control Model

The control model variables in the quasi-Altman-Sabato (2007) model are assumed to have the same effects as described by Altman and Sabato’s (2007) model. Namely, that increased leverage increases bankruptcy outcome, while increased liquidity, profitability, coverage, and activity decrease bankruptcy outcome. Thus, the following hypothesis can be established:

H1: Increases in leverage increases bankruptcy outcome, while increases in liquidity, profitability, coverage, and activity decreases bankruptcy outcome.

3.2.2 Agreeableness

Research shows that more agreeable individuals are associated with more risk-aversive behaviour and hold less leverage (Borghans et al., 2009; Gow et al., 2016; Joseph & Zhang, 2021; Nicholson et al., 2005). Risk-aversiveness is associated with lower leverage, and higher liquidity and coverage in a company (Diamond & Rajan, 2000; González et al., 2013; Lev, 1974; Lewellen, 2006). This implies that agreeableness is associated with decreased leverage, increased liquidity, and coverage in a company. Furthermore, Antoncic et al. (2018) found that managers of SMEs who are more agreeable associate with lower corporate profitability. This implies that agreeableness is associated with lower profitability in a company.

Moreover, Choi and Laschever (2017) found that more agreeable individuals are associated with inefficient co-holding of credit card debt and low yield liquid assets. Similarly, Matz and

Gladstone (2022) found that increased agreeableness is associated with financial hardship. These imply that agreeableness is associated with decreased activity, which measures ability to service debt.

The above descriptions imply that agreeableness moderates the relationship between the financial variables and bankruptcy outcome. Thus, the hypothesis becomes:

H2: Agreeableness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.

3.2.3 Conscientiousness

Tok (2011) found that a higher level of conscientiousness is associated with a decreased tendency to participate in risky sports, and thus more risk-aversive behaviour. Several other studies showed similar association of higher levels of conscientiousness associating with higher levels of risk-aversion (Benischke et al., 2019; Czerwonka, 2019; Goldberg, 1990; Joseph & Zhang, 2021; Nicholson et al., 2005). This association implies that increased conscientiousness is associated with decreased leverage, increased liquidity, and increased coverage (Diamond & Rajan, 2000; González et al., 2013; Lev, 1974; Lewellen, 2006).

Lauter et al. (2023) found that traders who are more conscientious outperform on a risk-adjusted basis. This association implies that increased conscientiousness is associated with increased profitability.

Donnelly et al. (2012) show that more conscientious individuals are better at managing money. Additionally, Choi and Laschever (2017) found that more conscientious individuals associate with lower levels of inefficient co-holding of credit card debt and low yield liquid assets. This association implies that increased conscientiousness is associated with increased activity.

The above descriptions imply that conscientiousness moderates the relationship between the financial variables and bankruptcy outcome. Thus, the hypothesis becomes:

H3: Conscientiousness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.

3.2.4 Emotional Stability

Several authors' research find that increased emotional stability is associated with increased risk-seeking (Joseph & Zhang, 2021; Kuhnen et al., 2013; Nicholson et al., 2005; Rustichini et al., 2016). This association implies that increased emotional stability is associated with increased leverage, decreased liquidity, and decreased coverage (Diamond & Rajan, 2000; González et al., 2013; Lev, 1974; Lewellen, 2006).

Antoncic et al. (2018) found that managers of SMEs that are more emotionally stable associate with lower profitability in a company. This association implies that increased emotional stability is associated with decreased profitability.

Ksendzova et al. (2017) found that increased emotional stability was associated with better money management. This association implies that increased emotional stability is associated with increased activity.

The above descriptions imply that emotional stability moderates the relationship between the financial variables and bankruptcy outcome. Thus, the hypothesis becomes:

H4: Emotional stability moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.

3.2.5 Extraversion

Tok (2011) found that a higher level of extraversion is associated with an increased tendency to participate in risky sports. Similarly, other studies found that a higher level of extraversion is associated with higher levels of risk-taking (Benischke et al., 2019; Czerwonka, 2019; Joseph & Zhang, 2021; Nicholson et al., 2005). This association implies that increased extraversion is associated with increased leverage, decreased liquidity, and decreased coverage (Diamond & Rajan, 2000; González et al., 2013; Lev, 1974; Lewellen, 2006).

Gow et al. (2016) found that more extraverted CEOs associate with lower return on assets and cash flow in a company. This association implies that increased extraversion is associated with decreased profitability.

Research show that individuals who exhibit extraversion as a personality trait are more likely to have financial management struggles (Hoffmann & Risse, 2020). Furthermore, Harrison and Chudry (2011) found that more extraverted students are more likely to have worse financial management in terms of debt. This association implies that increased extraversion is associated with decreased activity.

The above descriptions imply that extraversion moderates the relationship between the financial variables and bankruptcy outcome. Thus, the hypothesis becomes:

H5: Extraversion moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.

3.2.6 Openness

Tok (2011) found that a higher level of openness is associated a with an increased tendency to participate in risky sports. Furthermore, several studies found that a higher level of openness is associated with higher levels of risk-taking (Benischke et al., 2019; Joseph & Zhang, 2021; Nicholson et al., 2005). This association implies that increased openness is associated with increased leverage, decreased liquidity, and decreased coverage (Diamond & Rajan, 2000; González et al., 2013; Lev, 1974; Lewellen, 2006).

Gow et al. (2016) found that more open CEOs associate with a lower profitability, return on assets and cash flow in a company. This association implies that increased openness is associated with decreased profitability.

Troisi et al. (2006) found that increased openness is associated with worse money management. This association implies that increased openness is associated with decreased activity.

The above descriptions imply that openness moderates the relationship between the financial variables and bankruptcy outcome. Thus, the hypothesis becomes:

H6: Openness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.

3.2.7 Performance

Altman et al. (2017) found that including non-financial variables in a corporate bankruptcy prediction model increases the model's performance. Similarly, Laitinen and Suvas (2016) found that including Hofstede's cultural dimensions as moderating variables increases the bankruptcy model's performance. Furthermore, Bartram (2013) found that the Big Five personality traits are correlated with Hofstede's cultural dimensions. Additionally, Wang & Chen (2020) and Peterson et al. (2003) found that there is a connection between CEO personality and organizational performance.

The above description implies that the addition of non-financial variables, such as the Big Five personality traits, will improve the performance of a financial variables-exclusive corporate bankruptcy prediction model. Thus, the hypothesis becomes:

H7: The addition of the Big Five personality traits as moderating variables to the control model increases the performance vis-à-vis the control quasi-Altman-Sabato (2007) model.

3.2.8 Hypotheses Summary

Table 2 summarises the hypotheses.

Table 2
Summary of Hypotheses

Hypothesis	Intuition
H1	Increases in leverage increases bankruptcy outcome, while increases in liquidity, profitability, coverage, and activity decreases bankruptcy outcome.
H2	Agreeableness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.
H3	Conscientiousness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.
H4	Emotional stability moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.
H5	Extraversion moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.
H6	Openness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.
H7	The addition of the Big Five personality traits as moderating variables to the control model increases the performance vis-à-vis the control quasi-Altman-Sabato (2007) model.

4 Research Design and Methodology

This section describes the research design and methodology of this thesis. First, a thorough literature review is conducted, which is followed by data collection and statistical analysis.

4.1 Literature Review

Literature is collected primarily from peer reviewed articles from renowned journals. The search terms used in the literature collection of corporate bankruptcy prediction models are combinations of: “Bankruptcy Prediction”, “SMEs”, “Manufacturing”, “Industry”, “Cross-Country”, “Comparison”, “MDA Model”, “Logit Model”, “Probit Model”, “International” and “Non-Financial Data”. The search terms used for literature collection of personality traits models are combinations of: “Personality Traits”, “Big Five”, “Performance” and “Risk-Taking”. Different corporate bankruptcy prediction models are summarised to see which variables are most relevant for corporate bankruptcy prediction. Thereafter, previous research on the Big Five personality traits model is summarised to see which personality traits may have an effect on corporate bankruptcy, such as personality traits associated with risk-taking. Hypotheses are then established based on this literature review.

4.2 Data Collection

The following sections present how the Big Five personality traits data is collected, followed by a section on how data on active and bankrupt companies is collected.

4.2.1 Personality Traits

Data on personality traits by country is sourced from Bartram’s (2013) research on Big Five personality traits. Bartram’s (2013) study aggregates personality data for over one million people in 31 countries. Bartram’s (2013) study aimed to extend the research on whether differences in personality scale score averages across countries are meaningful or due to sources of systematic bias (Bartram, 2013). The study showed that the OPQ32i version is very robust in terms of construct equivalence across countries and found support for the scalar equivalence of the OPQ32 scores across countries (Bartram, 2013).

OPQ32 is a 32-scale personality inventory, while OPQ32i is an updated version widely used around the world developed by the UK consulting firm SHL (Bartram, 2013). It is constructed by four forced-choice item quads sets of four statements from which the test taker chooses one “most like me” and one “least like me” (SHL, 2009). The Big Five personality traits measures produced in Bartram’s (2013) paper are created through weighted aggregation using 25 of the

32 OPQ scales. The reliability of the OPQ32 measured was found to be strong (Bartram & Brown, 2005). Bartram's (2013) study recognized that countries are not homogeneous, which supports the findings of Rentfrow et al. (2008).

4.2.2 Company Data

Data on active and bankrupt companies is gathered through Orbis Europe (Bureau Van Dijk, 2023). The data is filtered in terms of the following criteria.

4.2.2.1 Industry

Companies are filtered to only include manufacturing firms, as this study is delimited to manufacturing and to exclude financial companies, since financial ratios can significantly differ between financial and non-financial industries (Altman et al., 2017; Laitinen & Suvas, 2016).

4.2.2.2 Ownership

The owners must have private limited liability, in order to exclude partnerships and sole proprietors from the data set (Altman et al., 2017).

4.2.2.3 Bankruptcy Status

Bankruptcy is defined as "when a company or entrepreneur gets into financial distress, or cannot pay its debts, specific proceedings are available in every country to address the situation inclusively, involving all the creditors" (European e-justice, 2022).

Orbis Europe (Bureau Van Dijk, 2023) classifies a company as either "Active" or "Inactive". The sample of active firms will be drawn from firms classified as "Active". Bankrupt companies will be selected from specific sub-categories in both the main categories, "Active" and "Inactive". Sub-categories from the "Inactive" category that are included as bankrupt companies in this study are: "Bankruptcy", "Dissolved (bankruptcy)" and "Inactive (liquidation)". Sub-categories from the "Active" category that are included as bankrupt companies in this study are: "Active (rescue plan)" and "Active (insolvency proceedings)". The rationale to include companies from these sub-categories in the "Active" category is that these firms often suffer serious financial distress and are thus close to bankruptcy. With this filtering of bankrupt companies, irrelevant data can be excluded, such as mergers, demergers, takeovers and administrability suspended companies. Similar procedures have already been adopted in similar studies (Altman et al., 2017; Laitinen & Suvas, 2016).

4.2.2.4 Company Size

In this study, the focus is on European SMEs. SMEs represent 99% of all businesses in the EU, employing around 100 million people and accounting for more than half of Europe's GDP (European Commission, n.d.a). Therefore, SMEs can be considered to "form the backbone of Europe's economy" (European Commission, n.d.a). In addition, for the purpose of data collection, the choice to study SMEs is beneficial since SMEs experience a higher bankruptcy risk vis-à-vis publicly traded companies for reasons such as increased difficulties in acquiring financing. Dietsch and Petey (2004) found that SMEs could be considered riskier than larger businesses. Furthermore, since a large majority of European companies are SMEs, this allows for a larger quantity of data to be collected on the topic of bankruptcy, than for example restricting the data collection to larger and publicly listed companies.

This study follows EU's definition of SMEs as of recommendation 2003/361 (European Commission, n.d.b), which defines SMEs as an enterprise that has a staff headcount of between 1-249 employees. In addition to the employment numbers, the company also must have less than EUR 43 million in assets or less than EUR 50 million in operating revenue, or both.

4.2.2.5 Time

The financial data is filtered to the time period of 2002-2021. This span of years is considered as compatible with Altman and Sabato's (2007) corporate bankruptcy prediction model. Additionally, this time period is used to gather an adequate number of observations of bankrupt companies.

4.2.2.6 Geography

The data set includes 16 European countries; Belgium, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, and United Kingdom. The choice of these countries is based on data availability and compatibility between Bartram's (2013) Big Five personality traits by country data and financial data available at Orbis Europe (Bureau Van Dijk, 2023). The countries Greece, Norway and Russia are filtered out due to low sample size in terms of bankrupt companies, leaving 13 countries in the data set.

4.3 Statistical Analysis Methodology

This section is divided into data pre-processing, which is followed by methodology on logistic regressions.

4.3.1 Data Pre-processing

The pre-processing of data includes several steps, which are data filtering, data splitting, data balancing and outlier processing.

4.3.1.1 Data Filtering

The company data retrieved from Orbis Europe (Bureau Van Dijk, 2023) is first filtered through the steps described in table 3, as to remove non-valid observations, such as observations where key variables were not available, not in the years of the study's focus (2002-2021), companies not classified as SMEs and filtering out observations in countries with less than 60 bankrupt company observations (Greece, Norway, and Russia).

Table 3
Data Filtering Process

Filtering Step	Active		Bankrupt		Total	
	N Remaining	N Filtered Out	N Remaining	N Filtered Out	N Remaining	N Filtered Out
Retrieved observations from Orbis Europe (Bureau Van Dijk, 2023)	351,201		47,408		398,609	
Filtered out observations with missing values for key variables	171,773	179,428	22,414	24,994	194,187	204,422
Filtered out observations not belonging in the years 2002-2021	165,501	6,272	18,503	3,911	184,404	9,783
Filtered out observations not classified as SMEs	150,517	14,984	17,885	618	168,402	16,002
Filtered out observations in countries with less than 60 bankrupt companies (Greece, Norway, and Russia)	144,549	5,968	17,830	55	162,379	6,023

4.3.1.2 Data Splitting

After the filtering process, the data is randomly split into a training set consisting of 70% of the data and a testing set consisting of 30% of the data. The training set is the set that is pre-processed, on which the logistic regression models are created. The testing set is only filtered, but not pre-processed in terms of data balancing and outlier processing, and solely used for validating the performance of the different logistic regression models. The training set data is pre-processed with the purpose of creating a balanced data set in terms of bankrupt and active companies, reducing the impact of outliers and in turn improving model performance for corporate bankruptcy prediction.

4.3.1.3 Data Balancing

Given that the data retrieved from Orbis Europe (Bureau Van Dijk, 2023) is heavily imbalanced, with active companies representing 89% of the data and bankrupt companies representing 11% of the data, several options to balance the data are tested. These options are over-sampling, under-sampling, and synthetic minority over-sampling technique (SMOTE). Over-sampling involves an algorithm that randomly duplicates observations in the minority class (Chawla et al., 2002). Under-sampling involves an algorithm that randomly removes entries in the majority class (Chawla et al., 2002). SMOTE involves an algorithm that creates synthetic observations based on the characteristics of the observations in the minority class (Chawla et al., 2002).

Over-sampling is chosen as the algorithm for balancing after testing all three algorithms, because it shows the best and most consistent performance for the logistic regression models. The over-sampling algorithm is thus applied on the minority class (bankrupt companies) in the training data set to create an equal sized data set of 101,184 active and 101,184 bankrupt companies. The rationale for balancing the data is that both methods were tried (balancing versus not balancing the data) and balancing the data improved performance for predicting bankrupt companies (true positives), which is more aligned to the purpose of this study, as predicting companies that are active (true negatives) adds little scientific value.

4.3.1.4 Outlier Processing

There are several options to process outliers, including trimming and winsorization (Lusk et al., 2011). Trimming involves discarding outlier observations beyond a certain retained extreme point (on both sides) on the distribution, while winsorization involves replacing outlier observation with values of the extreme retained point (on both sides) on the distribution (Wilcox, 2005). In this study, the data is winsorized on a 90-percentile level while the data is grouped by bankruptcy status. The rationale for using winsorization is that it keeps the sample size intact.

4.3.2 Logistic Regression

Logistic regression as a methodology is used to create six statistical models. In the control model, only the effects of financial variables on bankruptcy outcome are evaluated. Thus, this control model has the form:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(Intercept+\beta_1*Leverage+\beta_2*Liquidity+\beta_3*Profitability+\beta_4*Coverage+\beta_5*Activity)}}$$

In the subsequent five models, in addition to the financial variables as independent variables (as in the control model), the financial variables are moderated with the Big Five personality traits, one at a time. Thus, models 2-6 will be of the form:

$$P(Y=1|X) = \frac{1}{1+e^{-(Intercept+\beta_i * FV_j + \beta_6 * PTV + \beta_k * FV_j * PTV)}}$$

Where β_i refers to the coefficient for the financial variables and loops over i such that $i \in [1,5]$.

Where FV_j refers to the financial variable and loops over j such that $j \in [Leverage, Liquidity, Profitiability, Coverage, Activity]$.

Where β_k refers to the coefficient of the moderating variables and loops over k such that $k \in [7, 11]$.

Where PTV refers to the personality trait of the respective model.

The choice of logistic regression over other models such as multi-discriminatory model is for the following reasons. First, the logistic regression model's outcome is more intuitive due to its binary nature (Ohlsson, 1980). Second, models such as the multi-discriminatory model have requirements such as that the variance-covariance matrices of the predictors should be the same for both groups (Ohlsson, 1980). Third, the matching procedure with the multi-discriminatory model is somewhat arbitrary (Ohlsson, 1980).

4.3.3 Assumptions and Validity

There are several assumptions that need to be tested for the logistic regression model to be valid. The first assumption is that the dependent variable is binary, i.e., that the logit has only two possible outcomes (Stoltzfus, 2011). The second assumption is that the observations are independent and not paired, which can be evaluated by looking at the data set (Stoltzfus, 2011). The third assumption is that there is no multicollinearity between the independent variables, which is tested through a correlation matrix and variance inflation factor (VIF) test (Stoltzfus, 2011). The fourth assumption is that any continuous independent variable is linearly correlated to the log odds of the dependent variable, which can be tested graphically by looking for a monotonic relationship in a scatter plot (Stoltzfus, 2011). The fifth assumption is that there are no influential values (i.e., extreme outliers), which can be removed through trimming or winsorization (Stoltzfus, 2011).

Furthermore, the different logistic regression models are also k-fold cross-validated, which is a method that selects random subsamples of the data based on a value k and runs the classification on each subsample and outputs average performance numbers, which can then be compared with the logistic regression model (Jung, 2018).

4.3.4 Performance

The performance of the logistic regression models is measured by various R^2 measurements, such as McFadden R^2 (McFadden, 1974), Cox and Snell R^2 (Cox, 2018) and Nagelkerke R^2 (Nagelkerke, 1991), which are used to estimate the goodness-of-fit of a model. Goodness-of-fit for logistic regression models is based on maximum likelihood estimates. The traditional R^2 cannot be used for non-linear regression models, which is why these pseudo-R^2 measurements are used to measure the explanatory power of the logistic regression model (Allison, 2013). However, the interpretation for these pseudo-R^2 measurements is not the same as for the traditional R^2. The rationale for reporting all these three pseudo-R^2 measurements is that there is not one measurement that is superior to the other measurements (Allison, 2013). As such, including all three measurements provides for a clearer interpretation of the models' performances.

Furthermore, measurements such as receiver operating characteristics (ROC), area under curve (AUC), Akaike information criterion (AIC) and the relative likelihood, accuracy %, sensitivity %, specificity %, prevalence % and F1-score % are used to measure the classification power of the different models. Refer to appendix A for definitions of the performance measurements.

5 Empirical Data, Results, Performance and Validity

This section describes the empirical data as well as the results of the statistical analyses. First, descriptive statistics of the data are presented, which is followed by the different logistic regression models.

5.1 Descriptive Statistics

First, descriptive statistics of empirical data gathered by Bartram (2013) is presented. Afterwards, descriptive statistics of the company data gathered from Orbis Europe (Bureau Van Dijk, 2023) is presented.

5.1.1 Personality Traits Data

Table 4 shows the Big Five personality traits grouped by country. The key take-aways from this table is that the number of observations differ greatly between countries, from 811 in Poland to 370,955 in the UK. Additionally, the gender distribution is fairly even in the observations for each country, with the exception of Germany (71% male), Poland (71% male) and Portugal (22% male). Finally, the standard deviation for each personality trait is also relatively equal around the value 2 between the different countries. This helps confirm the validity of the data.

Table 4
Descriptive Statistics of the Big Five Personality Traits by Country (Bartram, 2013)

			Agreeableness		Conscientiousness		Emotional Stability		Extraversion		Openness		
Country	N	Male	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean SD
Belgium	14,368	53%	5.98	1.93	5.33	2.09	5.19	2.12	5.72	2.05	5.53	1.99	2.04
Denmark	7,773	63%	5.76	1.86	5.51	2.06	6.38	1.82	6.37	1.98	5.60	2.01	1.95
Finland	11,076	56%	6.08	1.98	5.44	2.14	6.22	2.06	5.94	2.24	5.01	2.11	2.11
France	9,075	57%	5.62	1.76	5.55	2.00	4.85	2.00	5.47	1.84	5.67	1.92	1.90
Germany	8,075	71%	5.87	1.80	5.32	1.84	6.08	2.00	5.96	1.88	6.08	1.86	1.88
Hungary	1,685	53%	5.52	1.81	5.43	2.12	5.73	2.08	5.21	1.96	4.97	2.07	2.01
Italy	8,605	65%	5.00	1.82	5.02	1.89	5.09	2.00	5.70	1.85	5.86	1.93	1.90
Netherlands	50,104	58%	5.73	1.84	4.49	1.89	6.01	2.05	5.84	1.98	6.08	1.93	1.94
Poland	811	71%	4.75	1.69	5.35	2.07	4.99	2.01	4.64	2.03	5.84	1.81	1.92
Portugal	1,026	22%	5.81	1.68	5.12	1.96	4.77	1.76	5.88	1.79	5.79	2.01	1.84
Spain	4,770	60%	5.40	1.80	5.13	1.94	4.79	1.78	5.75	1.90	5.61	1.88	1.86
Sweden	30,863	51%	6.28	1.89	5.83	2.00	6.24	1.90	6.37	1.91	5.61	1.87	1.91
UK	370,955	60%	5.79	1.94	5.61	2.05	5.38	2.04	5.52	2.00	5.66	2.03	2.01

5.1.2 Company Data

Table 5 shows the number of active and bankrupt companies per country. There is a wide range of observations between the countries, from 215 companies for Netherlands, to 70,474 companies for Italy.

Table 5

Descriptive Statistics of Active and Bankrupt Companies by Country After Filtering but Before Data Splitting, Data Balancing and Outlier Processing

Country	Active N	Bankrupt N	Total	Share Bankrupt
Belgium	5,617	1,090	6,707	16.3%
Denmark	1,406	112	1,518	7.4%
Finland	2,046	61	2,107	2.9%
France	8,422	3,028	11,450	26.5%
Germany	7,942	818	8,760	9.3%
Hungary	5,478	132	5,610	2.4%
Italy	60,301	10,173	70,474	14.4%
Netherlands	131	84	215	39.1%
Poland	8,486	224	8,710	2.6%
Portugal	8,369	860	9,229	9.3%
Spain	27,376	721	28,097	2.6%
Sweden	4,024	160	4,184	3.8%
UK	4,951	367	5,318	6.9%
Total	**144,549**	**17,830**	**163,480**	**10.9%**

The descriptive statistics in table 6 are created on the training data set after it has been filtered, split, balanced, and processed for outliers. Panel A shows the aggregate descriptive statistics of both active and bankrupt companies. Looking at kurtosis and skewness in panel A shows that the values for all the variables are in the range of [-10,10] for kurtosis and [-3,3] for skewness, except for extraversion which is slightly outside on kurtosis (10.02). This indicates that the variables follow a normal distribution, which is in parts an effect of winsorization. Examining all three panels shows that there are no significant outliers.

For panels B and C, which show the descriptive statistics for active companies and bankrupt companies respectively, one can see that leverage held by active companies is on average 2.9x higher than bankrupt companies. Furthermore, the variable liquidity is on average 4.7x higher for active companies as compared to bankrupt companies. For the profitability, coverage, and activity variables, one can see that active companies on average have positive values, while bankrupt companies on average generally have negative values for these variables.

Table 6

Descriptive Statistics of Variables in the Training Data Set After Filtering, Data Splitting, Data Balancing and Outlier Processing

Panel A

Descriptive Statistics of Active and Bankrupt Companies

Statistic	N	Min	Q1	Q2	Q3	Max	Mean	SD	Kurtosis	Skewness
Bankrupt	202,368	0.00	0.00	0.50	1.00	1.00	0.50	0.5	1.00	0.00
Leverage	202,368	-5.90	0.17	1.05	4.00	28.25	3.26	7.48	7.60	2.17
Liquidity	202,368	0.00	0.01	0.04	0.13	0.38	0.09	0.11	4.20	1.48
Profitability	202,368	-0.50	-0.01	0.05	0.11	0.25	0.01	0.18	5.15	-1.56
Coverage	202,368	-0.83	0.03	0.18	0.38	0.75	0.16	0.38	3.89	-0.90
Activity	202,368	-27.14	-0.28	3.70	16.24	263.38	26.34	66.35	9.53	2.72
Agreeableness	202,368	4.75	5.00	5.00	5.62	6.28	5.32	0.39	2.01	0.56
Conscientiousness	202,368	4.49	5.02	5.12	5.33	5.83	5.19	0.22	3.04	0.95
Emotional stability	202,368	4.77	4.85	5.09	5.09	6.38	5.13	0.37	6.09	1.91
Extraversion	202,368	4.64	5.70	5.70	5.75	6.37	5.67	0.26	10.02	-1.70
Openness	202,368	4.97	5.67	5.86	5.86	6.08	5.76	0.19	9.21	-1.96

Panel B

Descriptive Statistics of Active Companies

Statistic	N	Min	Q1	Q2	Q3	Max	Mean	SD	Kurtosis	Skewness
Bankrupt	101,184	0.00	0.00	0.00	0.00	0.00	0.00	0.00	N/A	N/A
Leverage	101,184	0.16	0.40	0.94	2.27	5.61	1.67	1.73	3.35	1.28
Liquidity	101,184	0.01	0.03	0.10	0.23	0.38	0.14	0.13	2.15	0.70
Profitability	101,184	0.01	0.05	0.09	0.16	0.25	0.11	0.08	2.13	0.50
Coverage	101,184	0.08	0.20	0.38	0.58	0.75	0.39	0.22	1.76	0.14
Activity	101,184	-0.19	4.72	16.24	61.35	263.38	56.03	83.19	4.46	1.71
Agreeableness	101,184	4.75	5.00	5.40	5.62	6.28	5.34	0.40	2.14	0.50
Conscientiousness	101,184	4.49	5.02	5.13	5.33	5.83	5.19	0.21	3.57	1.14
Emotional stability	101,184	4.77	4.85	5.09	5.09	6.38	5.15	0.41	4.66	1.61
Extraversion	101,184	4.64	5.70	5.70	5.75	6.37	5.66	0.32	7.36	-1.58
Openness	101,184	4.97	5.61	5.84	5.86	6.08	5.73	0.22	7.35	-1.81

Panel C

Descriptive Statistics of Bankrupt Companies

Statistic	N	Min	Q1	Q2	Q3	Max	Mean	SD	Kurtosis	Skewness
Bankrupt	101,184	1.00	1.00	1.00	1.00	1.00	1.00	0.00	N/A	N/A
Leverage	101,184	-5.90	-1.99	1.51	7.91	28.25	4.84	10.19	3.48	1.25
Liquidity	101,184	0.00	0.00	0.01	0.05	0.13	0.03	0.04	3.35	1.31
Profitability	101,184	-0.50	-0.18	-0.01	0.06	0.12	-0.09	0.20	2.86	-1.06
Coverage	101,184	-0.83	-0.22	0.03	0.16	0.34	-0.07	0.35	2.92	-1.01
Activity	101,184	-27.14	-7.63	-0.28	2.95	11.62	-3.35	11.00	3.07	-0.91
Agreeableness	101,184	4.75	5.00	5.00	5.62	6.28	5.30	0.38	1.84	0.63
Conscientiousness	101,184	4.49	5.02	5.02	5.33	5.83	5.18	0.23	2.59	0.80
Emotional stability	101,184	4.77	4.85	5.09	5.09	6.38	5.11	0.31	8.47	2.31
Extraversion	101,184	4.64	5.70	5.70	5.70	6.37	5.68	0.19	14.00	-1.33
Openness	101,184	4.97	5.67	5.86	5.86	6.08	5.79	0.15	10.03	-1.66

5.2 Results of Logistic Regression Models

This section covers the results of the logistic regression models. Table 7 presents an overview of the results of the different models. Refer to appendix B for a larger version of table 7. Model 1 shows the effect of the variables in the quasi-Altman-Sabato (2007) model, which is used as the control model. Model 2 adds the personality trait agreeableness as an independent variable and moderating variables of each financial variable with agreeableness. Models 3, 4, 5 and 6 repeat the same procedure as that for model 2, but instead with the other respective Big Five personality traits. Thus, a model is created for each of the Big Five personality traits.

Table 7

Overview of the Six Regression Models (Refer to Appendix B for a Larger Version)

	Control Model (1)		Agreeableness Model (2)		Conscientiousness Model (3)		Emotional Stability Model (4)		Extraversion Model (5)		Openness Model (6)	
Independent Variable	**Estimate**	**Sd. err.**	**Estimate**	**Sd. err.**	**Estimate**	**Sd. err.**	**Estimate**	**Sd. err.**	**Estimate**	**Sd. err.**	**Estimate**	**Sd. err.**
Intercept	1.82***	0.01	-0.12	0.19	-1.01***	0.33	0.52**	0.22	-0.66**	0.33	-2.57***	0.44
Leverage	0.08***	0.00	0.38***	0.02	0.58***	0.04	0.27***	0.03	0.04	0.05	-0.29***	0.06
Liquidity	-8.73***	0.12	-37.49***	1.61	-74.89***	2.58	-0.75	1.72	13.08***	2.15	30.79***	3.08
Profitability	-3.81***	0.13	-3.41*	1.79	3.46	3.00	4.89***	1.82	-24.17***	2.58	-34.03***	3.72
Coverage	-4.50***	0.04	-7.49***	0.62	-15.41***	1.01	-2.27***	0.61	-1.58*	0.82	-6.01***	1.27
Activity	-0.09***	0.00	-0.21***	0.02	-0.32***	0.03	-0.22***	0.02	0.15***	0.02	0.31***	0.04
Personality trait			0.37***	0.04	0.56***	0.06	0.25***	0.04	0.21***	0.06	0.76***	0.08
Moderating Variables												
Leverage * Personality trait			-0.06***	0.00	-0.10***	0.01	-0.04***	0.01	0.01	0.01	0.06***	0.01
Liquidity * Personality trait			5.29***	0.29	12.54***	0.49	-1.56***	0.34	-3.86***	0.38	-6.86***	0.54
Profitability * Personality trait			-0.13	0.33	-1.43**	0.57	1.70***	0.35	3.59***	0.46	5.26***	0.65
Coverage * Personality trait			0.52***	0.12	2.02***	0.19	-0.44***	0.12	-0.52***	0.15	0.27	0.22
Activity * Personality trait			0.02***	0.00	0.04***	0.01	0.03***	0.00	-0.04***	0.00	-0.07***	0.01
Model Performance Measurements												
McFadden R^2	0.5834		0.5901		0.5946		0.5837		0.5842		0.5853	
Cox and Snell R^2	0.5546		0.5587		0.5614		0.5548		0.5551		0.5557	
Nagelkerke R^2	0.7394		0.7449		0.7486		0.7397		0.7401		0.7410	
AUC	0.858		0.861		0.862		0.858		0.858		0.859	
AIC	116,901		115,019		113,770		116,793		116,685		116,376	
Accuracy %	80.11		80.41		80.75		80.17		80.14		80.16	
Sensitivity %	80.98		81.04		81.15		81.06		80.96		81.10	
Specificity %	80.00		80.33		80.70		80.06		80.04		80.04	
Prevalence %	26.69		26.40		26.09		26.65		26.66		26.67	
F1-score %	47.20		47.60		48.07		47.30		47.24		47.30	

Note: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.10: *.

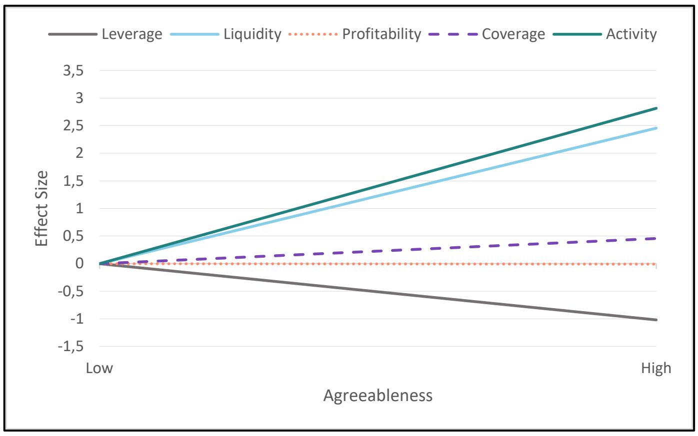
5.2.1 The Control Model (1)

The control model with only the financial variables shows statistically significant effects (at p-value < 0.01) for the intercept and all independent variables. They are also in the directions as expected by previous theory, namely that leverage increases bankruptcy outcome, while liquidity, profitability, coverage, and activity decreases bankruptcy outcome (Altman & Sabato, 2007). This shows strong empirical support for H1. The control model shows that the activity variable has the highest effect size on bankruptcy (when calculating the effect sizes at the variables' respective mean), followed by liquidity, coverage, leverage and lastly profitability.

5.2.2 The Agreeableness Model (2)

The agreeableness model shows that all the independent variables are statistically significant at a p-value < 0.01, except for profitability which is statistically significant at a p-value < 0.1. The financial variables are also all in the same direction as in the control model. Although the intercept of the agreeableness model is statistically insignificant. The moderating variables leverage, liquidity, coverage and activity are statistically significant at a p-value < 0.01, while profitability is statistically insignificant in this model. This shows empirical semi-support for H2 (not full support, due to profitability being statistically insignificant). The total effect sizes for the financial variables in the agreeableness model have changed compared to the control model. The leverage variable's total effect size is smaller in absolute terms, while the effect sizes for the liquidity, profitability, coverage, and activity variables are larger in absolute terms. Again, the activity variable has the largest effect size in the model, which is followed by liquidity, coverage, leverage, and profitability. Thus, this model shows that leverage is less important, while the other variables are more important, especially activity which has the largest gain in absolute terms. Figure 1 shows the total effect sizes for the agreeableness model calculated at the mean.

Figure 1
Total Effect Sizes for the Agreeableness Model



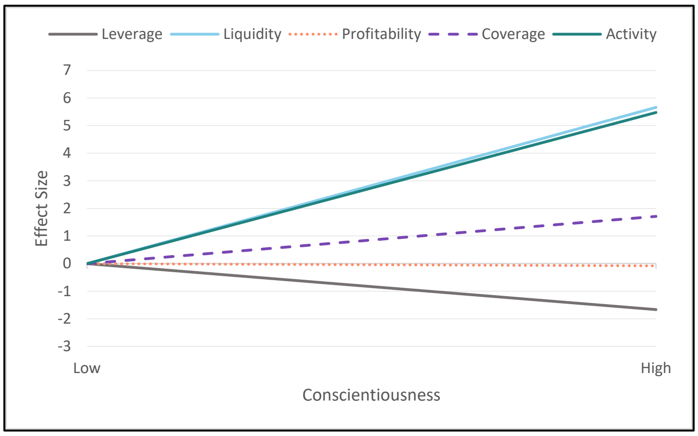
Note: these total effects are calculated at the mean.

5.2.3 The Conscientiousness Model (3)

The conscientiousness model shows that the intercept and all independent variables except for profitability are statistically significant at a p-value < 0.01. They are also all in the same direction as the control model (except for the profitability variable, but it is statistically insignificant). The moderating variables are all statistically significant at a p-value < 0.01, except for profitability which is statistically significant at a p-value < 0.05. This shows strong empirical support for H3. The total effect sizes for the financial variables in the conscientiousness model have changed compared to the control model. The leverage variable shows a smaller effect size in absolute terms, while the other variables show larger effect sizes in absolute terms. The directions of the total effects are the same as for the control model. Again, the activity variable has the largest effect size. This variable is followed by liquidity, coverage, leverage, and profitability. Figure 2 shows the total effect sizes for the conscientiousness model calculated at the mean. Given that the independent variable profitability loses its statistical significance in the conscientiousness model, while at the same time the profitability moderator variable is statistically significant in the model implies a complete moderation effect for the profitability variable.

Figure 2

Total Effect Sizes for the Conscientiousness Model



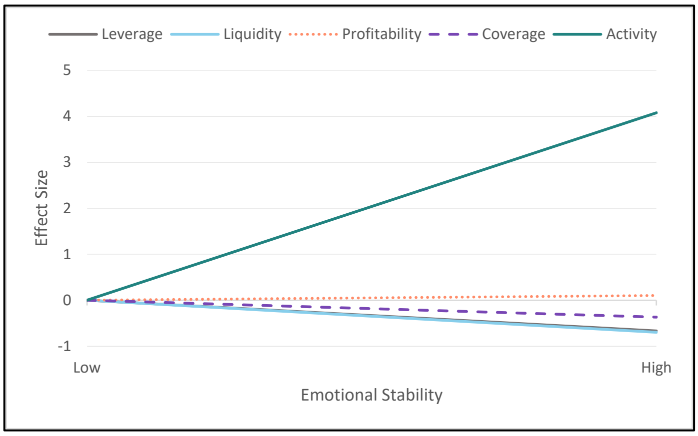
Note: these total effects are calculated at the mean.

5.2.4 The Emotional Stability Model (4)

The emotional stability model shows that the intercept is statistically significant at a p-value < 0.05. Furthermore, it shows that the all the independent variables except for liquidity are statistically significant at a p-value < 0.01. The independent variable profitability is in the opposite direction of the profitability variable in the control model. The moderating variables are all statistically significant at a p-value < 0.01. This shows strong empirical support for H4.

The total effect sizes for the financial variables in the emotional stability model have changed compared to the control model. The leverage and activity variables show smaller total effect sizes in absolute terms, while the other variables show larger total effect sizes in absolute terms. Additionally, the total effect size of the profitability variable is now directionally positive as opposed to directionally negative in the control model, meaning that increased profitability correlates with increased bankruptcy outcome in this model. Activity again has the largest effect size in absolute terms, followed by liquidity, coverage, leverage, and profitability. Figure 3 shows the total effect sizes for the emotional stability model calculated at the mean. Given that the independent variable liquidity loses its statistical significance in the emotional stability model, while at the same time the liquidity moderator variable is statistically significant in the model implies a complete moderation effect for the liquidity variable.

Figure 3
Total Effect Sizes for the Emotional Stability Model

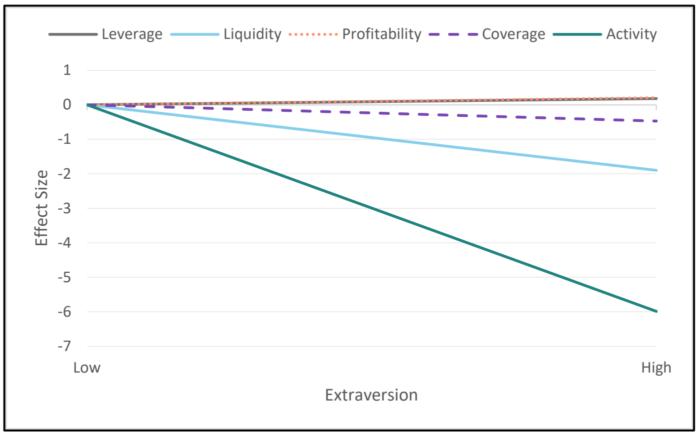


Note: these total effects are calculated at the mean.

5.2.5 The Extraversion Model (5)

The extraversion model shows that the intercept is statistically significant at a p-value < 0.05. The independent variables liquidity, profitability, activity, and personality trait are statistically significant a p-value < 0.01, while coverage is statistically significant at a p-value < 0.10. The independent variable leverage is statistically insignificant. The moderating variables are all statistically significant at a p-value < 0.01, except for leverage which is statistically insignificant. This shows empirical semi-support for H5 (not full support, due to leverage being statistically insignificant). The total effect sizes for the financial variables in the extraversion model have changed compared to the control model. The total effect sizes for leverage, liquidity, profitability, and coverage have increased in absolute terms (although the total effect for leverage is statistically insignificant). The total effect size for the activity variable has decreased. The directions of the total effects are the same as for the control model. Activity again has the largest effect size in absolute terms, followed by liquidity, coverage, leverage, and profitability. Figure 4 shows the total effect sizes for the extraversion model calculated at the mean.

Figure 4
Total Effect Sizes for the Extraversion Model

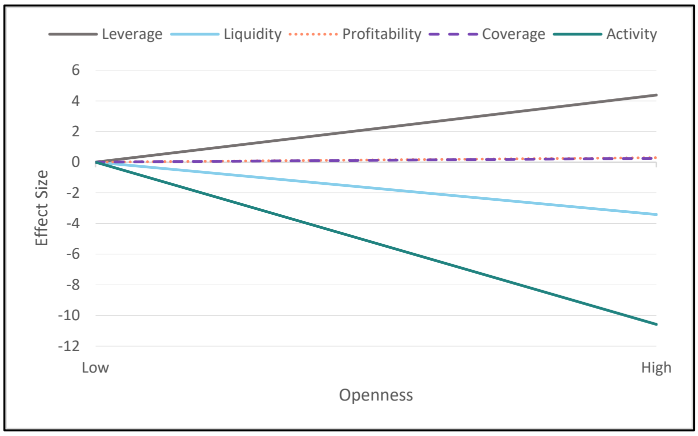


Note: these total effects are calculated at the mean.

5.2.6 The Openness Model (6)

The openness model shows that the intercept and all independent variables are statistically significant at a p-value < 0.01. The moderating variables are all statistically significant at a p-value < 0.01, except for coverage which is statistically insignificant. This shows empirical semi-support for H6 (not full support, due to coverage being statistically insignificant). The total effect sizes for the financial variables in the openness model have changed compared to the control model. The total effect sizes for leverage, liquidity and coverage have decreased in absolute terms, while the effect sizes for profitability and activity have increased in absolute terms. The directions of the total effect sizes are the same as in the control model. Activity again has the largest effect size in absolute terms, followed by liquidity, coverage, leverage, and profitability. Figure 5 shows the total effect sizes for the openness model calculated at the mean.

Figure 5
Total Effect Sizes for the Openness Model



Note: these total effects are calculated at the mean.

5.3 Performance

The Nagelkerke R^2 shows high values of more than 0.73 for all models. The McFadden R^2, and Cox and Snell R^2 show moderately high values of more than 0.58 for all models. These measurements imply that much of the variance is explained by all of the models.

Furthermore, the AIC is lower for each of models 2-6, compared to model 1, with the conscientiousness model having the lowest AIC of 113,770. The second-best model is the agreeableness model, with an AIC of 115,019. The rest of the models have a similar AIC in the 116,000 range. This implies that models 2-6 are better performing and that the conscientiousness is the best performing on this metric. Calculating the relative likelihood shows that models 1, 2, 4, 5 and 6 are ≈0 times as likely to minimize information loss as model 3 (conscientiousness model). Thus, one can conclude that the conscientiousness model is the best model based on the AIC measurement.

The AUC for models 2-6 is as good or better than model 1, which shows that they are better than the control model at classifying observations on the variables of sensitivity and specificity. All models have an AUC above 85%, which shows strong classification performance. Again, the conscientiousness model is the best performing out of the different models in terms of AUC. Refer to figure 7 for graphical illustrations of the AUC curves.

An analysis of the different classification performance measurements using a cut-off value of 0.5 shows that accuracy, sensitivity, specificity, prevalence, and F-1 score follow a similar trend, i.e., that models 2-6 outperform model 1 on all measures, with one exception of the extraversion model having a lower sensitivity (80.96%) than model 1 (80.98%). Sensitivity is an especially important performance measure since it shows the true positive rate. The higher the sensitivity measure, the less likely is it for the debt or equity provider to finance a company that is predicted to go bankrupt, which may result in large financial losses for the financing providers. All models are relatively well performing with accuracy measures above 80%. Again, the conscientiousness model is the best performing on these measures. These above-described findings show semi-support for H7 (not full support, due to lower sensitivity for the extraversion model compared to the control model and because the classification performance measurements are only slightly better). To conclude, there are performance improvements, although the performance improvements found are relatively small.

5.4 Assumptions and Validity

The first assumption of binarity of the logit outcome holds, because the dependent variable, bankrupt, can either be 1 (= bankrupt) or 0 (= not bankrupt). The second assumption of the observations being independent and not paired also holds, because the observations retrieved from Orbis Europe (Bureau Van Dijk, 2023) are independent of each other by design. The third assumption is that there is no multicollinearity. Table 8 shows that the VIF for all variables are below 5, which confirms that multicollinearity is not a problem (Gareth et al., 2013). This implies that the third assumption holds.

Table 8
Variance Inflation Factors for Independent Variables

Variable	VIF
Leverage	1.09
Liquidity	1.02
Profitability	1.72
Coverage	1.12
Activity	1.66
Agreeableness	4.03
Conscientiousness	3.70
Emotional stability	1.51
Extraversion	2.46
Openness	1.51

The correlation matrix in table 9 shows that all correlations are highly statistically significant at a p-value < 0.01, except for the correlation between coverage and leverage which is statistically significant at a p-value < 0.10. Most of the values show weak correlations, except for coverage and profitability at 0.73, and conscientiousness and agreeableness at 0.71. These values do not go below -0.8 or above 0.8, which are considered critical values (Laitinen &

Suvas, 2016). Thus, the correlation matrix shows that multicollinearity is not a problem in the data set.

Table 9
Pearson's Correlation Matrix

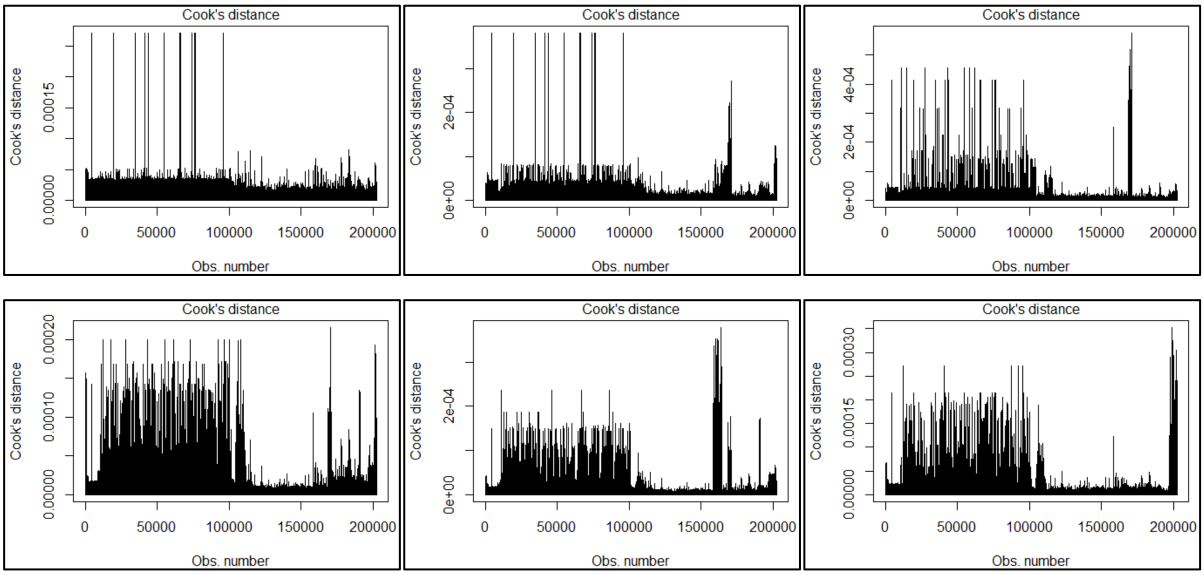
Bankrupt	1.00^{***}										
Leverage	0.21^{***}	1.00^{***}									
Liquidity	-0.49^{***}	-0.15^{***}	1.00^{***}								
Profitability	-0.54^{***}	-0.14^{***}	0.34^{***}	1.00^{***}							
Coverage	-0.62^{***}	0.00^{*}	0.42^{***}	0.73^{***}	1.00^{***}						
Activity	-0.45^{***}	-0.10^{***}	0.46^{***}	0.43^{***}	0.46^{***}	1.00^{***}					
Agreeableness	-0.04^{***}	-0.12^{***}	0.06^{***}	0.10^{***}	0.15^{***}	0.03^{***}	1.00^{***}				
Conscientiousness	-0.03^{***}	-0.11^{***}	0.06^{***}	0.08^{***}	0.15^{***}	0.02^{***}	0.71^{***}	1.00^{***}			
Emotional stability	-0.05^{***}	-0.02^{***}	0.01^{***}	0.09^{***}	0.07^{***}	0.04^{***}	0.37^{***}	0.30^{***}	1.00^{***}		
Extraversion	0.02^{***}	0.04^{***}	-0.02^{***}	-0.02^{***}	-0.04^{***}	0.02^{***}	0.37^{***}	-0.17^{***}	0.36^{***}	1.00^{***}	
Openness	0.14^{***}	0.09^{***}	-0.10^{***}	-0.10^{***}	-0.17^{***}	-0.03^{***}	-0.43^{***}	-0.47^{***}	-0.03^{***}	0.16^{***}	1.00^{***}
	Bankrupt	Leverage	Liquidity	Profitability	Coverage	Activity	Agreeableness	Conscientiousness	Emotional stability	Extraversion	Openness

Note: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.10: *.

The fourth assumption is that any continuous independent variable is linearly correlated to the log odds of the dependent variable. Scatter plots are created on the pre-processed data (i.e., after filtering, balancing and outlier processing of the training set data) and show all independent variables for respective model. Refer to appendix C for the scatter plots. Although the scatter plots for the different models look approximately the same, there are slight variations. Therefore, all scatter plots for all models are added for completeness. The scatter plots show that all the independent variables in all six logistic regression models are all fairly linearly correlated with the log odds of the dependent variable. This implies that the fourth assumption holds.

The fifth assumption is regarding influential values (extreme outliers). Although there are extreme outliers in the initial data set, these have been removed in the training data set through winsorization at a 90-percentile level. As such, influential values are not a problem in this study and the fifth assumption holds. Cook's distance plots are created on the winsorized data, which confirms that there are no influential values. One can see this because there are no values with a Cook's distance above $\frac{2}{\sqrt{Sample\ Size}}$ for any of the models, as suggested by Belsley et al. (1980).

Figure 6
Cook's Distance for the Six Logistic Regression Models



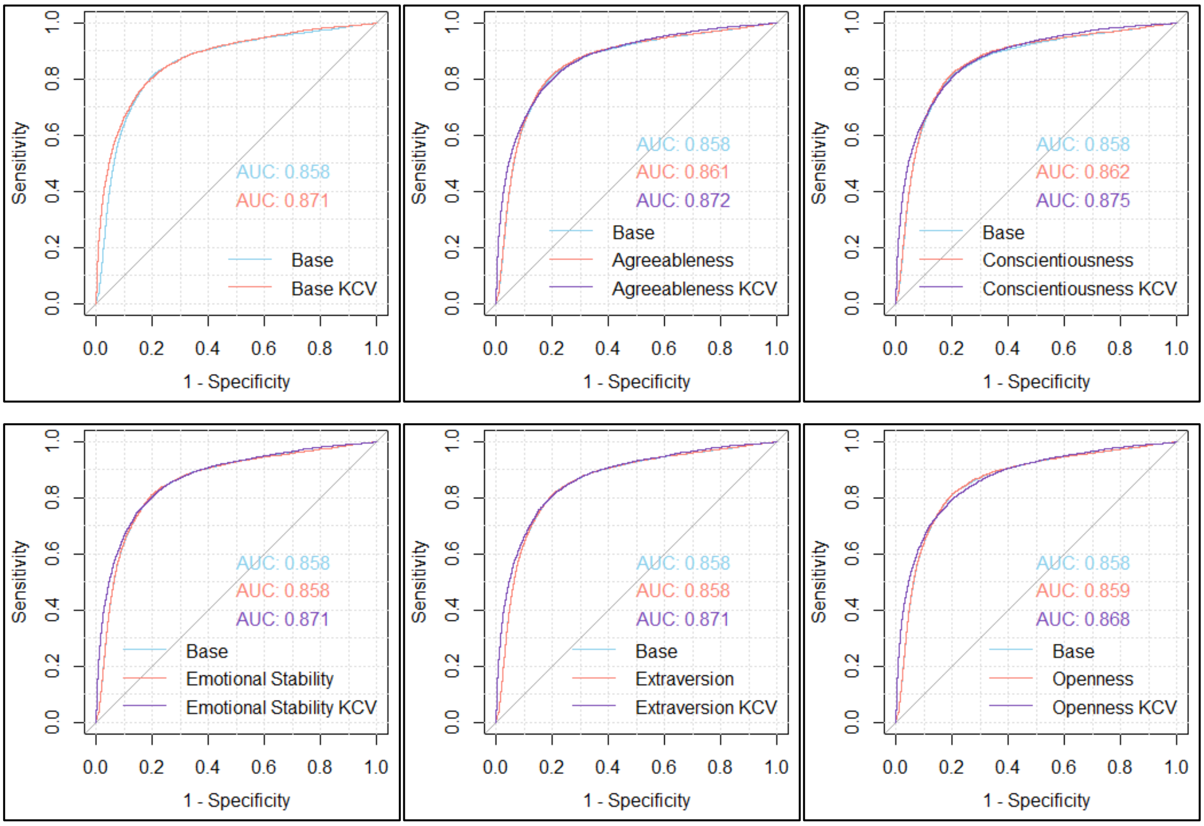
Note: In order from left to right: control model, agreeableness model, conscientiousness model, emotional stability model, extraversion model and openness model.

5.4.1 Receiver Operating Characteristics and K-fold Cross-Validation

Figure 7 shows the ROC-curves for all the different models. The AUC is larger for the agreeableness model, the conscientiousness model, and the openness model than the control model (the curves are slightly overlapping), with conscientiousness showing the strongest classification performance. The k-fold cross-validation is also the strongest for the conscientiousness model. Given that the k-fold cross-validation has a higher AUC for each model implies that the logistic regressions are validated also on smaller and random subsets of the data set.

Figure 7

ROC-Curves for the Six Logistic Regression Models and Respective K-fold Cross-Validation



6 Discussion and Conclusions

This section discusses the results and the implications they have on corporate bankruptcy prediction from an economic perspective. The section is divided into a results discussion, which is followed by limitations of the study, and ended with conclusions and further research topics.

6.1 Discussion

Several interesting discoveries are made with this study, with several interesting discussion topics. This study shows that adding personality traits as moderating variables have significant effects on almost all financial variables and increases the classification performance metrics for most personality trait models (i.e., not for all performance metrics in extraversion model), but the classification performance increase is relatively small vis-à-vis the control model). It is crucial for credit providers to correctly assess the future financial performance of potential debtors, thus adding personality traits provides additional information and serves to reduce the asymmetric information between the two parties. A creditor who wants to recoup their principal amount and interest may be more interested in the moderating effects on the activity ratio, while an equity provider such as a venture capital firm, who sees an increased valuation from increased profitability might be more interested in the moderating effects of the profitability variable. This cross-country study shows especially important results for credit and equity providers acting in an international context. With a more and more globalized world and international business practices, the importance of creating a valid bankruptcy prediction model for an international context becomes even more important.

The control model with only the financial variables shows significant effects for all variables and the variables' effects are in the direction as expected by previous theory. The control model shows that the activity variable has the highest effect size on bankruptcy (when calculating the effect sizes at the variables' respective means). This finding can be considered as reasonable, as not being able to service the company's debt is the definition of bankruptcy (European e-justice, 2022). Furthermore, profitability as a measurement has the lowest effect size, implying that profitability does not matter as much. Again, this can be seen as reasonable, since a company having low profitability does not necessarily mean that it has high leverage or high interest rates, which are the direct causes of bankruptcy.

The agreeableness model shows that the total effect sizes for the financial variables is approximately the same as in the control model, with the exception of the activity variable, which has increased considerably in effect size. One explanation for this can be that increased agreeableness means an increased need for rule conformity and meeting others' expectations, thus increasing the importance of being able to service debt (activity variable). The agreeableness model is one of the better performing models in this study.

The conscientiousness model shows larger effect sizes than the agreeableness model. Again, the activity variable has the largest effect size. Other strong effects can be seen for the liquidity and coverage variables, which show the strongest effects in this model vis-à-vis all the other models. Leverage as a variable has the weakest effect size. Conscientiousness as a model also completely moderated the profitability variable, confirming that conscientiousness is a predictor of profitability. An explanation for this could be that conscientiousness implies ambitiousness, thus taking on more leverage to amplify company performance and profitability, but due to high detail-orientation being able to manage this higher leverage. Additionally, conscientiousness as a model is the best performing model on all performance metrics. This result is in line with previous research (Duckworth et al., 2012; Ivcevic & Brackett, 2014; Kertechian, 2018; Ng & Feldman, 2010), that also found that conscientiousness as a personality trait is the most robust personality trait for predicting success. The reasons for this could be that conscientiousness has a more real impact on productivity than other personality traits, as has been shown in previous research (Barrick & Mount, 1991; Dollinger & Orf, 1991; Robertson et al., 2000). This finding implies that conscientiousness is the most valuable in decreasing information asymmetry.

The emotional stability model is one of the worst performing models in this study. One interesting thing about this model is that it predicts that increased profitability increases bankruptcy. Although this effect size in the model is small, it is statistically significant. One reason for this could potentially be that higher emotional stability is correlated with increased risk-seeking (Joseph & Zhang, 2021; Kuhnen et al., 2013; Nicholson et al., 2005; Rustichini et al., 2016) and thus increased leverage (Lev, 1974; Lewellen, 2006). Given that this study studies the years 2002-2021, which include several macroeconomic shocks (recession of 2008 and Covid-19), could imply that profitable, but highly leveraged companies, went bankrupt during these times. Other explanations could be additional factors such as size and company maturity. Another interesting finding is that the liquidity variable is completely moderated in this model, implying that emotional stability is a predictor of liquidity. Nevertheless, this result warrants further research.

The extraversion model is also one of the worst performing models. The result of this study shows that the extraversion model is the model that shows the worst sensitivity (i.e., the % correctly classified as true bankrupt companies) out of all models and was the only model that had a performance metric worse than the control model. This study shows that extraversion could potentially increase information asymmetry. This finding is especially important for the venture capital industry financing start-ups. Previous research show that entrepreneurs who are especially extravert are more likely to seek founding (Chapman & Hottenrott, 2023). Thus, using extraversion as an investment criterion may not provide additional information in terms of bankruptcy outcome.

The openness model is also one of the worst performing models. Openness is also a personality trait, similar to extraversion, that is found in start-up founders who are more inclined to seek funding (Chapman & Hottenrott, 2023). Although the openness model provides more information than the control model, this study does not find it as useful as a predictor compared to the stronger performing models such as conscientiousness and agreeableness.

6.2 Limitations

The main limitation of this study is that the Big Five personality traits were collected on an aggregated country level. This implies that these results are mainly correlations and not necessarily causal results. To improve results may be difficult and time consuming, as that would involve collecting personality traits of employees on a company level and industry level. This may not be possible with historical data, as bankrupt companies can no longer be surveyed to find out what the average personality traits of its employees are, as bankrupt companies by definition do not have any employees. Although this could be achieved by surveying employees in present non-bankrupt companies and then in the future evaluating which of these have filed for bankruptcy.

6.3 Conclusions and Further Research

The findings of this study are important, because they add a behavioural finance perspective beyond just financial numbers. The findings of this study are multifaceted and bring practical value. First, the variables leverage, liquidity, profitability, coverage, and activity in the developed quasi-Altman-Sabato (2007) are all statistically significant, showing that the model works well as a control model. Second, the Big Five personality traits have statistically significant moderating effects on almost all financial variables. More precisely, the conscientiousness and emotional stability models show statistical significance for all the moderating variables, while the agreeableness, extraversion and openness model show a majority of the variables as statistically significant. Third, the conscientiousness model decreased information asymmetry in the bankruptcy prediction model and is the best performing model, while the extraversion model increased information asymmetry and is one of the worst performing models, even performing worse than the control model in the sensitivity metric. Although it must be mentioned that the performance increases found in the personality traits models were relatively small for most classification performance measurements. Table 10 shows which hypotheses are supported.

Table 10
Summary of Hypotheses Results

Hypothesis	Intuition	Supported
H1	Increases in leverage increases bankruptcy outcome, while increases in liquidity, profitability, coverage, and activity decreases bankruptcy outcome.	Yes
H2	Agreeableness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.	Semi (not profitability)
H3	Conscientiousness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.	Yes
H4	Emotional stability moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.	Yes
H5	Extraversion moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.	Semi (not leverage)
H6	Openness moderates the effect that leverage, liquidity, profitability, coverage, and activity have on bankruptcy outcome.	Semi (not coverage)
H7	The addition of the Big Five personality traits as moderating variables to the control model increases the performance vis-à-vis the control quasi-Altman-Sabato (2007) model.	Semi (not extraversion)

Several interesting research ideas can be pursued. This study can be expanded by including more countries, industries, and time periods, and by clustering results by e.g., industry and size. Moreover, other incumbent corporate bankruptcy prediction models can be used as the control model, to validate the reproducibility and robustness of this study’s findings. Given that many incumbent bankruptcy prediction models are quite similar in terms of financial variables used, gives support that these moderating effects can be reproduced with other financial bankruptcy prediction models. Furthermore, other statistical models and machine learning algorithms can be applied to test if the results change. Moreover, the statistically significant effect found in the emotional stability model which showed that increased profitability increasing bankruptcy outcome warrants further research for why this is the case. Additionally, the completely moderated effects of profitability in the conscientiousness model and liquidity in the emotional stability are interesting findings that can be researched further.

7 References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723. https://doi.org/10.1109/TAC.1974.1100705

Allison, P. (2013). What's the best R-squared for logistic regression. *Statistical Horizons, 13*. Retrieved 07 May, 2023, from https://statisticalhorizons.com/r2logistic/

Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance, 36*, 278-285. https://doi.org/10.1016/j.jcorpfin.2015.12.009

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), 589-609. https://doi.org/10.2307/2978933

Altman, E. I., & Hotchkiss, E. (2010). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (Vol. 289). John Wiley & Sons.

Altman, E. I., & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the US market. *Abacus*, *43*(3), 332-357. https://doi.org/10.1111/j.1467-6281.2007.00234.x

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting*, *28*(2), 131-171. https://doi.org/10.1111/jifm.12053

Angeletos, G. M., Laibson, D., Repetto, A., Tobacman, J., & Weinberg, S. (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *Journal of Economic perspectives*, *15*(3), 47-68. https://doi.org/10.1257/jep.15.3.47

Antoncic, J. A., Antoncic, B., Grum, D. K., & Ruzzier, M. (2018). The big five personality of the SME manager and their company's performance. *Journal of Developmental Entrepreneurship*, *23*(04), 1850021. https://doi.org/10.1142/S1084946718500218

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and social psychology review*, *11*(2), 150-166. https://doi.org/10.1177/1088868306294907

Ashton, M. C., & Lee, K. (2008). The HEXACO model of personality structure and the importance of the H factor. *Social and Personality Psychology Compass*, *2*(5), 1952-1962. https://doi.org/10.1111/j.1751-9004.2008.00134.x

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405-417. https://doi.org/10.1016/j.eswa.2017.04.006

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*(1), 1-26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Bartram, D. (2013). Scalar equivalence of OPQ32: Big Five profiles of 31 countries. *Journal of Cross-Cultural Psychology*, *44*(1), 61-83. https://doi.org/10.1177/0022022111430258

Bartram, D., & Brown, A. (2005). Five factor model (Big Five) OPQ32 report: OPQ32 Technical manual supplement. *Thames Ditton, UK: SHL Group Ltd*.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons.

Benischke, M. H., Martin, G. P., & Glaser, L. (2019). CEO equity risk bearing and strategic risk taking: The moderating effect of CEO personality. *Strategic Management Journal*, *40*(1), 153-177. https://doi.org/10.1002/smj.2974

Borghans, L., Heckman, J. J., Golsteyn, B. H., & Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, *7*(2-3), 649-658. https://doi.org/10.1162/JEEA.2009.7.2-3.649

Bureau Van Dijk. (2023). Orbis Europe. Retrieved April 01, 2023, from https://www.bvdinfo.com

Chapman, G., & Hottenrott, H. (2023). Founder Personality and Start-up Subsidies. *ZEW-Centre for European Economic Research Discussion Paper*, (23-008). https://dx.doi.org/10.2139/ssrn.4404592

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of finance*, *8*(4), 537-569. https://doi.org/10.1093/rof/8.4.537

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357. https://doi.org/10.1613/jair.953

Chen, H. L., Liu, D. Y., Yang, B., Liu, J., Wang, G., & Wang, S. J. (2011). An adaptive fuzzy k-nearest neighbor method based on parallel particle swarm optimization for bankruptcy prediction. *Advances in Knowledge Discovery and Data Mining: 15th Pacific-Asia Conference, PAKDD 2011, Shenzhen, China, May 24-27, 2011, Proceedings, Part I 15* (pp. 249-264). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20841-6_21

Choi, H. S., & Laschever, R. A. (2018). The credit card debt puzzle and noncognitive ability. *Review of Finance*, *22*(6), 2109-2137. https://doi.org/10.1093/rof/rfx020

Cox, D. R. (2018). *Analysis of binary data.* Routledge. https://doi.org/10.1201/9781315137391

Czerwonka, M. (2019). Cultural, cognitive and personality traits in risk-taking behaviour: Evidence from Poland and the United States of America. *Economic research-Ekonomska istraživanja*, *32*(1), 894-908. https://doi.org/10.1080/1331677X.2019.1588766

Diamond, D. W., & Rajan, R. G. (2000). A theory of bank capital. *the Journal of Finance*, *55*(6), 2431-2465. https://doi.org/10.1111/0022-1082.00296

Dietsch, M., & Petey, J. (2004). Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs. *Journal of Banking & Finance*, *28*(4), 773-788. https://doi.org/10.1016/j.jbankfin.2003.10.006

Dollinger, S. J., & Orf, L. A. (1991). Personality and performance in "personality": Conscientiousness and openness. *Journal of Research in Personality*, *25*(3), 276-284. https://doi.org/10.1016/0092-6566(91)90020-Q

Donnelly, G., Iyer, R., & Howell, R. T. (2012). The Big Five personality traits, material values, and financial well-being of self-described money managers. *Journal of economic psychology*, *33*(6), 1129-1142. https://doi.org/10.1016/j.joep.2012.08.001

Duckworth, A. L., Weir, D., Tsukayama, E., & Kwok, D. (2012). Who does well in life?

Conscientious adults excel in both objective and subjective success. *Frontiers in psychology*, *3*, 356. https://doi.org/10.3389/fpsyg.2012.00356

European Commission. (n.d.a). *Entrepreneurship and small and medium-sized enterprises (SMEs)*. Retrieved 07 May, 2023, from https://single-market-economy.ec.europa.eu/smes_en

European Commission. (n.d.b). *SME Definition*. Retrieved 07 May, 2023, from https://single-market-economy.ec.europa.eu/smes/sme-definition_en

European e-justice. (2022). *Insolvency/bankruptcy*. Retrieved 07 May, 2023, from https://e-justice.europa.eu/447/EN/insolvencybankruptcy

Eurostat (2022). *Quarterly registrations of new businesses and declarations of bankruptcies – statistics*. Retrieved 07 May, 2023, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Quarterly_registrations_of_new_businesses_and_declarations_of_bankruptcies_-_statistics&oldid=504228

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, *27*(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

Feher, A., & Vernon, P. A. (2021). Looking beyond the Big Five: A selective review of alternatives to the Big Five model of personality. *Personality and Individual Differences*, *169*, 110002. https://doi.org/10.1016/j.paid.2020.110002

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Spinger.

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216. https://doi.org/10.1037/0022-3514.59.6.1216

González, M., Guzmán, A., Pombo, C., & Trujillo, M. A. (2013). Family firms and debt: Risk aversion versus risk of losing control. *Journal of Business Research*, *66*(11), 2308-2320. https://doi.org/10.1016/j.jbusres.2012.03.014

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, *37*(6), 504-528. https://doi.org/10.1016/S0092-6566(03)00046-1

Gow, I. D., Kaplan, S. N., Larcker, D. F., & Zakolyukina, A. A. (2016). CEO personality and firm policies (No. w22435). *National Bureau of Economic Research*. https://doi.org/10.3386/w22435

Grunert, J., Norden, L., & Weber, M. (2005). The role of non-financial factors in internal credit ratings. *Journal of Banking & Finance*, *29*(2), 509-531. https://doi.org/10.1016/j.jbankfin.2004.05.017

Gurven, M., Von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager–farmers in the Bolivian Amazon. *Journal of personality and social psychology*, *104*(2), 354. https://doi.org/10.1037/a0030841

Harrison, N., & Chudry, F. (2011). Overactive, overwrought or overdrawn? The role of personality in undergraduate financial knowledge, decision-making and debt. *Journal of Further and Higher Education*, *35*(2), 149-182. https://doi.org/10.1080/0309877X.2010.540080

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of accounting studies*, *9*, 5-34. https://doi.org/10.1023/B:RAST.0000013627.90884.b7

Hoffmann, A. O., & Risse, L. (2020). Do good things come in pairs? How personality traits help explain individuals' simultaneous pursuit of a healthy lifestyle and financially responsible behavior. *Journal of Consumer Affairs*, *54*(3), 1082-1120. https://doi.org/10.1111/joca.12317

Hong, R.Y., Ong, X.L. (2017). The Supernumerary Personality Inventory. *Zeigler-Hill, V., Shackelford, T. (eds) Encyclopedia of Personality and Individual Differences.* Springer, Cham. https://doi.org/10.1007/978-3-319-28099-8_89-1

Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of research in personality*, *52*, 29-36. https://doi.org/10.1016/j.jrp.2014.06.005

Joseph, E. D., & Zhang, D. C. (2021). Personality profile of risk-takers. *Journal of Individual Differences*. https://doi.org/10.1027/1614-0001/a000346

Judge, T. A., & Zapata, C. P. (2015). The person–situation debate revisited: Effect of situation strength and trait activation on the validity of the Big Five personality traits in predicting job performance. *Academy of Management Journal*, *58*(4), 1149-1179. https://doi.org/10.5465/amj.2010.0837

Jung, Y. (2018). Multiple predicting K-fold cross-validation for model selection. *Journal of Nonparametric Statistics*, *30*(1), 197-215. https://doi.org/10.1080/10485252.2017.1404598

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kertechian, S. K. (2018). Conscientiousness as a key to success for academic achievement among French university students enrolled in management studies. *The International Journal of Management Education*, *16*(2), 154-165. https://doi.org/10.1016/j.ijme.2018.02.003

Ksendzova, M., Donnelly, G. E., & Howell, R. T. (2017). A brief money management scale and its associations with personality, financial health, and hypothetical debt repayment. *Journal of Financial Counseling and Planning, 28*(1), 62-75. https://doi.org/10.1891/1052-3073.28.1.62

Kuhnen, C. M., Samanez-Larkin, G. R., & Knutson, B. (2013). Serotonergic genotypes, neuroticism, and financial choices. *PloS one*, *8*(1), e54632. https://doi.org/10.1371/journal.pone.0054632

Laitinen, E. K., & Suvas, A. (2016). Financial distress prediction in an international context: Moderating effects of Hofstede's original cultural dimensions. *Journal of Behavioral and Experimental Finance*, *9*, 98-118. https://doi.org/10.1016/j.jbef.2015.11.003

Lauter, T., Prokopczuk, M., & Trück, S. (2023). Conscientiousness and Performance: The Influence of Personality on Investment Decisions. *The Journal of Wealth Management*, *25*(4), 17-44. https://doi.org/10.3905/jwm.2023.1.199

Lehmann, B. (2003). Is it worth the while? The relevance of qualitative information in credit rating. *The Relevance of Qualitative Information in Credit Rating (April 17, 2003)*. https://dx.doi.org/10.2139/ssrn.410186

Lev, B. (1974). On the association between operating leverage and risk. *Journal of financial and quantitative analysis*, *9*(4), 627-641. https://doi.org/10.2307/2329764

Lewellen, K. (2006). Financing decisions when managers are risk averse. Journal of financial economics, 82(3), 551-589. https://doi.org/10.1016/j.jfineco.2005.06.009

Lim, A. G. (2023). Big Five Personality Traits: The 5-Factor Model of Personality. *Simply Psychology*. Retrieved 14 May, 2023, from https://www.simplypsychology.org/big-five-personality.html

Lusk, E. J., Halperin, M., & Heilig, F. (2011). A note on power differentials in data preparation between trimming and winsorizing. *Business Management Dynamics*, *1*(2), 23.

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, *5*(9), 1315-1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

Matz, S. C., & Gladstone, J. J. (2020). Nice guys finish last: When and why agreeableness is associated with economic hardship. *Journal of Personality and Social Psychology*, *118*(3), 545. https://doi.org/10.1037/pspp0000220

McCrae, R. R., & Costa Jr, P. T. (2008). *The five-factor theory of personality.* O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (pp. 159–181). The Guilford Press.

McCrae, R. R., Costa Jr, P. T., Del Pilar, G. H., Rolland, J. P., & Parker, W. D. (1998). Cross-cultural assessment of the five-factor model: The Revised NEO Personality Inventory. *Journal of cross-cultural psychology*, *29*(1), 171-188. https://doi.org/10.1177/0022022198291009

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic Press.

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691-692. https://doi.org/10.1093/biomet/78.3.691

Ng, T. W., & Feldman, D. C. (2010). Human capital and objective indicators of career success: The mediating effects of cognitive ability and conscientiousness. *Journal of Occupational and Organizational Psychology*, *83*(1), 207-235. https://doi.org/10.1348/096317909X414584

Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and domain-specific risk taking. *Journal of Risk Research*, *8*(2), 157-176.

https://doi.org/10.1080/1366987032000123856

Odom, M. D., & Sharda, R. (1990, June). A neural network model for bankruptcy prediction. In *1990 IJCNN International Joint Conference on neural networks* (pp. 163-168). IEEE. https://doi.org/10.1109/IJCNN.1990.137710

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131. https://doi.org/10.2307/2490395

Ooghe, H., & Balcaen, S. (2007). Are failure prediction models widely usable? An empirical study using a Belgian data set. *Multinational Finance Journal*, *11*(1/2), 33-76.

Paunonen, S. V. (2002). Design and construction of the Supernumerary Personality Inventory. *Research Bulletin*, *763*.

Paunonen, S. V., Haddock, G., Forsterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European journal of personality*, *17*(6), 413-433. https://doi.org/10.1002/per.496

Peterson, R. S., Smith, D. B., Martorana, P. V., & Owens, P. D. (2003). The impact of chief executive officer personality on top management team dynamics: one mechanism by which leadership affects organizational performance. *Journal of applied Psychology*, *88*(5), 795. https://doi.org/10.1037/0021-9010.88.5.795

Rentfrow, P. J., Gosling, S. D., & Potter, J. (2008). A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. *Perspectives on Psychological Science*, *3*(5), 339-369. https://doi.org/10.1111/j.1745-6924.2008.00084.x

Robertson, I. T., Baron, H., Gibbons, P., MacIver, R., & Nyfield, G. (2000). Conscientiousness and managerial performance. *Journal of Occupational and Organizational Psychology*, *73*(2), 171-180. https://doi.org/10.1348/096317900166967

Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, *64*, 122-137. https://doi.org/10.1016/j.socec.2016.04.019

SHL. (2009). *OPQ32 development and psychometric properties of OPQ32r (supplement to the OPQ32 technical manual)*. Thames Ditton, UK: SHL Group Ltd.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, *74*(1), 101-124. https://doi.org/10.1086/209665

Stoltzfus, J. C. (2011). Logistic regression: a brief primer. *Academic emergency medicine*, *18*(10), 1099-1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, *4*(3), 199-214. https://doi.org/10.1287/mksc.4.3.199

Tok, S. (2011). The big five personality traits and risky sport participation. *Social Behavior and Personality: an international journal*, *39*(8), 1105-1111. https://doi.org/10.2224/sbp.2011.39.8.1105

Troisi, J. D., Christopher, A. N., & Marek, P. (2006). Materialism and money spending disposition as predictors of economic and personality variables. *North American Journal of Psychology*, *8*(3).

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of personality*, *60*(2), 225-251. https://doi.org/10.1111/j.1467-6494.1992.tb00973.x

Wang, B. (2004). *Strategy changes and internet firm survival*. University of Minnesota.

Wang, S., & Chen, X. (2020). Recognizing CEO personality and its impact on business performance: Mining linguistic cues from social media. *Information & Management*, *57*(5), 103173. https://doi.org/10.1016/j.im.2019.103173

Wilcox, R. (2005). Trimming and winsorization. *Encyclopedia of biostatistics*, *8*. https://doi.org/10.1002/0470011815.b2a15165

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision support systems*, *11*(5), 545-557. https://doi.org/10.1016/0167-9236(94)90024-8

Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, *6*(1), 34-45. https://doi.org/10.1016/j.jcae.2010.04.002

Zajic, A. (2022). What Is Akaike Information Criterion (AIC)? *Built In*. Retrieved 14 May, 2023, from https://builtin.com/data-science/what-is-aic.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, 59-82. https://doi.org/10.2307/2490859

8 Appendix A – Performance Measurements Definitions

Appendix A describes the different performance measurements, starting with Pseudo R^2, followed by receiver operating characteristics and ending with Akaike information criterion.

8.1 Pseudo R^2

McFadden R^2 is defined as: $R^2_{McFadden} = 1 - \ln \frac{L_M}{\ln L_0}$, where L_0 is the value of the likelihood function for a model with no predictors (intercept only) and where L_M denotes the likelihood of the model being estimated (fitted likelihood value), as described by Allison (2013). Thus, it predicts how much the model improves compared to the model with no predictors.

Cox and Snell R^2 is defined as: $R^2_{Cox\ and\ Snell} = 1 - (\frac{L_0}{L_M})^{\frac{2}{n}}$, where n is the sample size, where L_0 is the likelihood of the null model and where L_M denotes the likelihood of the fitted model (Allison, 2013). L_M is the product of n such probabilities. The rationale behind Cox and Snell R^2 is that for normal theory linear regression, it is an identity. The traditional R^2 for a linear regression depends on the likelihoods for the model with and without predictors by this equation (Allison, 2013).

Nagelkerke R^2 is defined as: $R^2_{Nagelkerke} = \frac{1 - (\frac{L_0}{L_M})^{\frac{2}{n}}}{1 - (L_0)^{\frac{2}{n}}}$. A big issue with the Cox and Snell R^2 is that its upper bound is $1 - L_0^{\frac{2}{n}}$, meaning that the upper bound by definition is less than 1, and can give results significantly smaller than 1, even if the model predicts the variance perfectly (Allison, 2013). Thus, Nagelkerke adjusted the Cox and Snell R^2 by dividing it with its upper bound (Allison, 2013).

8.2 Receiver Operating Characteristics (ROC)

The ROC-curves visualises the classification performance of logistic regression models (and other types of classification models). The curve plots two parameters, which are the true positive rate (TPR) and false positive rate (FPR). Thus, the graph visualizes the trade-off between sensitivity on the y-axis and 1-specificity on the x-axis (Fawcett, 2006).

The true positive rate is defined as: $TPR = \frac{True\ Positive}{True\ Positive + False\ Negative}$

The true positive rate is defined as: $FPR = \frac{False\ Positive}{False\ Positive + True\ Negative}$

Where true positive refers to a bankrupt company classified as a bankrupt company by the classification model.

Where false positive refers to an active company classified as a bankrupt company by the classification model.

Where true negative refers to an active company classified as an active company by the classification model.

Where false negative refers to a bankrupt company classified as an active company by the classification model.

The larger the AUC, the better the average classification performance (Fawcett, 2006). The ROC plot produces a diagonal line which illustrates the outcome if the observation is randomly guessed to be in one of the classifications, i.e., bankrupt or active (Fawcett, 2006). Any point above and to the left of the line indicates a higher classification accuracy than random guessing (Fawcett, 2006). The greater the difference between the ROC-curve and the diagonal line, the better the accuracy of the classification model (Fawcett, 2006). The AUC can take on values between 0 and 1, where 0 indicates a perfectly inaccurate model performance and 1 indicates a perfectly accurate model performance (Mandrekar, 2010). An AUC of 0.7-0.8 is considered acceptable, while 0.8-0.9 is considered excellent, and more than 0.9 is considered outstanding (Mandrekar, 2010).

8.3 Akaike Information Criterion (AIC)

AIC is a test used to determine which of multiple regression models is the best performing based on a given data set (Akaike, 1974). AIC-scores of multiple regression models can thus be compared to determine the best regression model (Akaike, 1974).

The formula for AIC is: $AIC = 2k - 2\ln(\hat{L})$, where k refers to number of parameters in the model and where $\hat{L}$ refers to the value of the likelihood function, i.e., how well the model reproduces the data (Akaike, 1974). The smaller the AIC-score, the better the model fit. A relative likelihood can be calculated with the equation: $relative\ likelihood = e^{\frac{AIC_{min} - AIC_i}{2}}$, where AIC_{min} refers to the AIC-score of the model with the lowest AIC-score, and where AIC_i refers to AIC-score of the compared model (Akaike, 1974). Thus, the result of this equation can be interpreted as how many times model i is probable to minimize information loss as the best performing model (Zajic, 2022).

9 Appendix B – Overview of the Six Regression Models

Table 7

Overview of the Six Regression Models

Independent Variable	Control Model (1)		Agreeableness Model (2)		Conscientiousness Model (3)		Emotional Stability Model (4)		Extraversion Model (5)		Openness Model (6)	
	Estimate	Sd. err.	Estimate	Sd. err.	Estimate	Sd. err.	Estimate	Sd. err.	Estimate	Sd. err.	Estimate	Sd. err.
Intercept	1.82***	0.01	-0.12	0.19	-1.01***	0.33	0.52**	0.22	-0.66**	0.33	-2.57***	0.44
Leverage	0.08***	0.00	0.38***	0.02	0.58***	0.04	0.27***	0.03	0.04	0.05	-0.29***	0.06
Liquidity	-8.73***	0.12	-37.49***	1.61	-74.89***	2.58	-0.75	1.72	13.08***	2.15	30.79***	3.08
Profitability	-3.81***	0.13	-3.41*	1.79	3.46	3.00	4.89***	1.82	-24.17***	2.58	-34.03***	3.72
Coverage	-4.50***	0.04	-7.49***	0.62	-15.41***	1.01	-2.27***	0.61	-1.58*	0.82	-6.01***	1.27
Activity	-0.09***	0.00	-0.21***	0.02	-0.32***	0.03	-0.22***	0.02	0.15***	0.02	0.31***	0.04
Personality trait			0.37***	0.04	0.56***	0.06	0.25***	0.04	0.21***	0.06	0.76***	0.08
Moderating Variables												
Leverage * Personality trait			-0.06***	0.00	-0.10***	0.01	-0.04***	0.01	0.01	0.01	0.06***	0.01
Liquidity * Personality trait			5.29***	0.29	12.54***	0.49	-1.56***	0.34	-3.86***	0.38	-6.86***	0.54
Profitability * Personality trait			-0.13	0.33	-1.43**	0.57	1.70***	0.35	3.59***	0.46	5.26***	0.65
Coverage * Personality trait			0.52***	0.12	2.02***	0.19	-0.44***	0.12	-0.52***	0.15	0.27	0.22
Activity * Personality trait			0.02***	0.00	0.04***	0.01	0.03***	0.00	-0.04***	0.00	-0.07***	0.01
Model Performance Measurements												
McFadden R^2	0.5834		0.5901		0.5946		0.5837		0.5842		0.5853	
Cox and Snell R^2	0.5546		0.5587		0.5614		0.5548		0.5551		0.5557	
Nagelkerke R^2	0.7394		0.7449		0.7486		0.7397		0.7401		0.7410	
AUC	0.858		0.861		0.862		0.858		0.858		0.859	
AIC	116,901		115,019		113,770		116,793		116,685		116,376	
Accuracy %	80.11		80.41		80.75		80.17		80.14		80.16	
Sensitivity %	80.98		81.04		81.15		81.06		80.96		81.10	
Specificity %	80.00		80.33		80.70		80.06		80.04		80.04	
Prevalence %	26.69		26.40		26.09		26.65		26.66		26.67	
F1-score %	47.20		47.60		48.07		47.30		47.24		47.30	

Note: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.10: *.

10 Appendix C – Linearity Assumption for Logistic Regression

The below section shows scatter plots for independent variables in all six regression models.

Figure 8

Scatter Plots Showing That the Linearity Assumptions Holds for the Control Model

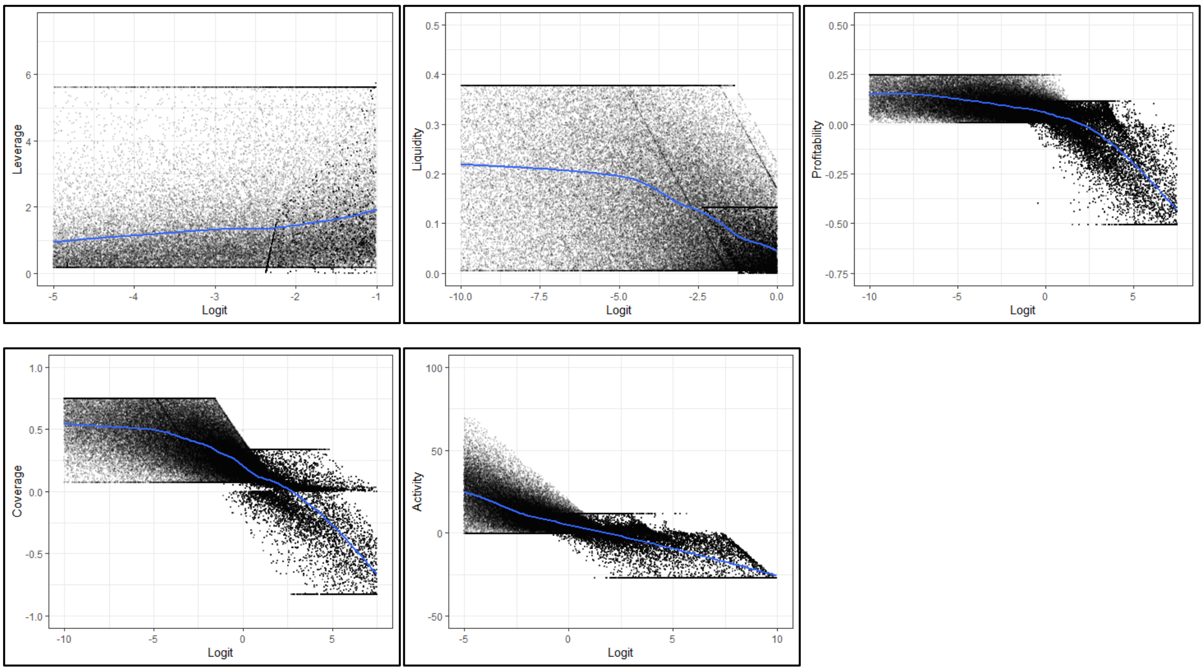


Figure 9

Scatter Plots Showing That the Linearity Assumptions Holds for the Agreeableness Model

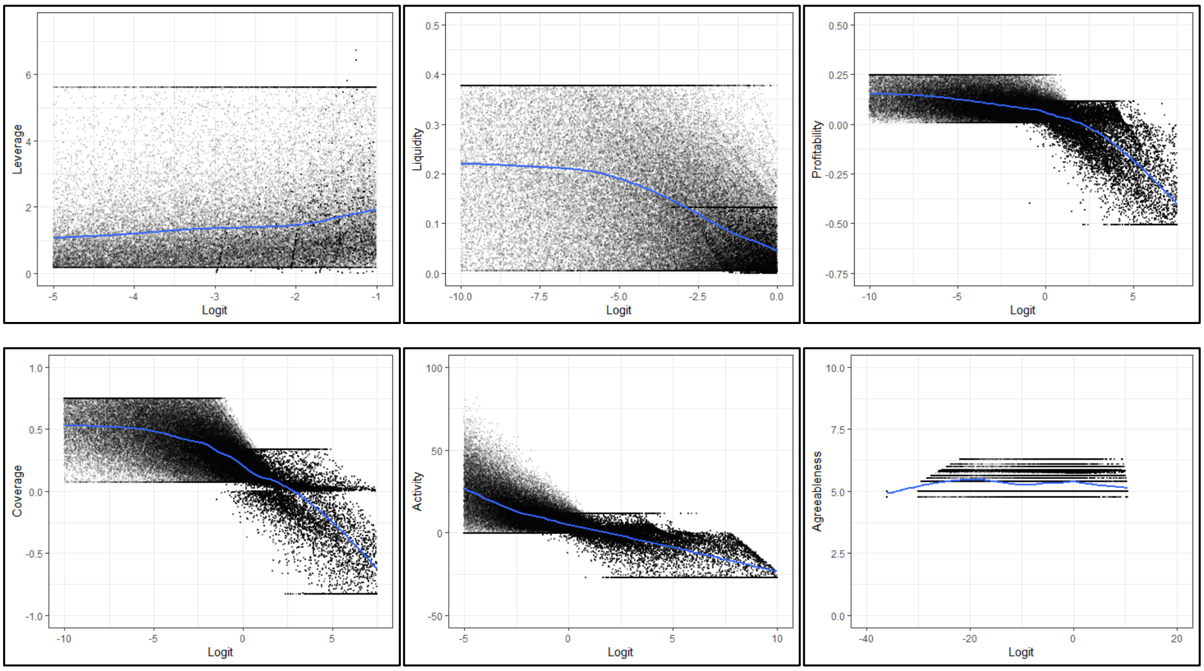


Figure 10

Scatter Plots Showing That the Linearity Assumptions Holds for the Conscientiousness Model

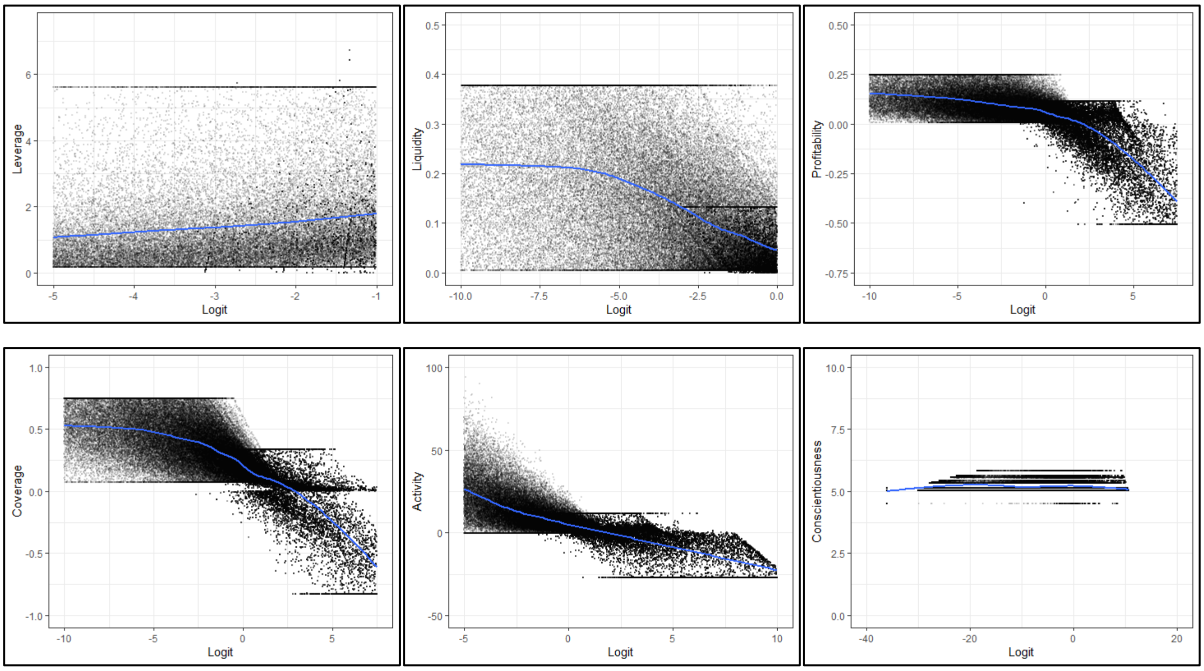


Figure 11

Scatter Plots Showing That the Linearity Assumptions Holds for the Emotional Stability Model

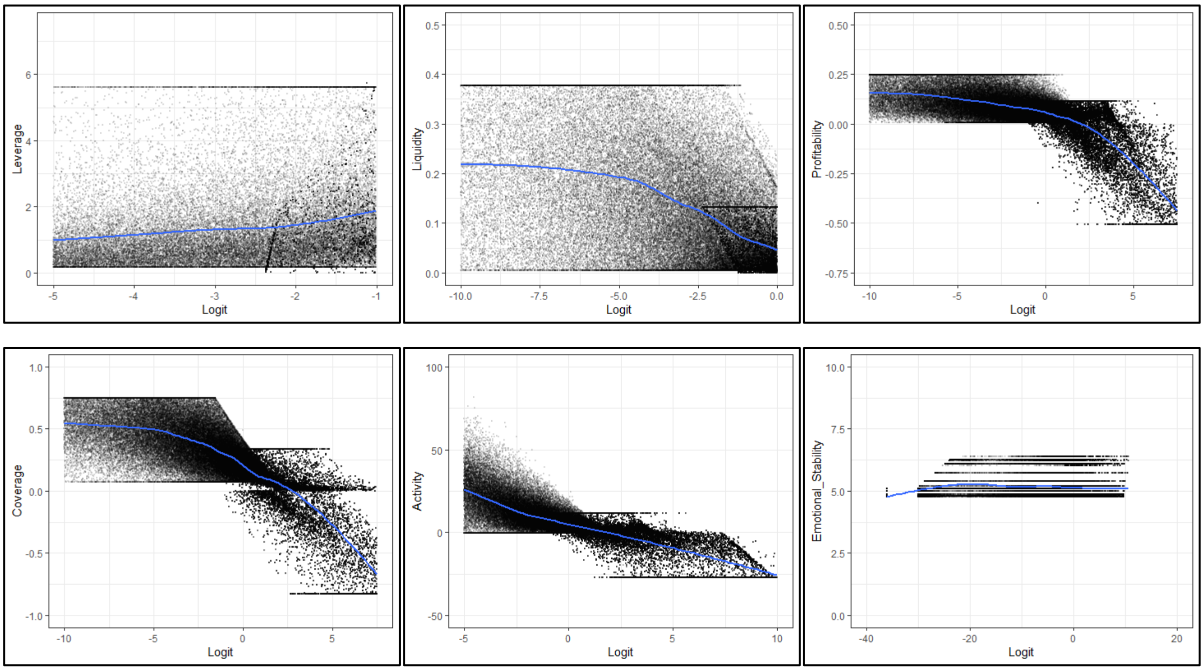


Figure 12
Scatter Plots Showing That the Linearity Assumptions Holds for the Extraversion Model

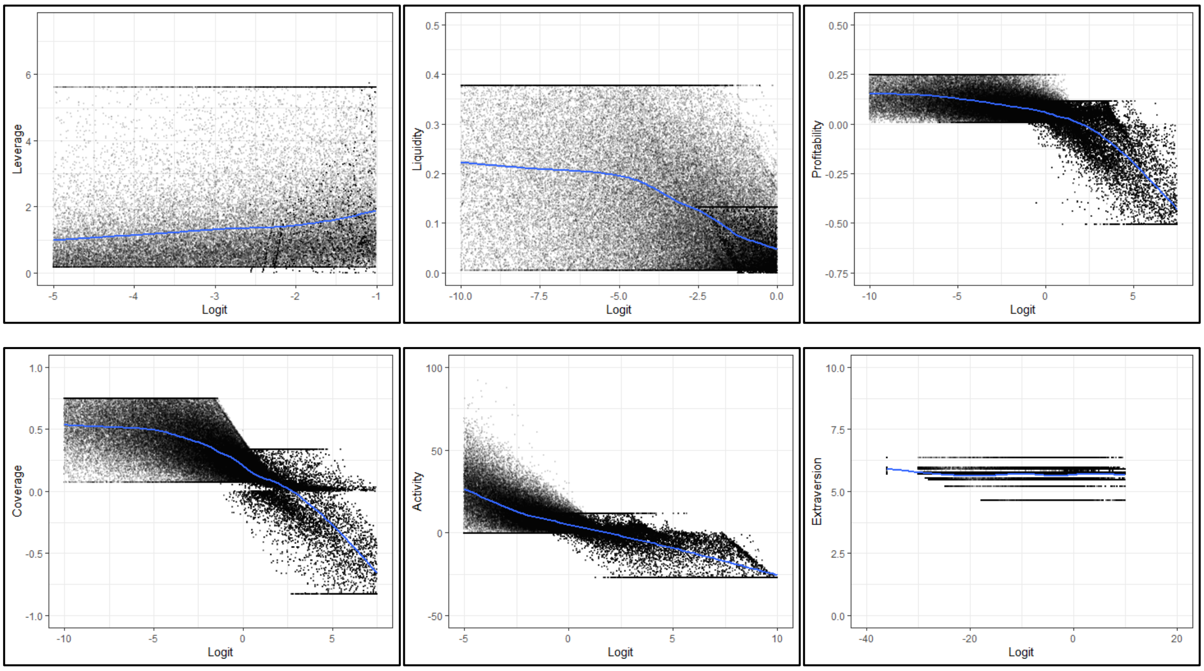


Figure 13
Scatter Plots Showing That the Linearity Assumptions Holds for the Openness Model