Probabilities of default: What is the impact of Covid-19 in Portuguese Micro-enterprises?

José Cruz

Dissertation written under the supervision of Professor Diana Bonfim, with the collaboration of industry expert Professor Nuno Silva

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, 07.04.2021

Abstract

This dissertation aims to quantify the impact of the current pandemic on Portuguese micro-enterprises credit risk. To this end, the probability of bankruptcy before and after the pandemic shock is calculated using a modified version of the structural model proposed by Eisdorfer, Goyal, & Zhdanov (2019). For this exercise, representative firms were created for each of the 14 sectors under analysis. These firms were created using a database of over 200 000 firms with observations ranging between 2006 and 2018. Sector specific shocks are calibrated using data from the quick survey on the impact of the pandemic shock on firms' business, by Statistics Portugal and by the Banco de Portugal. The results of this exercise suggest that the Accommodation and Food Services sector is the most affected sector with its probability of default within two-years increasing from 1.49% to 14.20%. On average the two-year cumulative probability of default increases by 6.46 percentage points with the current pandemic. Understanding the determinants of the shock impact, the initial liquidity level reveals to be an important factor. This dissertation demonstrates that a higher level of cash allows firms to minimize the impact of the current pandemic on the distance to default.

Title: Probabilities of default: What is the impact of Covid-19 in Portuguese Micro-enterprises?

Author: José Pedro Nolasco Amaral Gomes Cruz

Keywords: Probabilities of default; Distance to Default; Credit risk; Micro-enterprises; Covid-19

Resumo

Esta dissertação tem como objetivo quantificar o impacto da atual pandemia no risco de crédito das Micro-empresas portuguesas. Para tal, foi calculada a probabilidade de falência antes e depois do choque pandémico usando uma versão modificada do modelo estrutural proposto por Eisdorfer, Goyal, & Zhdanov (2019). Para este exercício, foram criadas empresas representativas de cada um dos 14 sectores em análise. Estas empresas foram criadas utilizando uma base de dados de mais de 200 000 empresas com observações entre 2006 e 2018. Os choques sectoriais são calibrados utilizando o inquérito rápido sobre o impacto pandémico nos negócios das empresas, publicado pelo Instituto Nacional de Estatística e pelo Banco de Portugal. Os resultados deste exercício sugerem que o sector de Alojamento e Serviços Alimentares é o sector mais afetado com a sua probabilidade de falência dentro de dois anos a aumentar de 1.49% para 14.20%. Em média, a probabilidade acumulada de falência de dois anos aumenta em 6.46 pontos percentuais com a atual pandemia. Compreendendo os determinantes do impacto do choque, o nível de liquidez inicial revela ser um fator importante. Esta dissertação demonstra que um nível mais elevado de liquidez permite às empresas minimizar o impacto da atual pandemia na distância até a falência.

Título: Probabilidades de falência: Qual o impacto da Covid-19 nas Micro-empresas portuguesas?

Autor: José Pedro Nolasco Amaral Gomes Cruz

Palavras-chave: Probabilidades de falência; Distância para a falência; Risco de Crédito; Micro-empresas; Covid-19

Acknowledgements

This dissertation culminated almost five years of learning and growth, a journey I will never forget. I will also not forget the support given by Professor Nuno Silva. I will always be grateful for his continuous commitment and dedication. Without him, this work would certainly not be the same. I also want to thank Professor Diana Bonfim for her feedback and review of this work.

Secondly, I would like to thank my father and my mother, Teresa, and Carlos, not only for allowing me to take this journey but mainly for raising me in their image. Their image will always be a reference for me and an inspiration.

Thirdly, to my brother, João, for his advice and concern towards me and to my sister Madalena for always supporting me, sometimes revealing more faith in my abilities than I.

Finally, I want to thanks to Madalena Gaspar. Your support and care will always be remembered with great fondness.

Table of Contents

List of Figures

List of Tables

1 Introduction

The first case of Covid-19 in Portugal was registered on 2 March 2020. Twenty days later, on 22 March, a state of emergency was declared, which forced thousands of Portuguese companies to abandon their physical presence with their clients. Not being prepared for this, most companies suffered a shock in their business volume. During this period, given the low capacity of businesses to generate liquidity, companies had great difficulties in meeting their obligations with business and financial partners and employees. In the months after the major lockdown, credit and government efforts helped firms tackle the shock consequences, but recovery expectations continued weak. At the beginning of February 2021, with another lockdown period, forecasted to last almost two months, the recovery expectations were lowered. For some firms, and despite government efforts, bankruptcy became the best option.

Economic shocks always have heterogeneous impacts across many dimensions. One of these dimensions is firm size with Small and Medium Enterprises (SME's) typically suffering more than others. According to the most recent statistics, in 2018, micro-enterprises assume a significant role in the Portuguese economy. They represent 96.2% of the total companies, employ 44.1% of the workers, and represent 17.6% of Portuguese companies' total turnover. For these reasons, it is fundamental to analyze the impact of the current pandemic in these companies.

This dissertation attempts to understand the severity of the shock caused by Covid-19 to Portuguese micro-enterprises credit risk. To quantify this shock, the probability of bankruptcy was used and compared with and without pandemic. To calculate the probability of default, a structural model is used. The application of structural models to the probabilities of default emerges from Black & Scholes (1973) and Merton (1974). Since then the application of structural models has mostly focused on large publicly traded firms for which the market value of equity is available. This does not have to be the case as shown in the context of equity valuation by Eisdorfer, Goyal, & Zhdanov (2019). In this model, the value assigned to assets comes from a fictitious security, whose value comes from gross margin as the state variable. Further, the authors consider two stages of growth. For its characteristics, a modified version of this model is used in this dissertation. One of the changes was to adjust the duration of the first stage. In the original article, the first period lasts five years. In this dissertation it lasts two years, to represent the duration of the crisis, before entering the steady-state. Besides, cash holdings were introduced, assigning liquidity buffers.

Although recent, the Covid-19 crisis has led to an extraordinary effort by the scientific community. One of these articles is the one by Welburn & Strong (2020). In this article, using a structural credit risk model, the authors analyzed the impact of extending the lockdown period in 3000 US listed firms. From the authors' results is understandable that extending the lockdown period leads to an increase in the average corporate default risk, which in the worst scenario can be two to three times higher than in the 2008 financial crisis. However fewer articles focus on SME’s. In this sense, it is important to highlight the work developed by Carletti, et al. (2020). The authors focused on the Italian case assuming a lockdown period of three months. The authors were able to demonstrate that small and medium-sized firms suffer a greater shock than large companies. According to the study, the authors further identifies the Manufactures and Wholesale trade sectors as the most affected. Another important paper is the one developed by Gourinchas, et al. (2020). The authors' objective was to understand the impact of Covid-19 on SME's in 17 countries. The authors' results suggest heterogeneity in impact felt at the industry level and at the regional level.

This dissertation arises to understand the impact of the current pandemic in Portuguese micro-enterprises, analyzing 14 sectors of economic activity. None of the articles mentioned above analyzed the impact on micro-enterprises or focused on the Portuguese case. With the importance highlighted above, it is fundamental to analyze the impact on these companies to fully understand the magnitude of the impact on the Portuguese economy. For this, it was necessary to access these companies’ accounting data and understand the size of the shock per sector. In both cases, the databases provided by the Banco de Portugal Microdata Research Laboratory - BPLIM were used. The accounting data was obtained through the Central Balance Sheet - Harmonized Panel (CBHP). To understand the impact of the shock, the Fast and Exceptional Enterprise Survey (COVID-IREE) was used. This database summarizes the questionnaire responses given to a representative set of companies in each sector. After understanding the impact in each sector, a series of sensitivity tests were carried out to assess the magnification of the impact on different companies.

As for the structure of this dissertation, it is divided into six chapters. In the next chapter, a review of the existing literature on probabilities of default is presented. Chapter 3 introduces the model used for the research. Chapter 4 presents the database and the variables used in the application of the model. Chapter 5 presents the results. Chapter 6 concludes this dissertation.

2 Literature Review

2.1 Probabilities of default – Credit scoring systems

The literature on credit risk is vast and mostly focuses on measuring debtors' probability of default. This is usually done using statistical models, which are commonly named credit scoring systems. In this sub-chapter, three of the principal methodologies are highlighted: linear discriminant analysis, logits and neural networks.

The application of linear discriminant analysis in credit risk has strong roots backing to Altman (1968) seminal work. In this paper, Altman applies linear discriminant analysis to 66 manufacturing companies for the period between 1946 and 1965, half of them failed. The author started with 22 ratios commonly used in the already existent literature. From these, five ratios were chosen based on their capacity to predict default. These ratios are: working capital divided by total assets; retained earnings divided by total assets; EBIT divided by total assets; book value of equity divided by the book value of the liabilities and revenues divided by total assets. The author then presented a score, which became known in the literature as the Altman Z-score. Based on this score, the author identifies three different types of companies. The ones close to default, the ones safe from default and the ones in between, could be called "grey-area". According to the author conclusions, his model accurately predicted 90% of the defaults in a time-window of 1 year. This figure was nevertheless obtained within sample. From Altman (1968) conclusions, the accuracy level decreases as the time horizon increases.

Logit models allow us to study the relation between binary variables (such as default or no-default) and some explanatory variables through a logistic function. In contrast to the Altman z-score, logit models produce probability of default estimates. Martin (1977) compared the accuracy of the logit models and the discriminant analysis models for a sample of companies between 1975 and 1976. He concluded that under a cut-off value of 0.4%, the models were equivalent in terms of accuracy (i.e., it leads to a similar percentage of correct predictions, both default and non-default). One of the most cited logit-based default prediction models is Ohlson (1980). The author based his analysis on 2163 US-listed companies between 1970 and 1976, constructing what is known as Ohlson's score. This score results from of a logistic function with nine firm-specific explanatory variables (to capture size, liquidity, leverage, and operational performance). Using a cut-off point of 0.38, the Ohlson score is reported as more precise in determining bankruptcy than the Z-score, especially for a 2-year horizon.

With the technological developments, more sophisticated statistical methods models were introduced to credit scoring. One of these is artificial neural networks. This technique allows to estimate complex non-linear relations between the variables. Altman, Marco, & Varetto (1994) analyzed 1000 healthy Italian firms between 1982 and 1992. They compared the results from linear discriminant analysis against the results of artificial neural networks (NN). The authors concluded that these two models result in similar accuracy in the one-year case. Comparing with discriminant analysis and logit regressions, the main disadvantage of NN, according to the authors, is the impossibility of recognizing, with certainty, if the variables in the model have the expected signs. Back, Laitinen, Sere, et al. (1996) compared a logit model, a discriminant analysis model, and a NN model. The authors agreed with Martin (1977) conclusions, however, according to their research, the NN model was the most accurate to predict a failure in both one-year and three-year horizons. Interestingly, for the longer horizon, the logit model was outperformed by the discriminant analysis model. The authors also investigated which variables are mostly correlated with default. They concluded that liquidity was the most important factor in predicting defaults, confirming Douglas Smith & Lawrence (1995) results. Yang, Platt, & Platt (1999) confirmed Altman, Marco, & Varetto (1994) conclusion using data from 122 oil and gas companies between 1984 and 1989. The conclusion was in line with the previous ones regarding the efficiency of the NN models. Nevertheless, the authors state that discriminant analysis models outperform the NN's when minimizing the error of false negatives.

Although these models are among the most widely used, Altman & Saunders (1997) highlight three main issues. Firstly, these models mainly use book value accounting data, so they cannot capture rapid changes in the borrower's conditions. Secondly, the real world is too complex to assume that a linear discriminant analysis can capture its complexity. At last, these models seem to be disconnected from the theoretical models.

2.2 Applicability of structural models on probabilities of default

Due to the limitations described in the previous sub-chapter, more complex and theoretical-based models started to be used. The majority of these models have their roots linked to the models presented by Black & Scholes (1973) and Merton (1974). The Black and Scholes base model was developed to price European options on public-traded firms. As the authors explain, their model only requires either observable or relatively easy to estimate inputs, namely, the strike price, the current share price, the time to maturity, an estimate of equity return volatility, and the risk-free. Behind their model is the idea that an investor can instantaneously replicate an option by buying or selling the underlying stock. This leads to option prices formulas that do not depend on preferences. Merton (1974), expanded this model into stock and bond pricing. Merton considers that in the case of a firm financed by equity and a single pure discount bond with a pre-determined liquidation date, its equity value is equivalent to a call option on the firm assets with strike equal to nominal debt. If the value of the assets becomes lower than the debt at maturity date, the company defaults. As Black and Scholes did, Merton considers that the value of the company assets follows a geometric Brownian motion (GBM) and thus, the Black-Scholes formulae can be used in his setting. In addition, the probability of default corresponds to the probability of the call option ending up out of the money.

After this initial breakthrough, the literature focus has been on the improvement of the Merton model. Black & Cox (1976) developed a model similar to the one proposed by Merton, but that allows a company to default any time before the firm liquidation date. For this reason, their model has been coined as a “first passage time” model. The authors justify this possibility because debtholders can ask for the firm’s insolvency before maturity to maximize their claim when default. Since, in reality, default can happen at any time, this change approximates the model to reality. Consequently, most of the literature that follows adopts first passage time models. This is Leland (1994) case, which improved the model by adding taxes on corporate profits and considering that the firm has no pre-determined liquidation date. Leland (1994) also compared two rationales for determining the default barrier, notably, an exogenous barrier equal to nominal debt and an endogenous barrier chosen by equity holders to maximize their claim. Leland (1994) focuses on the optimal capital structure of the firm. Brockman & Turtle (2003) proposed a slightly different model on the credit risk literature where they see the common share as a down-and-out call (DOC) option on the company assets. Reached the barrier, the debtholders would be entitled to the company assets.

Back to the optimal capital structure literature, Goldstein, Ju, & Leland (2001) adapt the Leland (1994) model in three ways. Firstly, they include the possibility of increasing the debt whenever the asset value grows sufficiently. Secondly, they use a closer to reality tax setup corporate income taxes, interest income taxes, and dividend taxes. Thirdly, they consider the firm assets as a fictive security, whose value comes from the firm’s capacity to generate earnings (EBIT). Despite this security not being traded, its value is considered observable as long as there is at least one traded contingent claims (e.g., equity). The fact that the asset value is directly linked to the firm capacity to generate earnings opens the door to apply these models to equity valuation. This is the case of the paper by Eisdorfer, Goyal, & Zhdanov (2019). These authors enhanced the previous model in three ways. Firstly, instead of using the EBIT as the state variable of the model, they choose to use the gross margin since the EBIT is often negative, which is not compatible with the GBM assumption. Secondly, they consider a two-stage growth model, allowing for a better fitness in non-mature firms. Thirdly, they consider the impact of financial distress costs before the company defaults.

All models presented so far assume that company assets' follows a continuous diffusion-based process, specifically a GBM. Sarkar & Zapatero (2003) applied a mean-reversion continuous process to the company earnings to understand the implications on the optimal capital structure. Intuitively, it is no longer assumed that the firm log-value follows a random walk, but instead, it converges to some mean. As a result, periods of higher growth are expected after lower growth periods. As its known from practice, some sectors, such as construction or tourism sector, witness this cyclical effect. By doing this, the authors demonstrated a negative correlation between earnings and leverage, which goes in line with the empirical evidence (something that the typical trade-off theory is not able to explain). As a suggestion for future research, it could be interesting to understand the impact of this assumption on probabilities of default.

Though an interesting extension, mean-reverting models cannot address the criticism in the literature regarding structural models' underestimation of credit risk in the case of investment-grade firms for short maturities. Indeed, assuming that the company's value follows a continuous process precludes the company's value to drop suddenly as it happens from time to time in reality. As such, it is impossible for an investment grade firm to default in a short time period. In Zhou (2001), the author stressed this hypothesis by allowing jumps on the firm value's continuous evaluation process. He demonstrated that this new model is much more flexible than previous ones allowing for a theoretical explanation of large credit spreads, especially for short maturity bonds. However, to introduce these jumps, it would be necessary to run Monte-Carlo simulations, not allowing the use of closed-formulas.

A popular alternative to address the underperformance of diffusion-based structural models in the case of short-term investment-grade debtors is the commercial model developed by KMV. In 1980's KMV Corporation adapted the Merton model in two major ways. Firstly, instead of evaluating the distance to default on the Normal distribution, they started using their empirical distribution. In this way, the KMV model can avoid unreasonable probabilities of bankruptcy (either too high or too low) that results from the Merton model. Secondly, they considered setting the option's trike price equal to the current liabilities and half of the non-current liabilities. Interestingly, this leads to barrier values more in line with endogenous barrier models. Later in 2002, Moody's Corporation acquired KMV to expand its credit risk management products.

The literature on the empirical evaluation of the capacity of structural models to predict default is relatively vast. In this dissertation, the papers by Brockman & Turtle (2003), Reisz & Perlich (2007) and Duffie, Saita, & Wang (2007) are explored. Brockman & Turtle (2003) explore the performance of their DOC model. They concluded that their model has a significant predictive power, dominating the Z-score by Altman in most cases. Reisz & Perlich (2007) confirmed this conclusion by looking at 5784 industrial firms between 1988 and 2002. However, they concluded that this is only true for long-term horizons. In the one-year horizon, the DOC framework is outperformed by the Altman Z-score (and similar accounting-based). Besides this confirmation, they also confirmed the superiority of this model regarding the Merton model. Duffie, Saita, & Wang (2007) showed the high explanatory power of distance to default (the measure proposed by Merton and following authors) on firms default. To do so, the authors used monthly information of 2770 companies between 1980 and 2004. Moreover, they highlighted the increase in predictive power when including macroeconomic variables (such as the risk free and the returns of the S&P500).

2.3 Studies on probabilities of default on SME's

It is important to highlight that throughout the two previous subchapters, all the articles mentioned have their focus on large public corporations. Since the object of analysis in this thesis is Portuguese Small and Medium-Sized Enterprises (SME's), specifically the micro-enterprises, it is fundamental to understand what characterizes these firms.

Rikkers & Thibeault (2009) consider that these companies have a set of characteristics that make them more vulnerable than Large-sized/Public companies, resulting in higher probabilities of default. According to these authors, these firms typically provide less information than public firms. Contracts are kept in private, and their suppliers are unknown. The quality of the information provided by these companies also tends to be lower due to lower accounting standards (frequently not audited) and the fact that personal expenses being frequently accounted as business expenses. As a result of these features, the financial statements figures tend to be more volatile and opaque. Hyytinen & Pajarinen (2008) add that firms with financial difficulties tend to hide their difficulties from external creditors, further contributing to the already mentioned opacity. According to these authors, since SME's operate on a smaller scale, they also do not have strategic positions with their suppliers and business partners and are less likely to have a valuable know-how. Consequently, they become more exposed to business cycles, which leads to higher chances of default.

The literature on the application of multivariate credit scoring systems for SME's is vast. In this dissertation Edmister (1972) and Altman & Sabato (2007) works are explored. Edmister (1972) studies the usefulness of financial ratios to predict small businesses default. To do so, he used 192000 financial statements between 1954 and 1969. He used seven firm-specific explanatory variables to capture size, leverage, liquidity, among other characteristics. Using a synthetic sample validation approach the model was able to discriminate 32 of the 35 cases correctly. Altman & Sabato (2007) used an accounting-based logit model to analyze the one-year default probability of 2000 U.S companies between 1994 and 2000. Their model was 30% more accurate than the generic Z-score model. The authors state that the use of qualitative information could be a source of improvement of the model. As pointed previously, the lack (and quality) of information from the SME's is one of the main differences from the typical large-sized company.

The literature on structural models applied to SME's is, however, scarcer. As these firms usually do not have public traded securities, it is difficult to apply most structural models without some adaptations. One exception is Rikkers & Thibeault (2009) work, which studied a structural model performance for a sample of 1238 Dutch SME's, from which 240 defaulted. They claim that their model correctly predicts 63.8% of the defaults out-of-sample. When compared to commonly used ratios, such as profitability, solvency, and size, they concluded that the model could add predictive power.

2.4 The impact of the pandemic crisis on SME's failure probabilities

Despite recent, there is already literature on the impact of Covid-19 on the SME's. One of these articles is the one presented by Carletti, et al. (2020). In this article, the authors study the impact of the current pandemic crisis on Italian companies. To do so, they collected a sample of 80000 Italian firms and analyzed the implications of a 3-month lockdown. The authors assumed that during this period both revenues and labor-costs became lower (due to the support of the Italian government). In the Italian case, the authors concluded an impact on profits close to 170 billion euros and an impact on equity value close to 148 billion euros. Due to the drop in profits, about 17% of the firms on the sample became distressed, impacting 9% of the employment in the sample. The authors also concluded that the shock had a very heterogeneous impact across companies. As expected, the impact on small and medium firms was strong, 18.1% and 14.3% became distressed respectively, compared to larger firms, 6.4%. The impact has been felt differently across sectors of activity. The Manufactures and Wholesale trade sector are the sectors with the highest shortfall. However, the authors arrived at a curious result. According to the study, the tourism sector is one of the least affected. For the authors, this was because it is a labor-intensive sector with low fixed costs (when not accounting for wages).

In the previous article, the obtained results were strongly affected by the consideration of government support. Gourinchas, et al. (2020) focus was to estimate the impact of Covid-19 on SME's liquidity in 17 countries, without government intervention. To do so, they used a partial equilibrium economic model calibrated using the ORBIS global database. Their exercise is done considering a lockdown period of 8 weeks. During this period, they consider that the economy suffers a supply and a demand contraction (allowing for sector-specific demand shocks), which translates into a substantial reduction in firms capacity to generate cash flow. In the 9th week, the supply and the productivity go back to the pre-Covid levels, but the demand evolves according to the IMF projections. During these weeks, if the available cash and the future cash-flow are lower than the fixed costs, taxes, and financial expenses the authors assumed the company would face a liquidity shortfall since they would not be able to borrow. Under this liquidity problem, a firm would be insolvent. This last assumption creates a clear distinction between this model and some of the structural contingent claims models presented in subchapter 1.2, notably the ones by Goldstein, Ju, & Leland and Eisdorfer, Goyal, & Zhdanov. In the models presented above, if necessary, the shareholders inject money into the company to maximize their welfare. From their research, they were able to conclude an 8.8 percentage point increment of the default rate due to Covid-19, from 9.4% to 18.2%. However, the impact of the current pandemic is not uniform between sectors or between regions. In contrast with the previous study, the Accommodation and food service is one of the most affected sectors (probably because government intervention was disregarded) together with the Entertainment and recreation and with the Educational sector. The authors also identify heterogeneity between regions, Portugal being the 6th most-affected country.

3 Model

This chapter explains the model used in this dissertation to assess the pandemic crisis's impact on the distance to default and the probability of default of Portuguese SME's. The model used is a modified version of the option pricing-based equity valuation model proposed by Eisdorfer, Goyal and Zhdanov (2019), EGZ hereafter. In section 3.1, the model is explained in-depth. Section 3.2 covers the changes made to the initial model, and, in section 3.3, it is explained how the impact of the shock in analysis is measured.

3.1 Eisdorfer, Goyal and Zhdanov original model

In EGZ model, it is considered that the firm holds a project that generates cash flow to the shareholders according to the following equation:

Equation 1

$$CF_{i,t} = \left[(1-\tau)\left(x_{i,t} - Int_{i,t} - Foc_i\right) + \tau Dep_{i,t} - Capex_{i,t}\right] * \left[1 + \eta \mathbf{1}_{(1-\tau)\left(x_{i,t} - I_{i,t} - F_i\right) + \tau Dep_{i,t} - Capex_{i,t} < 0}\right] - D_{i,t}$$

where $x_{i,t}$ represents the gross profit of firm *i* at time *t* (i.e. revenues less cost of goods sold), $Int_{i,t}$ is total interest payments to debtholders due at time t, Foc_i is the total operating fixed costs incurred by the company, τ is the corporate tax rate, $\tau Dep_{i,t}$ is the tax shield due to depreciation expenses, $Capex_{i,t}$ refers to the capital expenditures, and finally, $D_{i,t}$ is the principal repayment due at time t. If the cash flow is negative, it is assumed that the company incurs in financial distress costs, amplifying this negative cash flows by a percentage η. Since, the model does not correctly consider the existence of cash-holdings, every project cash-flow represents an equal cash-flow to the shareholders. It follows that, when the cash flow is positive, the shareholders receive a dividend and when it is negative, the shareholders choose between a capital injection or firm liquidation.

It is assumed that $x_{i,t}$, the model state variable, follows a geometric Brownian motion with $\mu_{i,p}$ as the drift parameter, σ_i as the volatility and W_t^p is the Wiener process:

Equation 2

$$\frac{dx_{i,t}}{x_{i,t}} = \mu_{i,p} dt + \sigma_i dW_t^p.$$

In order to price contingent claims, it is useful to present the process under the risk neutral measure, Q. This is done by subtracting to $\mu_{i,P}$ the market price of risk times the volatility ($\overline{m}_i * \sigma_i$):

Equation 3

$$\mu_{i,Q} = \mu_{i,p} - \overline{m}_i * \sigma_i$$

Regarding depreciations and capex, EGZ write these as linear functions of $x_{i,t}$. They do this in two steps. First, they assume a constant gross margin ratio implying that future revenues are proportional to the state variable:

Equation 4

$$GMR_{i,t} = \frac{Revenues_{i,t} - COGS_{i,t}}{Revenues_{i,t}} = \frac{x_{i,t}}{Revenues_{i,t}} = \overline{GMR_\iota}$$

Second, the authors assume that depreciations and capex are also kept as a constant proportion of revenues:

Equation 5

$$Capex_{i,s} = Revenues_{i,s} * \overline{CSR_\iota} \ ,$$

$$Dep_{i,s} = Revenues_{i,s} * \overline{DSR_\iota} \ , \text{with } s \geq t.$$

EGZ set $\overline{CSR_\iota}$ and $\overline{DSR_\iota}$ as the last three years industry average of capex over revenues and depreciation over revenues. Substituting "Revenues" by $x_{i,s}/\overline{GMR}$, it is possible to write these variables as linear functions of our state variable, making it easier to derive contingent claims prices.

Equation 6

$$Capex_{i,s} = \frac{x_{i,s} * \overline{CSR_\iota}}{\overline{GMR_\iota}}$$

$$Dep_{i,s} = \frac{x_{i,s} * \overline{DSR_\iota}}{\overline{GMR_\iota}}$$

For the growth rate of the state variable under the physical measure, $\mu_{i,p}$, the authors use a standard approach in the literature. They assume that the capex invested during a certain period of time generates an operating cash flow of $Capex_{i,t} * R_A dt$. Under this hypothesis it is possible to show that $\mu_{i,p}$ is given by, *Equation 7*

$$\mu_{i,p} = \frac{\mathbb{E}_t^p(dx_{i,t})}{x_{i,t}dt} = \frac{\overline{CSR} * R_A}{(1-\tau)GMR_{i,t} + \tau\overline{DSR}},$$

where R_A is the return on assets, which in this model can be written as:

Equation 8

$$R_A = r + \overline{m} * \sigma$$

As explained in the previous chapter, this model has two different stages. On the first stage, until the fifth year, the firm issues two types of debt: short-term debt (maturity of one year) and long term-debt (maturity of five years). For the short-term debt, the authors are not clear whether this debt is then rolled-over. The coupon on the long-term debt is paid annually with the first coupon payment due in one year. The coupon rate is computed using the risk-free rate plus the yield-spread corresponding to the level of risk of the firm (given by the credit rating). On the fifth year, if the company is solvent, shareholders pay all the debt and issue new debt, this time perpetual debt with constant coupon. The amount received from the debt issue is used to pay an extraordinary dividend. The amount of debt issued is chosen to match the industry leverage ratio.

The objective of EGZ is to find the equity value under this model. Intuitively, this corresponds to the discounted sum of all futures expected dividends until the process reaches the default threshold, x_d:

Equation 9

$$E_{i,0}(x_0) = \sup_{T_{x_d(t)}} \boldsymbol{E}_{x_0}^{Q} \int_0^{T_{x_d(t)}} e^{-rt} CF_{i,t} dt$$

The value of equity at time zero is found by computed all expected cash flows under the risk-neutral measure and discounting these at the risk-free rate. Risk aversion is taken into account by changing the probability space rather than discounting at a risk adjusted rate. In Equation 9, x_d (t) is the optimal default barrier (i.e. the level of gross profit at which the shareholders give up the firm) and $T_{x_d(t)}$ is the first passage time of the process x to the boundary $x_d(t)$. In this model, the default barrier is a function of time because the coupons are not continuously paid, and debt has a final maturity (five years).

The fact that x_d is constant after the perpetual debt issuance (stage 2) lead EGZ to find equity value at time zero following a backwards strategy. This is done in four steps. First, the authors compute equity value at the end of the fifth year, after the refinancing. Since debt is perpetual, the endogenous default boundary is now constant allowing the authors to find almost closed-formulas for the equity value. Second, they compute the value of the perpetual debt issued on the fifth year. This is also the amount of the extraordinary dividend that shareholders receive at this time in case the firm is solvent. Third, they compute the equity value before the issuance of the new debt and the dividend payment. This corresponds to the unlevered firm value. Fourth, having the value of equity at the end of fifth year, before the stage change, the authors use a binomial tree to compute the value of equity at year zero. Once again this is done backwards. Explaining each of these points in detail in detail:

1. The authors state that the value of equity at time *t=5* is given by:

Equation 10

$$E(x_{i,t}) = \begin{cases} Ax_{i,t}^{\beta_1} + Bx_{i,t}^{\beta_2} + \left[\frac{\tau\overline{DSR}-\overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{x_{i,t}}{r-\mu_Q} - (1-\tau) * \frac{Int_i+Foc_i}{r-\mu_Q} & \text{if } x_{it} \geq x^* \\ Ax_{i,t}^{\beta_1} + Bx_{i,t}^{\beta_2} + (1+\eta)\left\{\left[\frac{\tau\overline{DSR}-\overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{x_{i,t}}{r-\mu_Q} - (1-\tau) * \frac{Int_i+Foc_i}{r-\mu_Q}\right\} & \text{otherwise} \end{cases}$$

Where $x*$ is the minimal value of x_i before the financial distress costs penalty,

Equation 11

$$x^* = \frac{(Int_t + Foc_t)(1-\tau)}{1 - \tau + \frac{\tau * \overline{DSR} - \overline{CSR}}{GMR_{i,t}}}$$

and β_1 and β_2 are the roots of the quadratic equation $\frac{1}{2}\beta\sigma^2(\beta-1) + \mu_Q\beta - r = 0$. A, B, C, D, and x_d are constants and found through a five equations system (see appendix equation A1). This system of equations make sure that no bubbles comes up with x increase (first equation); the value functions and their first derivatives match at x* (second and third equation); at last, they allow us to compute the optimal default threshold x_d at the end of fifth year (fourth and fifth equation).

2 and 3. After solving the system of equations, the value of equity at year five is discovered, i.e. immediately after issuing the perpetual coupon bond. The cash inflow associated with the debt issuance at year five ($D(x_{i,5})$) can be computed through Equation 12:

Equation 12

$$D(x_{i,5}) = \frac{Int_i}{r} + \left(\frac{Int_i}{r}\right)^{\beta_2}\left[(1-\alpha)(V_u(x_d) - \frac{Int_i}{r}\right],$$

where α is the cost of bankruptcy and $V_U(x_d)$ is the value of the unlevered firm.

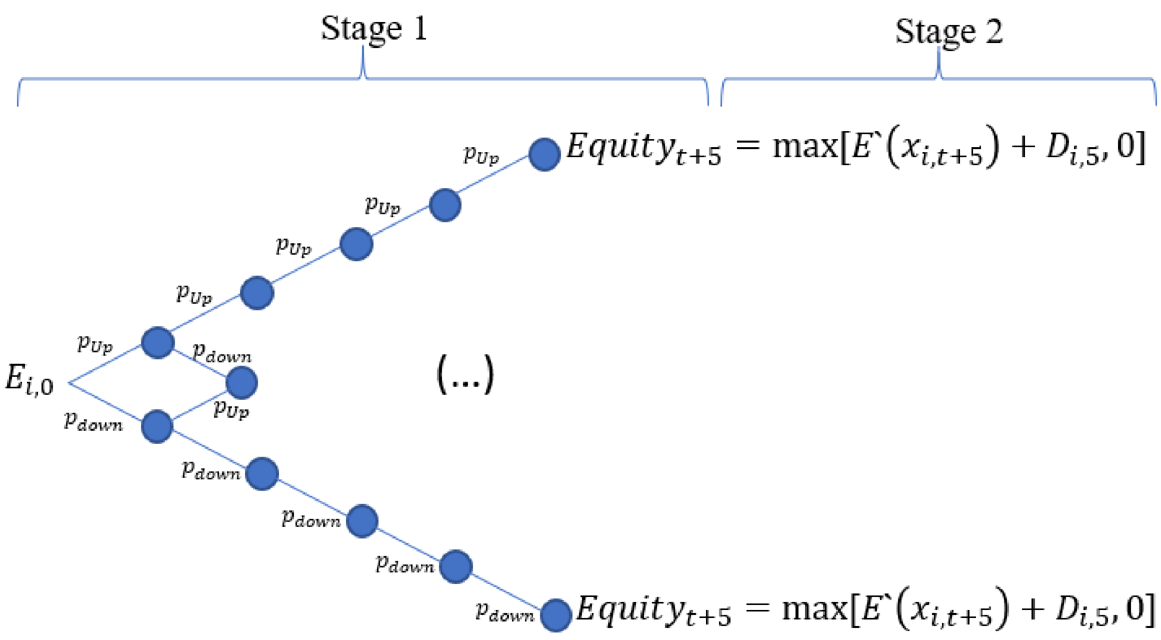
With this it is possible to compute the value of equity before the re-financing ($E'(x_{i,5})$), as in Equation 13,

Equation 13

$$E'(x_{i,5}) = E(x_{i,5}) + D(x_{i,5})$$

4. Once the equity value at year five before the re-financing ($E'(x_{i,5})$) is computed, the authors compute the value of equity at time zero using a binomial tree, again following the risk neutral pricing approach, as illustrated in Figure 1.

Figure 1- Binomial tree



In each node, the equity value is given by the maximum of the expected value of the next two nodes or zero. Since shareholders have limited responsibility, zero is the minimum value of equity. In this case, the value of equity is zero if the gross margin is equal or smaller than the respective x_d (i.e when the default barrier is met):

Equation 14

$$Equity_t = \max\left(PV(Equity_{t+dt}), 0\right), \text{ with,}$$

$$PV(Equity_{t+dt}) = e^{-rdt}[p_{up}Equity_{t+dt}^{up} + p_{down}Equity_{t+dt}^{down}],$$

where,

Equation 15

$$p_{up} = 0.5 + \left(\mu_Q - 0.5\sigma^2\right)\frac{\sqrt{dt}}{2\sigma}\ , p_{down} = 1 - p_{up}\ , U = e^{\sigma\sqrt{dt}}\ , D = \frac{1}{U}$$

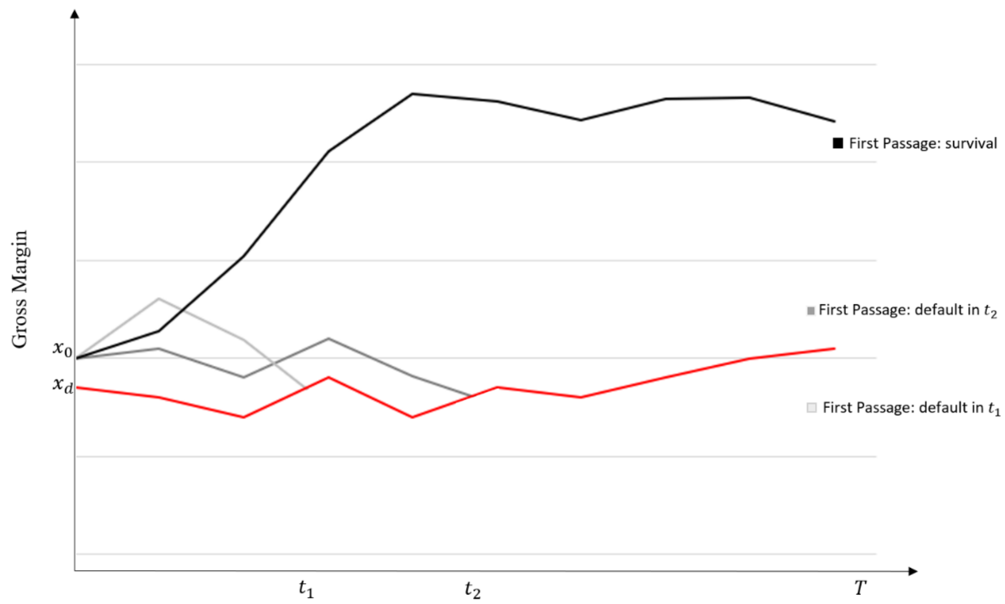
The EGZ model was designed for equity valuation. Nevertheless, one can also use the model for credit risk. As in most structural models, the distance to default (DD) at time T can be computed as:

Equation 16

$$DD = \frac{\ln\left(\frac{x}{x_{d(T)}}\right) + \left(\mu_p - 0.5 * \sigma^2\right) * T}{\sigma * \sqrt{T}}$$

The computation of the probabilities of default is less straightforward, though. As explained before, the default barrier, x_d, is a function of time. For this reason, there is no explicit formula for computing the probability of default as in Merton model. Moreover, it is not possible to compute the probability of default on a first-time passage model using a binomial tree. This occurs because there are many paths to reach the same node. Figure 2 illustrates two defaults using the presented model:

Figure 2- First passage models default timing



An alternative is to follow the below listed steps:

1. Use the binomial tree to compute the x_d that is associated with each moment in time. This can be done by finding the higher level of $x_{i,t}$ that leads equity value to be zero at each moment in time.
2. Use Equation 2 to simulate a large number of potential gross profit trajectories.
3. Since the EGZ model is a first passage time model, the probability of default until time T is computed by counting the number of times that the simulated values fall below $x_{d,t}$ at any moment in time between t=0 and T.

3.2 Model extensions

During this chapter, further explanation on model changes are provided. These changes are not specific to this dissertation's exercise and so they could be used in different circumstances. These changes are: 1. the introduction of cash-holdings; 2. the removal of the distress costs; 3. the introduction of a more detailed cost structure; 4. a simplification of the debt dynamics, and finally, 5. The introduction of a tax on dividends.

1. **Introduction of cash-holdings in the first stage of the model.** Since the Covid-19 crisis started as a liquidity crisis, cash-holdings were introduced in the model. During the first stage, the firm cash-holdings correspond to the maximum between zero and the previous cash-holdings plus the corresponding moment cash-flow. It is assumed that during this first stage the shareholder does not receive any dividend. However, if the cash-holding becomes zero and the cash-flow is negative the shareholder must inject new capital (on the same amount as the negative cash-flow) or, alternatively, let the firm default, similar to EGZ model. At the end of this stage, the cash-holding is set equal to some minimum cash level (see section 4.3 on how this level is chosen). In this way, depending on the cash-holdings value, the firm may have an extraordinary dividend or a capital injection (done by the shareholders) at the end of stage 1. Importantly, cash-holdings are introduced in the model by means of a heuristic. Notice that the cash-holdings value on the up and down case of the binomial tree can be different from the down and up case, something that cannot be correctly accounted in a simple binomial tree. As a solution, in both cases the cash holdings at the beginning of each period are set as the simple average of the two cases (further explanation available in the appendix A2).

2. **Removal of distress costs.** To simplify the model, the distress costs whenever the free cash flow is negative are not considered. In this way, a distress term is no longer present on the cash-flow equation and the piecewise-defined equity valuation function as Equation 10 is no longer valid. Therefore, its now possible to use closed formulas (i.e. there is no need to solve the system of equations in A1).

3. **More detailed operating cost structure.** Operating costs outside the gross profit equation are divided in two: fixed operating costs (Foc) and "Other operating costs" (Ooc). Two differences justify this change. First, though widely used in the literature, SG&A may underestimate firms' flexibility in adjusting their cost structure. Second, for the companies under analysis it is not possible to use the SG&A (data is not provided). In this way, the cash flow of a company i at time t is given by,

Equation 17

$$CF_{i,t} = (1-\tau)x_{i,t} + \tau Dep_{i,t} - (1-\tau)(Int_{i,t} + Foc_{i,t} + Ooc_{i,t}) - Capex_{i,t}$$

The difference between Equation 1 and Equation 17 is the inclusion of Ooc and the exclusion of the distress costs.

Similar to EGZ, its assumed a constant gross margin ratio, capex to revenues ratio and depreciations to revenues ratio. In addition, its also considered that the "other operating cost" to revenues ratio is also constant and given by its mean historical values. With this, in addition to Equation 6, $Ooc_{i,t}$ is given by:

Equation 18

$$Ooc_{i,t} = \frac{x_{i.t} * \overline{OSR_\iota}}{\overline{GMR_\iota}}$$

By doing this, the cash flow of the company *i* at time *t* is given by,

Equation 19

$$CF_{i,t} = x_{i,t}\left(1 - \tau + \tau * \frac{\overline{DSR}}{\overline{GMR}} - \frac{\overline{CSR}}{\overline{GMR}} - (1-\tau)\frac{\overline{OSR}}{\overline{GMR}}\right) - (1-\tau)\big(Int_{i,t} + Foc_{i,t}\big)$$

4. **Simpler debt dynamics.** In EGZ original model the authors assume three types of debt. During the first stage, a short debt (yearly issue) and a long-term debt (issue in the beginning of the first stage and fully paid at the end). At last, a perpetual debt, issued to meet the industry average leverage ratio. To simplify the debt dynamics, it is assumed that corporate debt is perpetual since period zero.

As result of these changes, the equity valuation formula used in the end of the first stage, after adding the tax on dividends, can be expressed as (complement in the appendix A3):

Equation 20

$$E_{i,0}(x_{i,0}) = (1 - \tau^{Div}) * [(1 - \tau^{Corp})(A_{Solv} - A_{Int} - A_{Foc} - A_{Ooc}) + \tau^{Corp} A_{Dep} - A_{Capex},$$

where A_{Solv} is the value of all future gross margins as long as the company stays solvent. A_{Int}, A_{Foc} and A_{Ooc} are, respectively, the value of the debt holders interest claim, the fixed costs suppliers claim, and the other operating costs suppliers claim on the company assets between today and the default day. Equivalently, $\tau^{Corp} A_{Dep}$ is the added value due to the depreciations tax shield, and A_{Capex} is the sum of all future investments in capex between today and the insolvency day.

The default barrier x_d in the second stage of the model is found by using the smooth pasting condition (Equation 21):

Equation 21

$$\left.\frac{\partial E}{\partial x}\right|_{x=x_d} = 0$$

This is done in three steps. First, by finding the expression of equity (Equation 20). Second, by doing the first derivative in respect to x_i. Third, substituting x_i by x_d and isolate it (complement in the appendix A4).

3.3 The introduction of the pandemic shock

As stated before, the aim of this dissertation is to understand the impact of the current pandemic on Portuguese SME's default probability. To do so, the EGZ model is adapted to include, 1. an initial shock on the gross margin; 2. a two-years first stage, and 3. a specific gross margin growth rate for the first period.

1. **Initial shock**. To understand the impact of the current pandemic, two cases are considered. First, considering that $x_{i,1}$ corresponds to the last observable x_i, i.e. considering the first gross margin to revenues ratio equal to the one on the last observable year. Second, considering that at the begging of the first stage (March 2020-represeting the beginning of the lockdown) companies face a shock on their gross margin. The impact of this shock is given by *k*. This initial shock aims to capture the impact of the Covid-19 on the firm gross profit. To perform a more

realistic exercise, a different k is assigned to each sector based on sector expectations.

Equation 22

$$x_{i,1} = e^{-k_i} * x_{i,0}$$

2. **Shorter first stage**. In the original EGZ, the first stage is set to last five years. Following Gourinchas, et al. (2020) expectation the time horizon for the first stage was set at two years. According with the authors research[1], the economy is expected to take 50 weeks to closely recover half of the total demand lost due to the lockdown. Following this rationale, within 100 weeks, closely two years, its expected a full recover.

3. **Growth rate**. During the first stage, the gross margin growth rate is denoted by $\mu^{P,St}$,

Equation 23

$$\mu^{P,St} = \frac{k}{N}$$

where N denotes the number of years during the first stage, in this case two. This growth rate is used on the GBM's during the Monte-Carlo simulations and on the binomial tree as explained in chapter 3.1.

4 Data

In this chapter, a description of the initial dataset used and the cleaning methodology is given. Second, a description of the final dataset is provided. Third, an explanation of how the model is calibrated. Fourth, two of the model assumptions are tested.

4.1 Data and data cleaning

The sample used in this dissertation was provided by Banco de Portugal Microdata Research Laboratory - BPLIM. The dataset used is designated "Central Balance Sheet Harmonized Panel Data (CBHP)". It contains annual longitudinal data on the Portuguese firms, covering the 2006 to 2018 period. As 2019 data was not available, the 2018 values are used as

[1] According to Figure 4 in the report

proxy. The initial dataset contains more than five million observations belonging to almost 730 thousand companies. Each observation has the company correspondent anonymization code, the year, the sector (according to CAE-Rev. 3), the size, and the remaining financial information provided by the annual company reports. This dissertation analyzes the impact in the micro-sized companies, and for this reason the remaining dimensions are excluded. According to the European Commission (2003), the Micro-sized companies have less than ten staff members and a turnover lower than two million euros or total assets lower than two million euros.

As explained in the literature review, SME's data is typically very noisy and thus some data treatment is required. Table 1 summarizes the data cleaning process.

Table 1-Cleaning process results

Step	Before		Change		After	
	Number of companies:	Number of observations:	Number of companies:	Number of observations:	Number of companies:	Number of observations:
Removed the observations with sales either negative or 0	727 944	5 028 327	-84 083	-936 734	643 861	4 091 593
Removed the observations with negative gross margin (given by sales-cogs)	643 861	4 091 593	-8 832	-97 421	635 029	3 994 172
Removed companies with less than 5 observations	635 029	3 994 172	-288 892	-663 417	346 137	3 330 755
Removed financial and public administration companies	346 137	3 330 755	-23 319	-258 509	322 818	3 072 246
Ensured that each company has only one sector/dimension and re-allocate the observations on which the sector/dimension of this company is not the most common.	322 818	3 072 246	0	0	322 818	3 072 246
Filter only Micro-size companies	322 818	3 072 246	-50 515	-588 090	272 303	2 484 156
Removed companies with 0 employees (remunerated and non-remunerated)	272 303	2 484 156	-68 128	-477 358	204 175	2 006 798

The data cleaning process started by removing all the observations with negative or zero revenues and gross margins. By doing so, more than one million observations and almost 100 thousand companies were removed. This happens because most of these companies present zero revenues in this period. Companies with less than five observations were also removed. This minimum value of observations per company was set to avoid having excessively young companies when setting-up the representative firm. Almost 300 thousand companies and more than 600 thousand observations were lost during this step. After, financial and public administration companies were removed since they are not the target of this dissertation. To provide more reliable results, companies whose observations presented divergences of sectors were treated to present their most common sector, resulting in zero companies being lost. Up to this step, the micro-companies represented more than 80% of the observations. Finally, companies with no employees were removed from the sample, attempting to avoid having shell companies. Following these steps, the used dataset has 204 175 companies and 2 006 798 observations (accounting for 70% of the total revenues for micro-companies in the original dataset).

4.2 Summary statistics

Each company has an economic activity sector code given by its CAE (*Classificação Portuguesa das Actividades Económicas*-Industry sector code). In the final sample, it is possible to identify 14 different sectors: Agriculture, Mining and quarrying, Manufacturing, Electricity and gas, Water supply, Construction, Wholesale and retail trade, Transportation and storage, Accommodation and food services, Information and communication, Real estate, Professional scientific and technical, Education, Human health. Table 2 presents the number of companies, observations and the number of observations per company, both the average and the median values for each of these sectors.

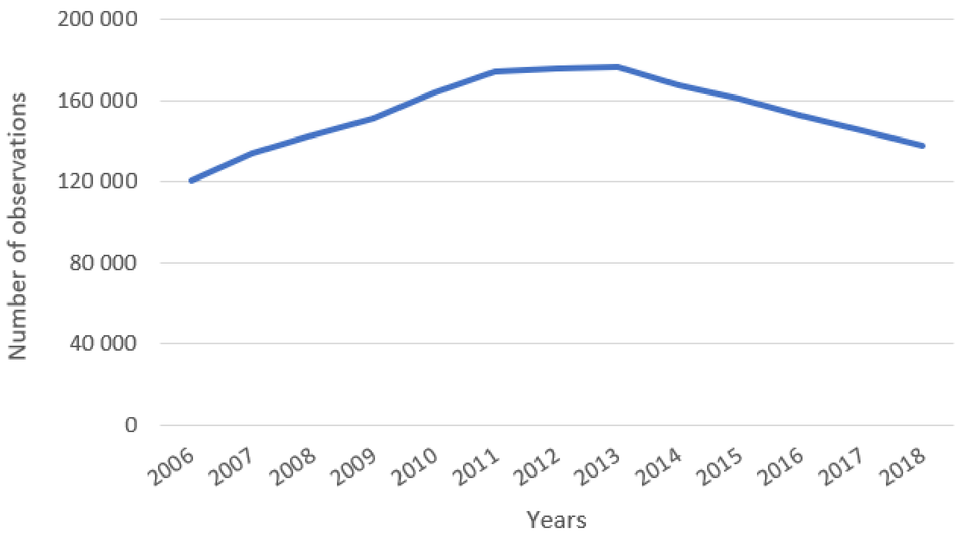
Table 2- Observations by sector

Sector of economic activity	Number of Companies	Number of observations	Number of observations per company (Average)	Percentage of total Gross Profit
Agriculture	6 522	62 455	9.6	3.4%
Mining and quarrying	375	3 660	9.8	0.3%
Manufacturing	20 421	201 877	9.9	9.4%
Electricity and gas	109	989	9.1	0.5%
Water supply	376	3 205	8.5	0.3%
Construction	22 233	206 954	9.3	12.7%
Wholesale and retail trade	65 968	657 309	10.0	27.5%
Transportation and storage	14 188	149 531	10.5	8.2%
Accommodation and food services	20 562	201 933	9.8	5.6%
Information and communication	5 054	46 258	9.2	3.0%
Real estate	7 191	66 276	9.2	6.2%
Professional scientific and technical	24 110	235 805	9.8	12.7%
Education	3 019	28 941	9.6	1.2%
Human health	14 047	141 605	10.1	9.0%

The number of companies (and subsequentially the number of observations) varies from sector to sector. In this sense, it is important to highlight the weight of "Wholesale and retail trade", near 32% of the sample. On the other hand, "Electricity and gas, Water supply and Mining and quarrying" are the smaller ones, representing 0.05%, 0.18%, and 0.18%, respectively. When measuring each sector's weights by their gross margin, the values are in line with the previous ones. The average number of observations per company is stable, fluctuating between 8.5 and 10.5 depending on the sector.

Figure 3 represents the number of observations per year. The year with the most observations is 2013, 177 062 (an increase of 42% when compared to 2006). In 2018, the number of observations is 137 536 (14% more companies than in 2006).

Figure 3- Observations per year



4.3 Model calibration

To understand the impact of the current pandemic on the Portuguese economy's different sectors, representative companies were created. These companies were created from the database described above. Each representative company resulted from the median values of each variable.

Corporate and dividend tax rates

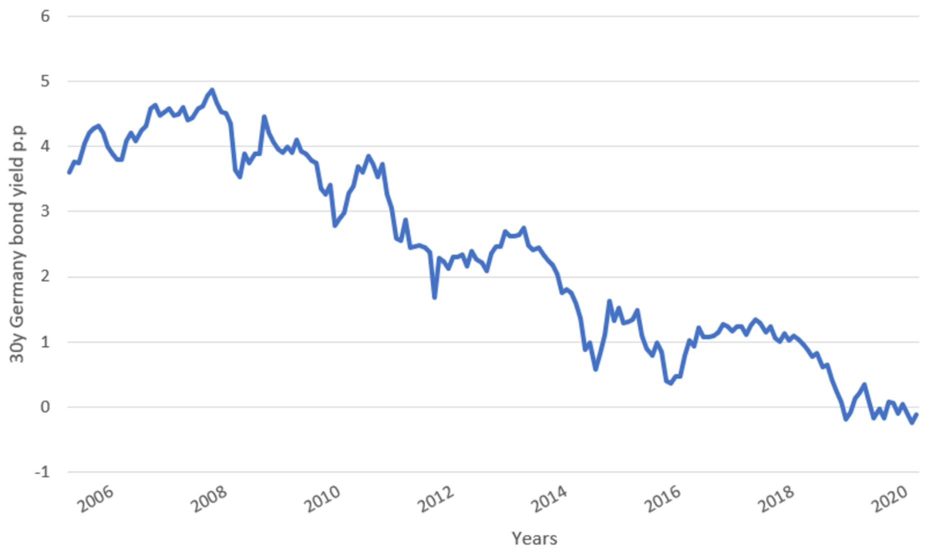
The model used in this dissertation considers a corporate tax rate applied to corporate profits and a dividend tax rate applied to dividend payments. The corporate tax rate was set at 21%, while the dividend tax rate was set at 28%. These values correspond to the corporate income tax rate (*IRC*) and the withholding dividends and capital gains rate in Portugal, respectively. [2] All in all, for each euro of profits, the shareholder receives 0.56€, in contrast to 0.65€ in EGZ original model.

[2] An effective corporate tax rate was computed per sector and is available in appendix B1. However, due to meaningful differences between sectors, a statutory tax rate was used to avoid having a distortion impacting the results.

Risk-Free interest rate

In this dissertation, the 30-year German bond yield was used as the risk-free interest rate. As presented in Figure 4, between 2006 and 2020, this rate is decreasing. Consequently, the risk-free rate is set by the last 12 months simple average, setting it at -0.02%.

Figure 4- Risk-free interest rate



Gross margin

The gross margin corresponds to total revenues minus costs of goods sold (cogs). Since the database was not updated, the values of 2018 were used as a proxy to the values of 2019. The gross margin of 2018 is computed per company and grouped by sector, resulting in the statistics presented in Table 3, after standardizing the values based on each company's revenues (gross margin to revenues - GMR). It is important to highlight that in seven of the 14 sectors, most companies do not report any costs of goods sold. This is the case of Electricity and gas, Transportation and storage, Information and communication, Real estate, Professional scientific and technical, Education, and Human health. For example, more than 70% of the observations reported zero cogs in the Electricity and gas sector. As a result, the gross profit over revenues ratio of these sectors representative firm is 1. Some sectors present wider asymmetries than others. For example, Education and Health show relatively low standard deviations (about 8%) and strongly concentrated quartiles. However, Agriculture and Water supply sectors have higher standard deviations (27%) and a significantly larger difference between quartiles. These asymmetries are important when considering the quality of the representative company. The median presents better each sector whenever asymmetries are less pronounced.

Table 3- Gross margin to revenues

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviation
Agriculture	1.00	0.00	0.48	0.72	0.93	0.69	26.90%
Mining and quarrying	1.00	0.07	0.85	0.97	1.00	0.91	16.59%
Manufacturing	1.00	0.02	0.47	0.63	0.80	0.64	21.71%
Electricity and gas	1.00	0.14	0.77	1.00	1.00	0.88	24.38%
Water supply	1.00	0.03	0.49	0.76	0.98	0.72	27.75%
Construction	1.00	0.06	0.54	0.69	0.88	0.70	19.23%
Wholsesale and retail trade	1.00	0.00	0.22	0.33	0.51	0.37	20.33%
Transportation and storage	1.00	0.04	1.00	1.00	1.00	1.00	10.37%
Accomodation and food services	1.00	0.00	0.39	0.51	0.63	0.52	21.33%
Information and communication	1.00	0.04	0.88	1.00	1.00	0.92	19.32%
Real estate	1.00	0.03	1.00	1.00	1.00	0.95	20.72%
Professional scientific and technical	1.00	0.02	1.00	1.00	1.00	0.97	12.78%
Education	1.00	0.08	0.99	1.00	1.00	0.98	8.80%
Human health	1.00	0.13	0.96	1.00	1.00	0.97	8.07%

Interest expenses

Table 4 shows 2018 interest expenses. To simplify the reading, these were standardized by revenues. Table 4 shows that the median company has relatively low interest payments as a percentage of revenues. These values are justified by a low level of debt in most of the companies and a low interest rates environment. It is important to highlight that while all the medians (as well as quantiles and averages) are relatively low, some sectors present much higher standard deviations than others. This is true for the Real estate, Construction, and Professional scientific and technical sectors, reflecting higher asymmetry within these sectors.

Table 4- Interests payments to revenues

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviation
Agriculture	6.57	0.00	0.00	0.0069	0.0301	0.0210	19.44%
Mining and quarrying	0.46	0.00	0.00	0.0057	0.0186	0.0126	5.08%
Manufacturing	8.21	0.00	0.00	0.0028	0.0101	0.0063	8.84%
Electricity and gas	0.27	0.00	0.00	0.0198	0.0689	0.0373	6.50%
Water supply	0.25	0.00	0.00	0.0040	0.0141	0.0100	3.50%
Construction	84.34	0.00	0.00	0.0021	0.0091	0.0060	59.31%
Wholsesale and retail trade	52.76	0.00	0.00	0.0025	0.0105	0.0067	20.74%
Transportation and storage	6.91	0.00	0.00	0.0019	0.0079	0.0049	7.35%
Accomodation and food services	13.70	0.00	0.00	0.0009	0.0055	0.0039	16.22%
Information and communication	8.79	0.00	0.00	0.0014	0.0073	0.0046	4.96%
Real estate	75.32	0.00	0.00	0.0049	0.0329	0.0288	122.35%
Professional scientific and technical	59.32	0.00	0.00	0.0017	0.0090	0.0059	57.54%
Education	1.82	0.00	0.00	0.0029	0.0113	0.0073	4.51%
Human health	3.21	0.00	0.00	0.0017	0.0077	0.0040	4.59%

Depreciations

Similar to EGZ, the depreciation to revenues ratio was computed per company using the observable values in the considered period. Table 5 presents the results obtained for the depreciation by sector. Some sectors present extremely high levels of standard deviation, such as Construction or Real estate. The standard deviations presented are not only the result of the heterogeneity within the sector but also the heterogeneity of the depreciation over revenues through time. However, analyzing the differences between quartiles, it is possible to conclude that outliers are driving this outcome.

Table 5- Depreciations to revenues

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviation
Agriculture	32.94	0.00	0.05	0.12	0.23	0.16	66.21%
Mining and quarrying	2.67	0.00	0.04	0.09	0.16	0.10	18.18%
Manufacturing	4.29	0.00	0.01	0.03	0.06	0.04	11.74%
Electricity and gas	1.32	0.00	0.04	0.11	0.30	0.17	21.60%
Water supply	1.64	0.00	0.02	0.05	0.10	0.07	13.92%
Construction	457.02	0.00	0.01	0.02	0.04	0.03	307.50%
Wholsesale and retail trade	84.59	0.00	0.00	0.01	0.03	0.02	36.85%
Transportation and storage	2.91	0.00	0.01	0.06	0.01	0.06	9.27%
Accomodation and food services	74.45	0.00	0.01	0.02	0.06	0.04	61.83%
Information and communication	29.15	0.00	0.01	0.03	0.07	0.04	50.94%
Real estate	70.68	0.00	0.01	0.05	0.14	0.09	172.43%
Professional scientific and technical	44.64	0.00	0.01	0.04	0.08	0.05	40.31%
Education	6.48	0.00	0.01	0.04	0.09	0.55	16.69%
Human health	2.52	0.00	0.03	0.06	0.11	0.07	8.34%

Capital Expenditures

Regarding the capex, by analyzing the authors' results in the original model (EGZ) it is possible to establish a relationship between depreciation and capital expenditures. From their results, we find that the capex was set equal to the depreciation plus 2 p.p., possibly to consider the inflation rate. Thus, inspired by the authors' result, in this dissertation, the same assumption was made.

Fixed operating costs

Regarding the fixed operating costs, the values used were provided by Félix, Moreira, and Silva (2021). In this article, the authors estimated the fixed costs of the Portuguese companies per sector. Following a similar approach to the one proposed by Gu, Hackbarth and Johnson (2018), the authors compute firms' fixed operating costs using Equation 24:

Equation 24

$$FixedCosts_{i,t} = \beta_i + \beta_{1,i}TotalOperatingCosts_{i,t-1} + \beta_{3,i}Revenues_{i,t-1},$$

whose coefficients were estimated using equation 25:

Equation 25

$$\begin{aligned} TotalOperatingCosts_{i,t} \\ &= \beta_i + \beta_{1,i}TotalOperatingCosts_{i,t-1} + \beta_{2,i}Revenuess_{i,t} \\ &+ \beta_{3,i}Revenues_{i,t-1} + e_{i,t} \end{aligned}$$

The authors reached the values of fixed operating costs to revenues presented in Table 7. Different sectors have relatively different fixed costs. The Transportation and storage sector has the lowest median fixed costs, representing 5% of total revenues. On the contrary, in the Accommodation and food services sector, the fixed costs represent 30% of revenues. In general, the calculated values show relatively high standard deviations, between 8% and 20%.

Table 6- Fixed operating costs to revenues

Sector of economic activity	Q1	Median	Q3	Average	Standard deviation
Agriculture	0.05	0.14	0.24	0.17	14%
Mining and quarrying	0.09	0.16	0.22	0.18	12%
Manufacturing	0.05	0.11	0.17	0.13	11%
Electricity and gas	0.13	0.22	0.38	0.25	15%
Water supply	0.06	0.13	0.34	0.21	18%
Construction	0.03	0.06	0.21	0.13	14%
Wholesale and retail trade	0.03	0.07	0.14	0.09	8%
Transportation and storage	0.03	0.05	0.13	0.10	12%
Accomodation and food services	0.20	0.29	0.42	0.31	16%
Information and communication	0.05	0.10	0.19	0.13	11%
Real estate	0.12	0.22	0.35	0.26	20%
Professional scientific and technical	0.04	0.13	0.27	0.17	16%
Education	0.13	0.19	0.26	0.20	10%
Human health	0.18	0.27	0.37	0.28	13%

Other operating costs

The other operating costs were calculated by the difference between the total operating costs to revenues[3] and the median fixed operating costs to revenues by sector, as summarized in Table 7. As shown in the table, most sectors are subject to a high-cost structure. Half of the sectors incur in other operating costs higher than 50% of total revenues. The Electricity and gas

[3] In the original data set, total operating costs are not provided. For this reason, they were computed by subtracting to the revenues the costs of goods sold and the operating income. In the Appendix B3, more information regarding the total operating costs is provided.

sector presents the lowest value, around 14% of revenues. The Transportation and storage sector presents a considerable value, about 82% of revenues. Both cases are expected given the business operation. However, given this difference, some sectors stand out with higher operating results. The Energy and Agriculture sectors stood out for presenting a relatively low-cost structure and a high gross margin.

Table 7- Other operating costs to revenues

Sector of economic activity	Other operating costs
Agriculture	0.27
Mining and quarrying	0.61
Manufacturing	0.44
Electricity and gas	0.14
Water supply	0.43
Construction	0.57
Wholsesale and retail trade	0.23
Transportation and storage	0.82
Accomodation and food services	0.21
Information and communication	0.73
Real estate	0.60
Professional scientific and technical	0.73
Education	0.73
Human health	0.48

Cash-holdings

As explained in sub-chapter 3.2, two levels of cash-holdings are considered in the model; the initial level, used by the companies to face the current crisis and the minimum level, necessary to start the second stage of the model. Similarly to each company's remaining variables, it was assigned the median level of the cash-holdings observed. Table 8 presents the results obtained aggregating by sectors. The initial cash-holding level was set at the median value of each sector. The minimum level was set at the first quartile of each sector. The Human health sector presents the highest median and minimum cash-holdings to revenues levels, 0.2 and 0.07 respectively. Accommodation and food services is the sector with the lowest levels of median and minimum cash-holdings, 0.04 and 0.02 respectively.

Table 8- Cash-holding to revenues

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviation
Agriculture	6.10	0.00	0.06	0.17	0.46	0.31	78.71%
Mining and quarrying	3.56	0.00	0.03	0.07	0.17	0.13	41.95%
Manufacturing	4.69	0.00	0.03	0.08	0.18	0.12	69.69%
Electricity and gas	2.46	0.00	0.05	0.12	0.47	0.26	39.13%
Water supply	3.86	0.00	0.03	0.08	0.16	0.11	32.56%
Construction	2.12	0.00	0.04	0.09	0.20	0.14	14.87%
Wholsesale and retail trade	5.00	0.00	0.03	0.06	0.14	0.10	19.63%
Transportation and storage	9.83	0.00	0.06	0.18	0.46	0.29	83.99%
Accomodation and food services	1.70	0.00	0.02	0.04	0.12	0.08	42.41%
Information and communication	8.12	0.00	0.05	0.12	0.27	0.18	48.89%
Real estate	2.50	0.00	0.06	0.15	0.37	0.26	113.47%
Professional scientific and technical	4.70	0.00	0.05	0.12	0.27	0.18	47.04%
Education	1.67	0.00	0.04	0.09	0.21	0.14	62.44%
Human health	5.75	0.00	0.07	0.20	0.49	0.33	10.16%

Covid-19 impact on the initial gross margin, *k*

To estimate the Covid-19 impact, the "Fast and Exceptional Enterprise Survey (COVID-IREE)"- provided by Banco de Portugal Microdata Research Laboratory- BPLIM was used. More than 8 thousand companies answered this survey. In the beginning, between April and May, this survey was answered every week. Then, up to July, every 15 days. In this questionnaire, the companies were asked to assess the pandemic's impact on the company's business. As Table 9 shows, ten possible answers were allowed, five of them representing a decrease in the business activities and the remaining an improvement. As the answers were given in the form of a range, it was decided to use each range's median values. However, it was decided to use 75%, as Table 9 shows for the upper range.

Table 9- Used values from the questionnaire

Answer	Value used
(negatively) less than 10%	-5%
(negatively) between 10%-25%	-18%
(negatively) between 26%-50%	-38%
(negatively) between 51%-75%	-63%
(negatively) more than 75%	-75%
(positively) less than 10%	5%
(positively) between 10%-25%	18%
(positively) between 26%-50%	38%
(positively) between 51%-75%	63%
(positively) more than 75%	75%

Using the results provided in July and aggregating by sector, using simple average, the impact of the Covid-19 crisis results in Table 10 values for *k*. The agriculture sector was excluded from this survey. To continue analyzing the impact in this sector, a simple average between all the sectors was used as proxy for *k*. This is an important assumption, impacting the final results, and for this reason, it cannot be compared with the remaining sectors.

Table 10- Impact in gross margin

Sector of economic activity	k
Agriculture	-26.0%
Mining and quarrying	-19.4%
Manufacturing	-25.2%
Electricity and gas	-6.2%
Water supply	-22.2%
Construction	-27.1%
Wholsesale and retail trade	-22.0%
Transportation and storage	-31.4%
Accomodation and food services	-47.8%
Information and communication	-25.2%
Real estate	-33.6%
Professional scientific and technical	-28.6%
Education	-21.1%
Human health	-27.9%

Expected growth rates

As explained in Chapter 3, two gross margin growth rates are used. First, a long-term growth rate is set to be used after the pandemic crisis is over. Second, a short-term growth rate is used during the pandemic stage. When computing each sector growth rate, multiple methodologies were considered[4]. However, none of these methodologies were used due to the model constraints, either by having negative rates or due to unreasonable values. For this reason, and following a common practice for perpetuities, a standard growth rate of 2% was set. This can be thought of as the expected inflation in the economy. As explained in sub-chapter 3.3, the annual short-term growth rates during the pandemic stage corresponds to the shock size *k* divided by the number of years the shock lasts, two years. The results are presented in Table 11.

[4] In appendix B4 further methodologies on the computation of the long-term growth rates are provided.

Table 11- Short term expected annual growth rates

Sector of economic activity	Short-term growth rate
Agriculture	13.0%
Mining and quarrying	9.7%
Manufacturing	12.6%
Electricity and gas	3.1%
Water supply	11.1%
Construction	13.6%
Wholsesale and retail trade	11.0%
Transportation and storage	15.7%
Accomodation and food services	23.9%
Information and communication	12.6%
Real estate	16.8%
Professional scientific and technical	14.3%
Education	10.5%
Human health	14.0%

Volatility

To compute the volatilities, the observable log-changes in the gross margin were grouped by sector. The volatility was then estimated as the standard deviation of these log-changes. The results are presented in Table 12. This is equivalent to treating all firms within the same sector equally. The presented standard deviations are relatively high, however, in some sectors, the figures are considerably higher. Therefore, it is important to highlight sectors such as Construction, Real Estate and Agriculture. These three sectors present standard deviations of over 60%. Being sectors with a high degree of seasonality, these results are in line with expectations. On the contrary, the Human health sector presents the smallest standard deviation, in line with the expectations.

Table 12- Standard deviation of the state variable

Sector of economic activity	Standard deviation
Agriculture	64.46%
Mining and quarrying	51.77%
Manufacturing	41.01%
Electricity and gas	52.20%
Water supply	47.03%
Construction	64.67%
Wholsesale and retail trade	42.67%
Transportation and storage	26.05%
Accomodation and food services	44.08%
Information and communication	48.89%
Real estate	66.40%
Professional scientific and technical	39.54%
Education	37.89%
Human health	22.89%

Market price of risk

EGZ uses the CAPM to set the market price of risk. However, working with private companies, the market price of risk cannot be computed in the same way. As an alternative, the market price of risk was computed by solving Equation 7 in respect to *m,* resulting in,

Equation 26

$$m_i = \frac{\frac{\mu_{i,p} * ((1-\tau)GMR_i + \tau\overline{DSR}_i)}{\overline{CSR}_i} - r_f}{\sigma_i},$$

Where $\mu_{i,p}$ is the growth rate of the representative company and the GMR_i, $\overline{DSR}_i$, $\overline{CSR}_i$ correspond to the representative company median values of the, gross margin to revenues, depreciations to revenues and capex to revenues respectively.

Table 13 provides the respective market price of risk to each sector. Analyzing the table, it is evident the certain degree of homogeneity between sectors since most of them present a market price of risk between 0.23 and 0.41. However, the Agriculture sector registers the lowest value, 0.10, while the Human health sector presents the highest value, 0.51.

Table 13- Market price of risk

Sector of economic activity	Market price of risk
Agriculture	0.10
Mining and quarrying	0.22
Manufacturing	0.32
Electricity and gas	0.19
Water supply	0.26
Construction	0.28
Wholsesale and retail trade	0.22
Transportation and storage	0.49
Accomodation and food services	0.28
Information and communication	0.35
Real estate	0.25
Professional scientific and technical	0.45
Education	0.43
Human health	0.51

4.4 Testing the model assumptions

State variable distribution

In the EGZ model, assuming that the state variable follows a geometric Brownian motion implies that the log changes of this variable follow a normal distribution. Due to the high standard deviations presented before it was decided to study this assumption. To do so a

Kolmogorov-Smirnov test, an alternative to the Shapiro-Wilk test due to the high number of observations. As Table 14 presents, the p-value is 0 for all the sectors. In this case, the null is rejected, and it is not possible to state that the variable is normally distributed. Consequently, it is necessary to recognize a limitation of the model in representing of the state variable. Analyzing the high kurtosis values, between 11 and 90, confirms the presence of extreme cases. Similarly, since most sectors present a negative skew, the companies have greater chances of registering negative growths.

Table 14- Skewness, Kurtosis and Kolmogorov test of the state variable

Sector of economic activity	Skewness	Kurtosis	p-value
Agriculture	0.08	11.99	0.00
Mining and quarrying	-0.68	16.76	0.00
Manufacturing	-0.67	44.28	0.00
Electricity and gas	-0.54	17.12	0.00
Water supply	0.36	23.25	0.00
Construction	-0.36	16.42	0.00
Wholsesale and retail trade	-0.73	60.60	0.00
Transportation and storage	-1.11	74.72	0.00
Accomodation and food services	-0.51	24.64	0.00
Information and communication	-0.20	19.02	0.00
Real estate	-0.04	18.31	0.00
Professional scientific and technical	-0.31	24.77	0.00
Education	-0.08	22.63	0.00
Human health	0.72	90.79	0.00

Constant depreciation to revenues ratio

As explained in the model chapter, the EGZ original model assumes a constant depreciation to revenues ratio. However, in the original article, this hypothesis is not tested. Therefore, it was tested for the companies under analyzes if this hypothesis holds. Consequently, Equation 27 was estimated company by company. To do so, each company must have at least ten observations, reducing our analysis to 105 656 companies.

Equation 27

$$DepToSales_{i,t} = \beta_{0,i} + \beta_{1,i} * DepToSales_{i,t-1} + e_{i,t}$$

Then, it was tested if β_1 is statistically different from 1. As represented in Table 15, using an alpha of 5%, in 64% of the companies it was not possible to reject the nulls hypotheses. However, some sectors present high standard deviations of the error term, indicating high deviations from the mean. Thus, it is not possible to say if the hypothesis considered is valid or not.

Table 15- Depreciations to revenues, assumption test

Sector of economic activity	Percentage of companies with β_1 equal to 1	Standard deviation of the error term
Agriculture	62.97%	32.34%
Mining and quarrying	64.64%	40.08%
Manufacturing	66.36%	6.06%
Electricity and gas	59.14%	12.29%
Water supply	52.45%	6.29%
Construction	57.94%	11.85%
Wholsesale and retail trade	73.62%	3.81%
Transportation and storage	73.38%	7.98%
Accomodation and food services	68.00%	7.94%
Information and communication	64.95%	26.06%
Real estate	55.86%	44.22%
Professional scientific and technical	68.77%	7.84%
Education	59.77%	25.25%
Human health	66.26%	10.67%

5 Results

In this chapter, it is first analyzed the credit risk of Portuguese micro firms. This is done by looking at sector-representative firms calibrated using median values of the input variables. Second, the results on the pandemic shock's impact on the probabilities of default and distances to default of these companies are presented. Third, a sensitivity analysis is performed to explore the impact of changing some of the most important inputs, notably, the gross margin to revenues ratio, the initial cash holdings, the volatility of gross profits log-changes, the recovery duration, and by relaxing the minimum cash requirement at the end of the period. A binomial tree with daily steps was used for these exercises, and 10 000 paths were simulated[5].

5.1 Portuguese micro companies credit risk assessment

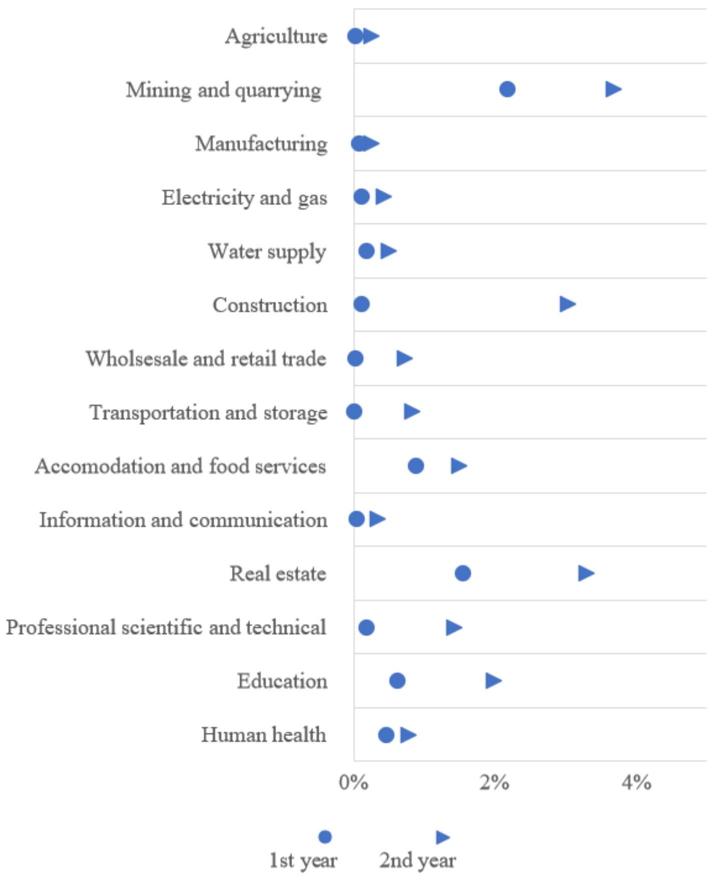
Analyzing the model results without considering the effects of Covid-19, it is possible to identify the different sectors that present different credit risk levels. Figure 5 displays the cumulative probabilities of default of the representative companies in one year and two years. From these sectors, it is possible to highlight four sectors that present higher probabilities of bankruptcy at one year. The Mining and Quarrying sector presents the highest bankruptcy rate, 2.17%, followed by the Real estate sector with the second highest probability of bankruptcy, 1.55%. The Accommodation and food services and Education sector then stand out with rates

[5] When tested with 200000 simulations, the results obtained were similar.

of 0.89%, and 0.62%, respectively. The remaining sectors have a probability of bankruptcy between 0.03% and 0.47%.

Expanding the analysis to the second year, it is possible to identify that from the previous four sectors, now three stand out from the rest by showing a higher probability of bankruptcy. The Mining and quarrying sector presents the highest probability of default (PD) in two years, 3.69%. The Real estate and Education sector present the second and fourth highest rate at two years, 3.30%, and 1.98%, respectively. The Construction sector is also highlighted with the third highest PD, 3.04%. The Agriculture sector stands out for the lowest two-year rate of 0.25%. Analyzing the PDs between the first and the second year, notice that the companies' rank order does not hold. One of the factors explaining this result is the difference between the volatilities and the increase in volatilities, given by the standard deviation and the period's square root.

Figure 5- 1- and 2-years Probabilities of default (cumulative probabilities)



It is well known from the literature that the PD increases non-linearly with firms fundamentals. Alternatively, it is possible to present these results through the distance to default (DD). The DD measures the number of standard deviations that the state variable is from the default barrier. The traditional distance to default formula cannot capture the effect of default before maturity. For this reason, the inverse of the Normal distribution applied to the respective probability of default is used. This is often called the DD because in the Merton model, the two things are the same. So, the DD computed in this way can be seen as the Merton model equivalent DD. Figure 6 illustrates the distance to default (DD) in one and two years for each sector. As expected, the interpretation made for the probabilities of default is identical for the DD's. Thus, the sectors with the highest PD's are now the sectors with the lowest DD's.

Figure 6- 1- and 2-years Distance to default

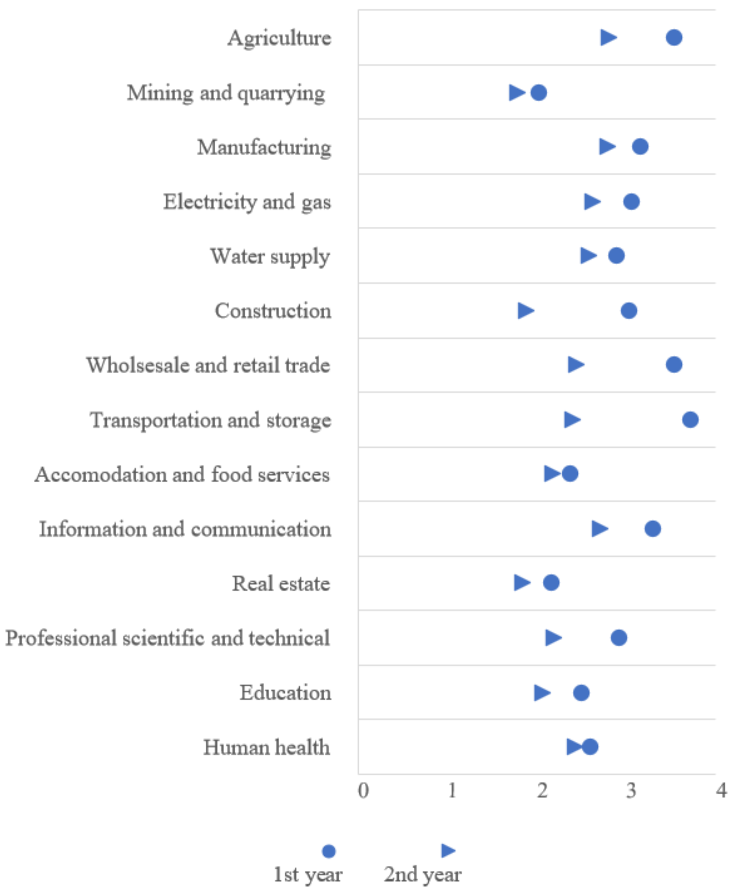


Table 16 presents the threshold level of the gross margin and the distance to default, calculated using the inverse of the normal and the formula suggested by Merton. As explained before, the gross margin threshold is not constant during the periods, increasing with time. Standardizing the values by the gross margin, all sectors present an initial gross margin equal to 1. As Table 16 shows, using the DD formula proposed by Merton leads to consistently higher values when comparing with the formula used. This result was expected considering the effect of the first-time passage model and as explained in previous chapters. On average, the difference between the two is 1.13. However, they are strongly correlated at 77.34%. As expected, the relationship between threshold level and DD is negative, with a correlation of -66.80%. Thus, a lower threshold levels lead to a higher DD, highlighting the Agriculture sector with an x_D of 0.02 and a DD of 3.54 in the first year. With the increase in the gross margin threshold in the second year, the distance to default decreases.

Table 16- Gross margin threshold levels per sector

Sector of economict activity	Time 0	1st year			2st year		
	x	x_D	DD^{Merton}	DD	x_D	DD^{Merton}	DD
Agriculture	1	0.02	5.60	3.54	0.03	3.09	2.81
Mining and quarrying	1	0.09	4.19	2.02	0.11	2.34	1.79
Manufacturing	1	0.08	4.66	3.16	0.10	3.27	2.79
Electricity and gas	1	0.03	4.71	3.06	0.04	3.80	2.63
Water supply	1	0.08	4.91	2.89	0.09	2.91	2.58
Construction	1	0.02	4.58	3.04	0.06	2.48	1.87
Wholsesale and retail trade	1	0.04	5.04	3.54	0.10	3.17	2.44
Transportation and storage	1	0.07	5.24	3.42	0.18	4.00	2.40
Accomodation and food services	1	0.14	4.01	2.37	0.16	2.29	2.17
Information and communication	1	0.02	4.04	3.29	0.11	3.52	2.71
Real estate	1	0.1	3.00	2.16	0.16	2.33	1.84
Professional scientific and technical	1	0.09	4.59	2.91	0.16	2.61	2.19
Education	1	0.2	3.80	2.50	0.25	2.75	2.06
Human health	1	0.13	4.45	2.60	0.19	3.45	2.42

5.2 The impact of Covid-19 on the representative company

Figure 7 shows the impact on the cumulative probabilities of default of each sector in the first and second year of the Covid-19 shock.

Focusing on the first year, three sectors stand out. First, the Accommodation and food services sector, which suffers an increase of 5.24 percentage points (p.p.) and presents a PD of 6.13% in the first year. Then, the Real estate sector presents a difference of 3.35 p.p., increasing the PD to 4.90%. Finally, the Mining and quarrying sector has the third-largest impact with a PD of 6.2% after the shock, an increase of 3.16 p.p. Three factors help to justify the increase in these sectors. First, all sectors suffer from considerable shocks. Second, these are sectors with relatively low values of initial cash. Third, these are sectors with low operating margins.

As explained in the chapter about the model, it was defined that the shock lasts two years. For this reason, it is important to extend the analysis of the probability of bankruptcy until the second year. In a similar way to the 1-year analysis, it is possible to highlight three sectors. First, the Accommodation and food service sector presents a cumulative PD of 14.20% with the pandemic, versus the 1.49% without, a 12.71 p.p. increase. Second, the Real estate sector with a PD of 14.76% compared to 3.30% without a pandemic, a 11.46 p.p. increase. Third, the

Professional scientific and technical sector with a PD of 11.09%, (9.66 p.p. increase). It is also important to highlight the Electricity and gas, with the lowest impact, notably 0.14 p.p. However, these results are predictable since they are exposed to some of the greatest and smallest shocks, respectively.

Figure 7- Impact on Probabilities of default in 1- and 2- years in p.p.

Figure 8 demonstrates the pandemic's impact on DD in each sector in the first and second years. Regarding the first year DD, all sectors present a DD between 2 and 3.5. After the shock, the values fall and are mostly between 2 and 2.5. On average, DD goes down 0.66 after the shock. The Transportation and storage sector, the Construction sector and the Information and communication sector exhibit the largest impacts on DD, 1.61, 1.05, and 1.04, respectively.

Similar to the PD's, the average variation of DD is higher when analyzed at two years. On average, the DD drops by 0.81, instead of the previous 0.66, compared with and without the current pandemic. However, some sectors have a more significant impact on 1-year DD than at year 2. This result's interpretation is identical to the one used for the analysis between the first and second year without pandemic.

As it is possible to understand from the Figure 8, the Information and communication and Accommodation and food services sectors are the most impacted sectors, presenting a shock in the second year DD of 1.16 and 1.10, respectively. Similarly, the Transportation and storage, Human health, and Manufacturing sectors suffer significant changes varying between 1.09, 1.08, and 1.03, respectively.

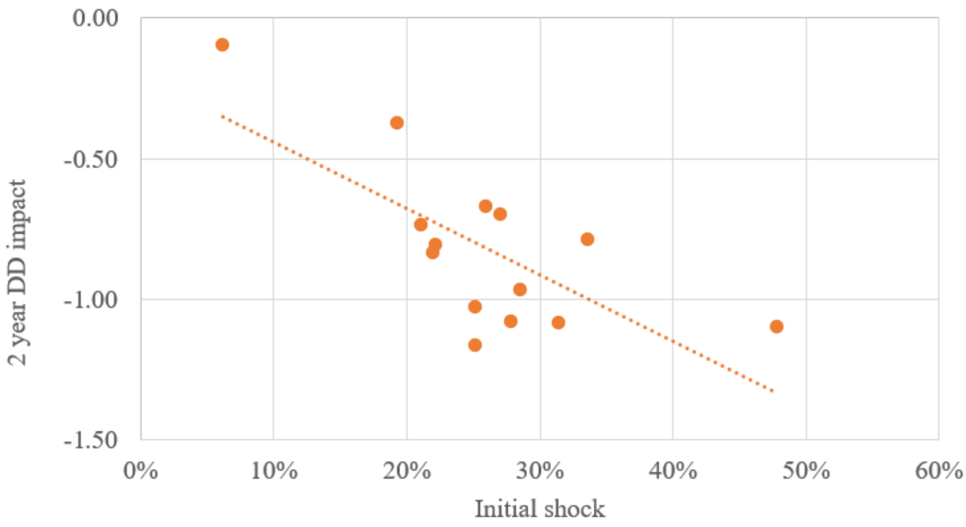
The Electricity and gas sector with the smallest change in DD, 0.09, should also be highlighted. Notice that this is the sector reporting the smallest shock. The remaining sectors suffer shocks varying between -0.38 and -0.96.

It is important to remember that the Agricultural sector was exposed to the average shock due to the lack of information. However, after this shock, DD in two years decreased by 0.67.

Figure 8- Impact on Distance to default in 1- and 2-years

Figure 9 shows the relationship between the size of the initial shock and the impact on two-year DD. As expected, companies facing a larger shock also present a larger impact on the distance to default; this means the two-year DD becomes smaller with larger shocks. This said, the correlation between these two variables is -71.53%. Note that knowing this relationship can be used to calibrate similar shocks in other models and compared. As the graphic also shows, some sectors facing similar shocks present lower impacts on DD. This phenomenon can be explained by the other variables used, such as gross margin to revenues ratio, initial cash-holding value, or volatility.

Figure 9- Relation between the initial shock and the impact in the DD



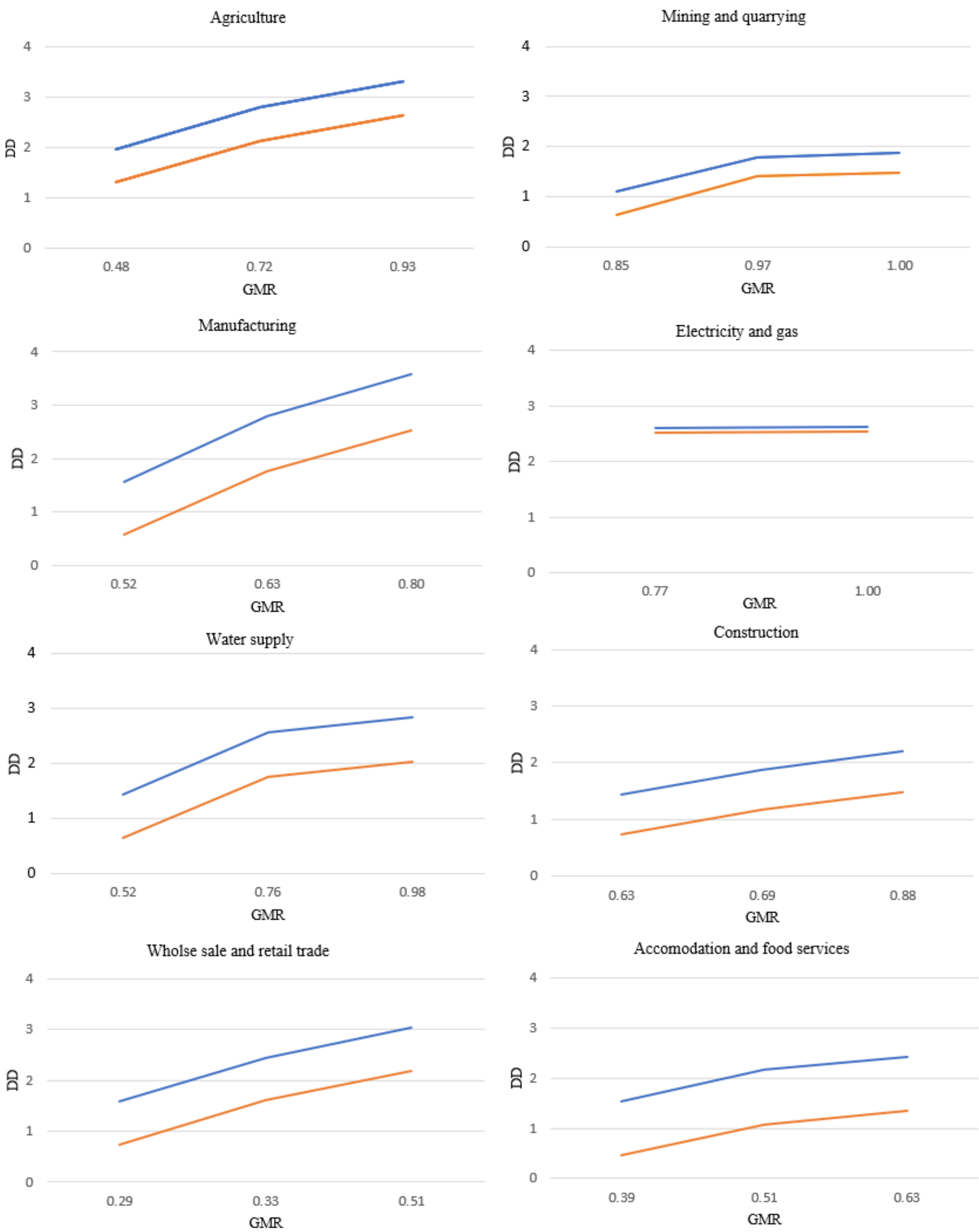
5.3 Sensitivity analysis

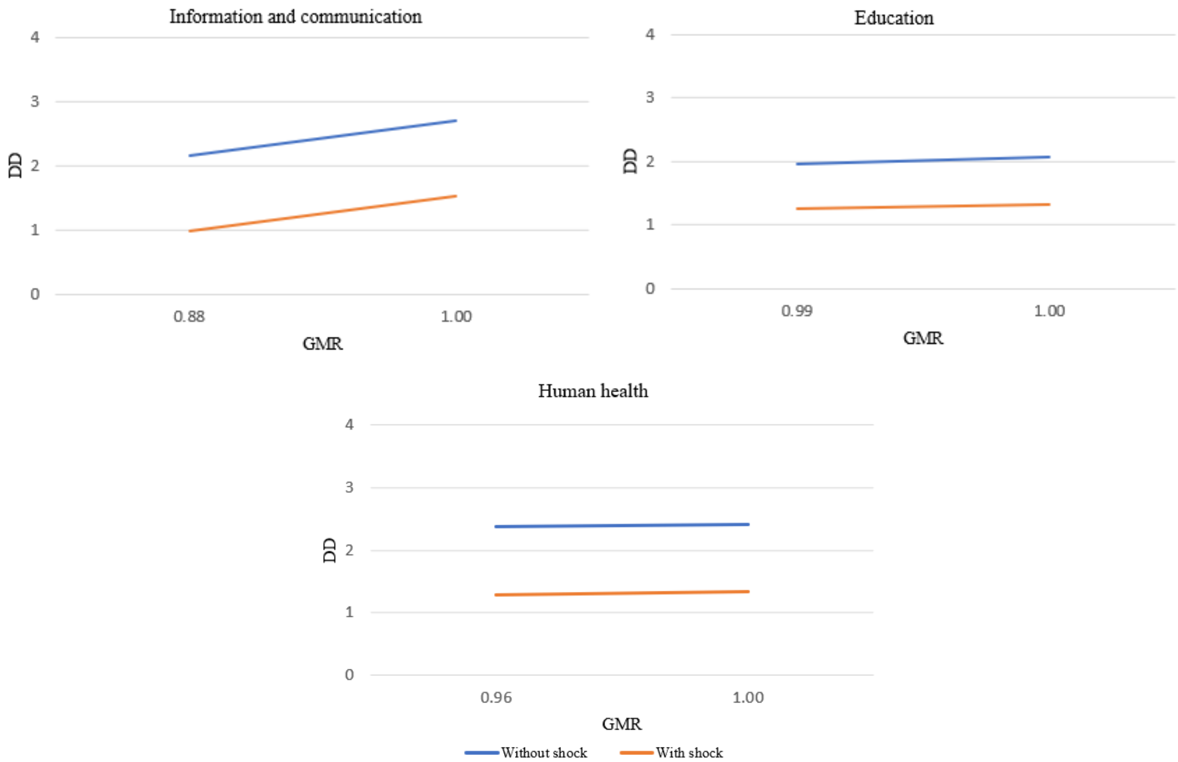
The objective of this sub-chapter is to understand the variables impact on the survival of companies. For this purpose, five variables are studied. The gross margin to revenues ratio, the initial cash value, the gross profit volatility, the shock duration, and relaxing the minimum cash requirement. For the first two variables, the effects on DD is tested using the first and third quartiles. In the case of standard deviation, the effect of increasing and decreasing it by 10 p.p. is tested.

Gross margin to revenues ratio

To better comprehend the impact of the gross margin to revenues ratio on the distance to default, a similar exercise to the previous one was done using the first and third quartile values for each sector instead of the me. The Transportation and storage, Real estate, and Professional scientific and technical sectors were excluded from this analysis because they present the same value in all quartiles. Figure 10 presents the results of this exercise for the various sectors. As expected, a lower GMR leads to a decrease in DD and a higher GMR to an increase in DD. However, a more curious result emerges from this exercise. As one can see, the variation of the DD with and without shock is reasonably constant in all sectors in the different quartiles. This result is of great importance because it allows us to expand our analysis and understand the current pandemic's effect on a larger group of companies. Based on this and keeping everything else constant, a company in the same sector but with a lower GMR suffers the same change in DD as one with a higher GMR.

Figure 10- Distance to default in 2 years with Q1, Median and Q3 of the GMR

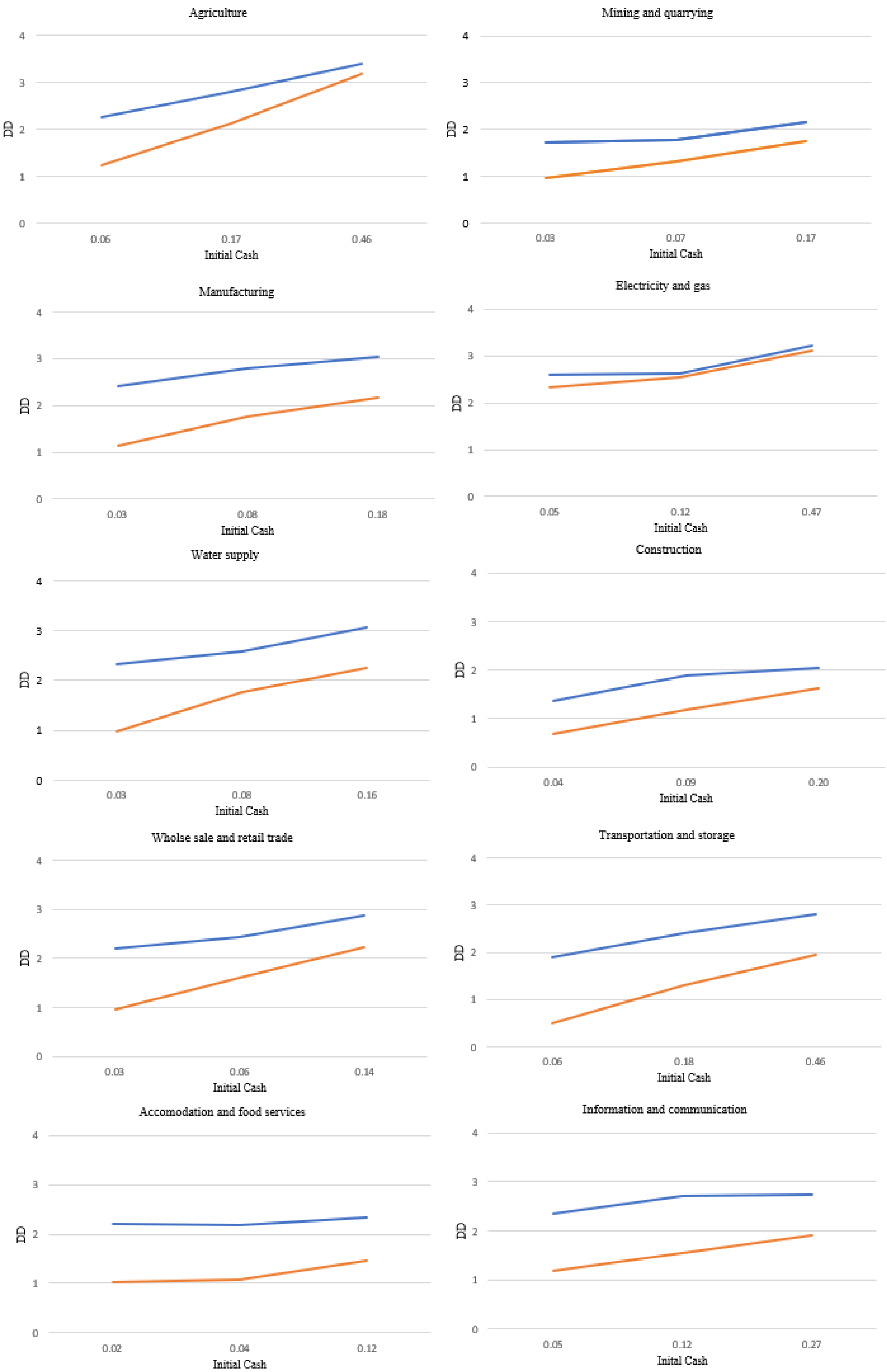


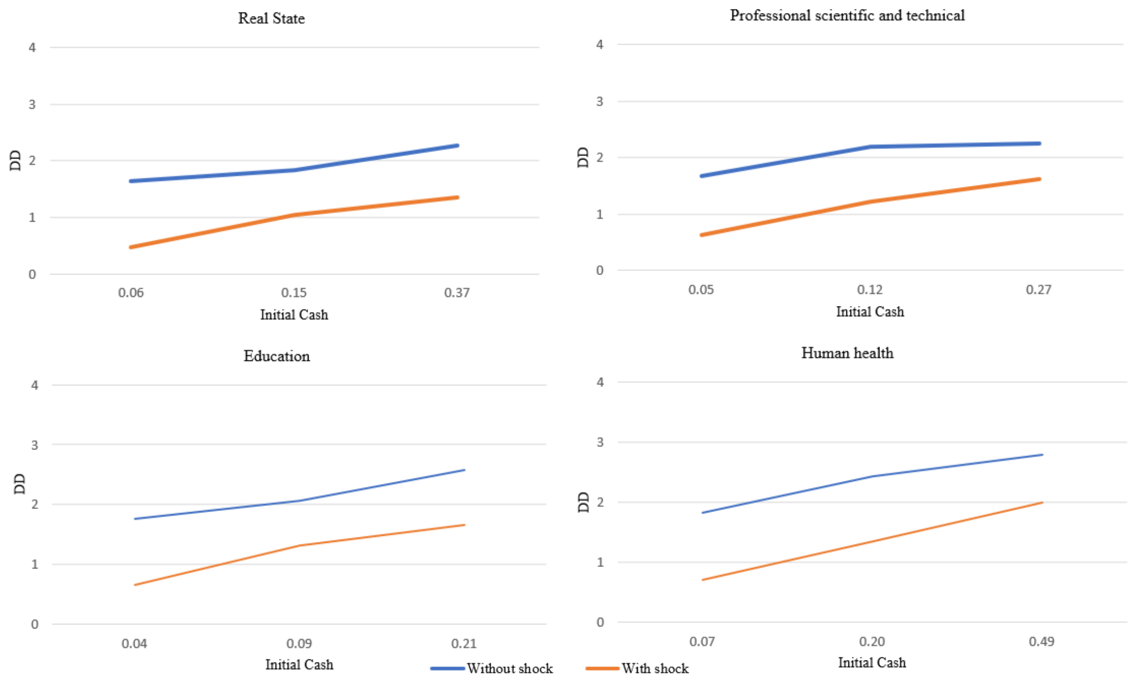


Initial cash holding

Figure 11 shows the distance to default in two years per sector with and without the Covid-19 shock with different initial levels of cash. As expected, the increase in the cash holding level leads to an increase in the distance to default. Interpreting the Covid-19 crisis as a liquidity shock, by giving companies a higher level of cash helps them facing this crisis with a higher survival rate. However, in contrast to the GMR, the impact of Covid-19 is smaller with higher levels of cash. On average, when using the third quartile of cash-holding as the initial cash level, the pandemic shock impacts the distance to default by less than 0.35 when comparing to the first quartile levels. Some sectors have interquartile intervals much higher than others, leading to higher differences. In these circumstances, this result becomes very important to perceive the effect of the pandemic since the shock does not linearly impact the companies as the previous one.

Figure 11- Distance to default in 2 years with Q1, Median, Q3 of the initial cash

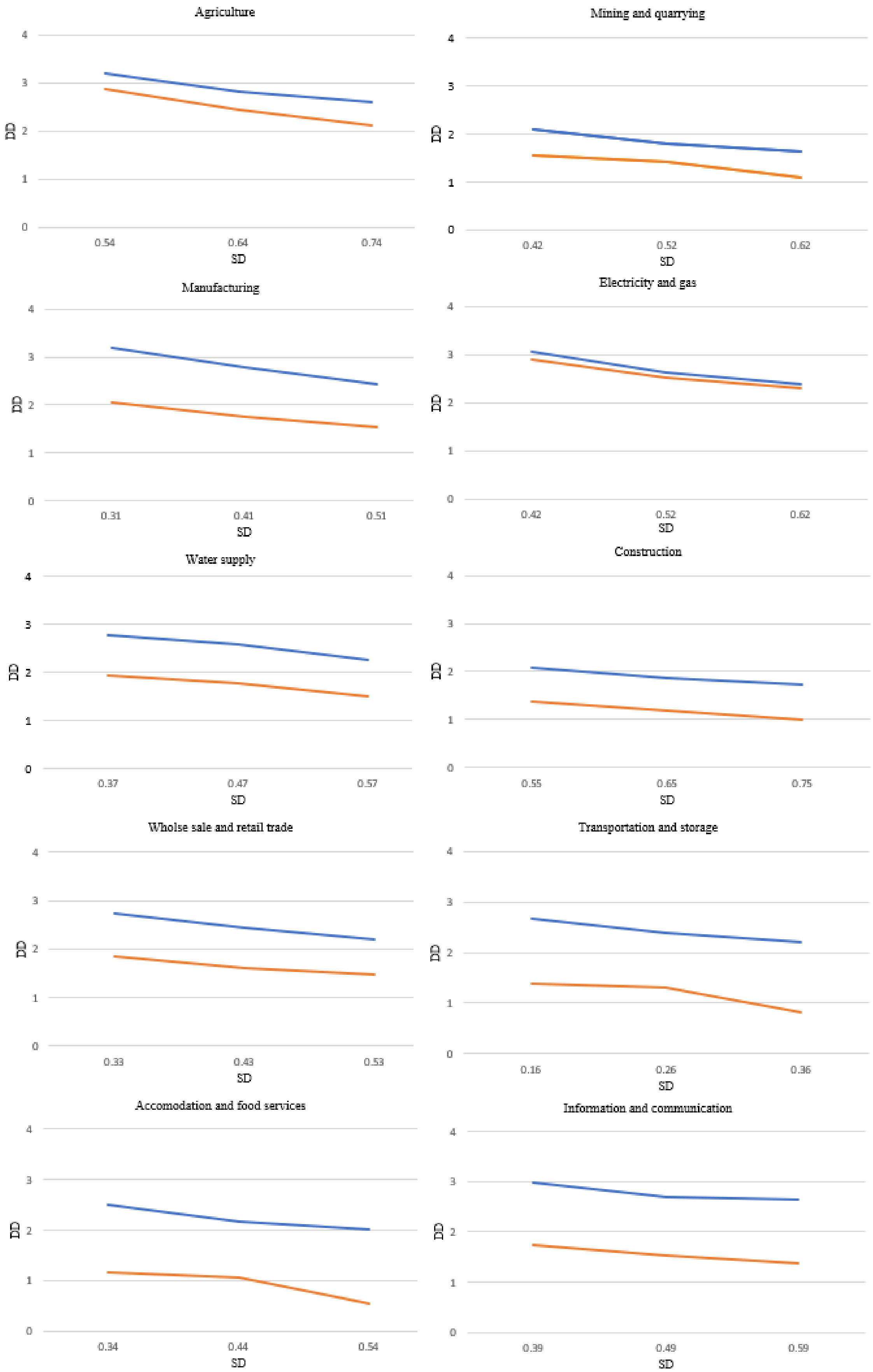


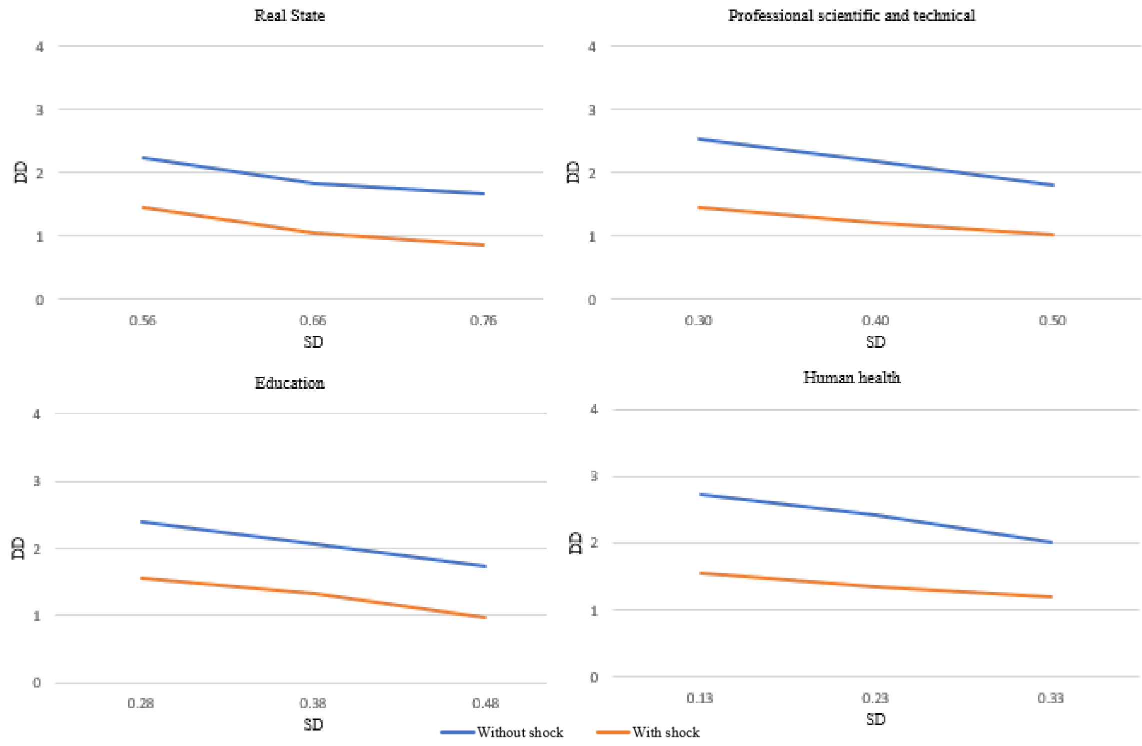


Volatility

Finally, the impact of varying the standard deviation (SD) in 10 p.p. in the DD was tested with and without pandemic. Figure 12 illustrates this test and shows that an increase of 10 p.p. on the standard deviation leads to a decrease in DD, and the symmetric with a similar decrease. Knowing this, it is possible to state that with an increase in the volatility the probability of default also increases. However, it is not possible to identify a clear effect caused by the shock on DD's differences. In some sectors, a lower standard deviation translates into a lower difference of DD's, with and without shock (example of the Agriculture, Mining and quarrying, and Real estate sectors). However, other sectors present a greater difference (example of the Electricity and gas, Manufacturing and Human health). On average, with a higher volatility, the difference in the DD's is about 0.8, while with a lower volatility is 0.9. In this way, it is possible to say that on average, with a lower volatility the effect caused by Covid-19 in the DD at two years is higher.

Figure 12- Distance to default in 2 years with different standard deviation



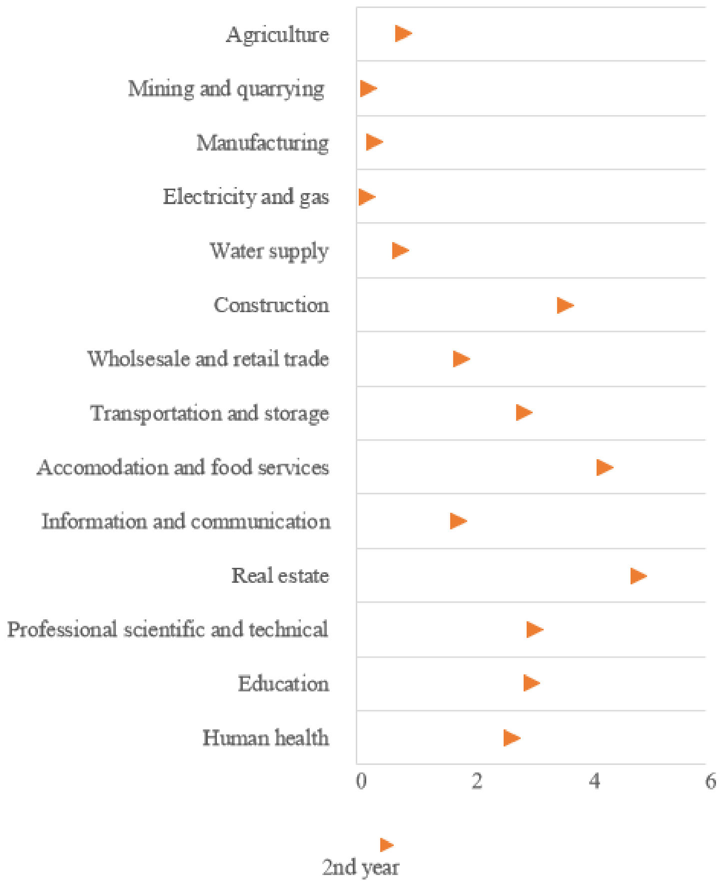


A 3-year shock

With the most recent information on the number of Covid-19 cases in Portugal and the confinement measures, it is possible to admit that a two-year shock is an optimistic view of the duration of the current crisis. Thus, the hypothesis that the economy will take three years to recover from this shock was explored.

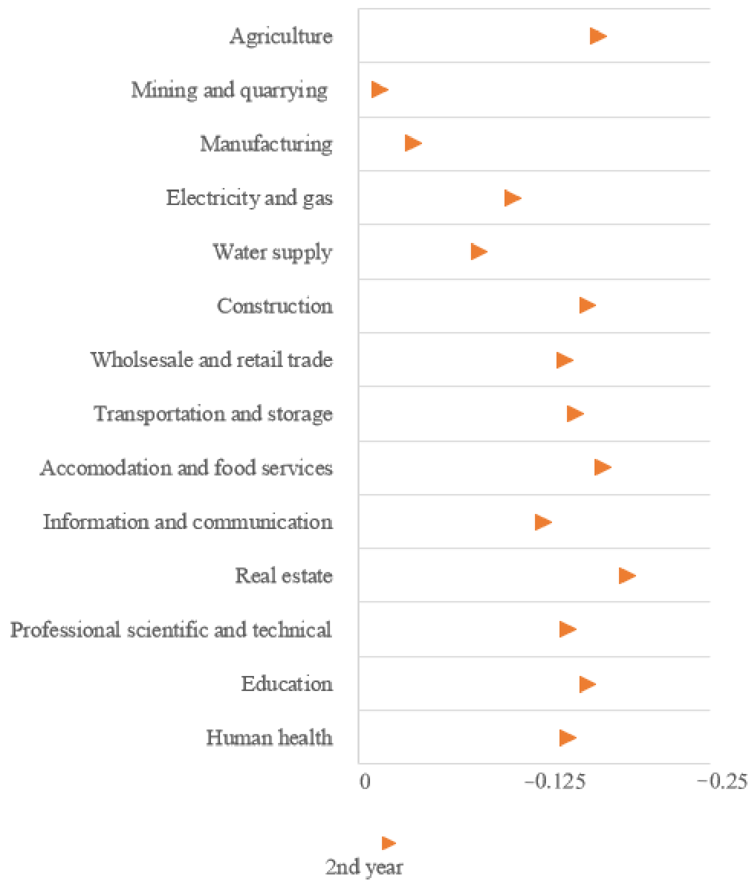
As expected, a more prolonged shock implies higher probabilities of bankruptcy, as shown in Figure 13. In this case, and if the shock lasts three years, the average PD at one year is expected to increase by 0.65pp relative to the impact of a two-year crisis. This effect is amplified when the two-year PD is analyzed since on average the 2 years cumulative PD increases by 2.16 p.p. The Real estate, Accommodation, and food services, and Construction sectors suffer the most from an increase in the shock duration, increasing by 4.84 p.p., 4.27 p.p., and 3.59 p.p. the two years cumulative PD, respectively.

Figure 13- Impact in the 2-years cumulative PD with the increase in the shock duration, in p.p.



Similarly, Figure 14 shows the difference between the DD's. In this case, if the shock lasts three years, the DD at one year decreases by 0.09 on average. In the second year, the shock in the DD is stronger, showing an average decrease of 0.13.

Figure 14- Impact in 1- and 2-years DD's with the increase in the shock duration



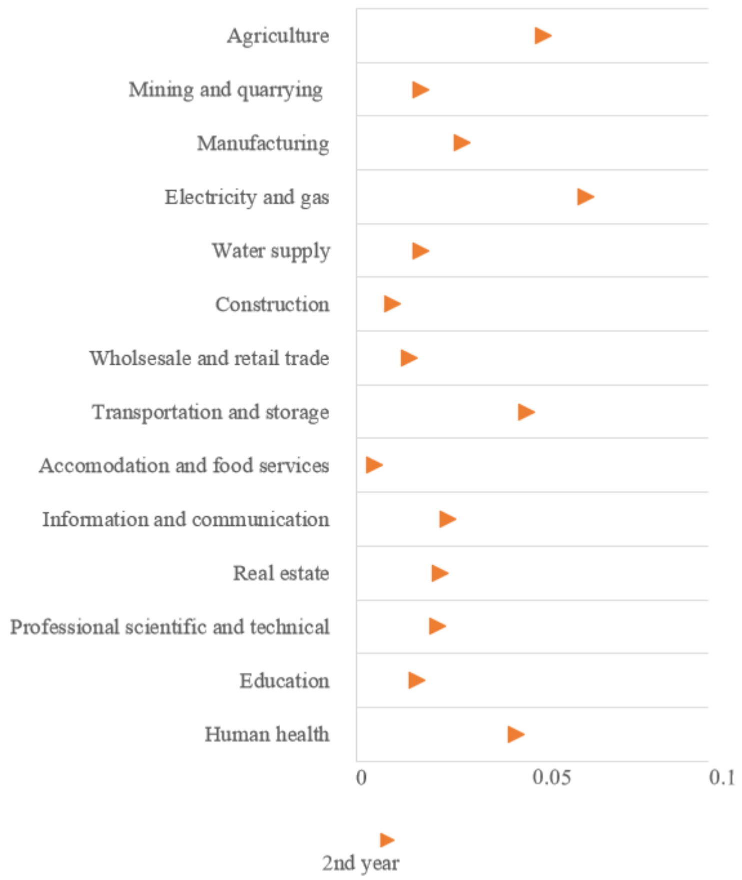
Removing the minimum cash-holding requirement

During the exercise, the companies were required to maintain a minimum cash-holding level at the end of the second year. This would give them liquidity to face possible shocks during the second phase (in perpetuity). However, the shareholders may not require a minimum cash-holding level to maintain activity. For this reason, this assumption was relaxed, and its impact on the cumulative probability of default in the second year was analyzed. Figure 15 shows the impact on the cumulative probability of default in the second year with shock. As one would expect, relaxing this hypothesis decreases the PD. On average, the two-year PD decreases by 0.31 p.p., but it is evident that this impact is stronger in cases of higher minimum levels. In this sense, the Human health sector, which had the highest minimum cash holdings, suffers the greatest impact with a decrease of 0.64 p.p. On the contrary, the Accommodation and food services sector suffers the smallest impact -0.11 p.p., as it had the lowest cash minimum requirements.

Figure 15- Impact in 2-years PD's with the removing the minimum cash-holding, in p.p.

Figure 16 extends the analysis through the impact on the distance to default. On average, the DD increases by 0.03. The Electricity and gas, and Agriculture sectors stand out from the others. However, this highlight is essentially due to how the distance to default is calculated, suffering greater changes with lower PDs. Otherwise, the greatest impact sectors remain the same, highlighting Human health and Transportation and storage. Suffering the least impact is the Accommodation and food services sector, which arises from the low minimum cash value.

Figure 16- Impact in 2-years DD's with the removing the minimum cash-holding



6 Conclusion

This dissertation aimed to understand the impact of Covid-19 on Portuguese micro-enterprises by analyzing its impact on the probability of default. This exercise was carried using representative firms from 14 sectors of activity and applying a shock corresponding to their respective expectations. The probability of default was computed using a modified version of the Eisdorfer, Goyal, & Zhdanov (2019) model.

As suggested by the existing literature, it is possible to identify a significant degree of heterogeneity in the impact of the pandemic by sector of economic activity. In Portugal, the Accommodation and food services sector stands out by its shock size, losing 47.8% of business volume. This sector is followed by the Real estate and the Transportation and storage sectors that lost, respectively, 33.6% and 31.4% of their business volume. The Electricity and gas sector stands out for the opposite reasons as the sector that reports the smallest losses in its activity, 6.2%. After estimating the shocks by sector, it was necessary to calculate the probabilities of default, taking the shock into account and understand its impact. The results obtained suggest that the Accommodation and food services sector is the most negatively impacted. The two-year cumulative probability of default increases by 12.71 p.p., from 1.49% to 14.2%. On the other hand, Electricity and gas is the sector that registers the smallest impact, 0.14 p.p., moving from 0.43% to 0.57%. On average, the probabilities of default increase by 6.46 p.p. under the current crisis up to the second year. However, while significant, the size of the impact is not the only determinant variable for the impact of the current pandemic. The Agricultural sector exposed to an average shock suffered a much smaller impact than the average, namely 1.38 p.p.

The different impact of the pandemic was further analyzed through a sensitivity analysis. By doing this, it is possible to identify the impact of other key variables. In this respect, measuring the impact through the distance to default, it was possible to identify that a higher level of initial cash allows to diminish the impact of the pandemic, highlighting the importance of access to liquidity during this crisis. An important result emerges from this sensitivity analysis. Except for different levels of cash, it is noteworthy that the impact of the current pandemic is relatively constant independently of the initial gross margin to revenues ratio. This result allows us to take the estimated change in the distance to default as a reference value that can be used in the context of stress test exercises. It is therefore a key result from this dissertation.

Nevertheless, it is important to recognize the limitations of this exercise. The model used relies on a set of assumptions that may not be suitable for the reality of micro-enterprises. Namely, and as shown, it is not possible to demonstrate that the state variable follows a normal distribution and the growth rate chosen. As a suggestion for future research, two topics are worth mentioning: expand the analysis to the remaining dimensions and incorporate the impact of the measures taken by the government to support Portuguese companies.

References

Altman, E.I., Marco, G., & Varetto, F. (1994). Corporate Distress Diagnosis. In *Journal of banking & finance* (Vol. 18, Issue 3, pp. 505–529). http://linkinghub.elsevier.com/retrieve/pii/0378426694900078

Altman, Edward I., & Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking and Finance*, *21*(11–12), 1721–1742. https://doi.org/10.1016/S0378-4266(97)00036-8

Altman, Edward I. (1968). FINANCIAL RATIOS, DISCRIMINANT ANALYSIS AND THE PREDICTION OF CORPORATE BANKRUPTCY. *The Journal of Finance*, *23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, Edward I, & Sabato, G. (2007). Modelling Credit Risk for SMEs: Evidence from the U.S. Market. *Abacus*, *43*(3), 332–357. https://doi.org/10.1111/j.1467-6281.2007.00234.x

Back, B., Laitinen, T., Sere, K., & van Wezel, M. (1996). Choosing bankruptcy predictors using discriminant analysis, logit analysis, and genetic algorithms. *Turku Centre for Computer Science Technical Report*, *40*(2), 1–18.

Black, F., & Cox, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *The Journal of Finance*, *31*(2), 351–367. https://doi.org/10.1111/j.1540-6261.1976.tb01891.x

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654. http://www.jstor.org/stable/1831029

Brockman, P., & Turtle, H. J. (2003). A barrier option framework for corporate security valuation. *Journal of Financial Economics*, *67*(3), 511–529. https://doi.org/https://doi.org/10.1016/S0304-405X(02)00260-X

Carletti, E., Oliviero, T., Pagano, M., Pelizzon, L., & Subrahmanyam, M. G. (2020). The Covid-19 shock and equity shortfall: Firm-level evidence from Italy. *CPER PRESS*, *25*, 23–54.

Douglas Smith, L., & Lawrence, E. C. (1995). Forecasting losses on a liquidating long-term loan portfolio. *Journal of Banking & Finance*, *19*(6), 959–985. https://doi.org/https://doi.org/10.1016/0378-4266(94)00065-B

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, *83*(3), 635–665. https://doi.org/https://doi.org/10.1016/j.jfineco.2005.10.011

Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *The Journal of Financial and Quantitative Analysis*, *7*(2), 1477–1493. http://www.jstor.org/stable/2329929

Eisdorfer, A., Goyal, A., & Zhdanov, A. (2019). Equity Misvaluation and Default Options. *Journal of Finance*, *74*(2), 845–898. https://doi.org/10.1111/jofi.12748

European commission. (2003). *Commission recommendation concerning the definition of micro, small and medium-sized enterprises*.

Félix, S., Moreira, P., & Nuno, S. (2021). *On the measurement of Portuguese firms' fixed operating costs*. 29–41.

Goldstein, R., Ju, N., & Leland, H. (2001). An EBIT-Based Model of Dynamic Capital Structure. *The Journal of Business*, *74*(4), 483–512. http://www.jstor.org/stable/10.1086/322893

Gourinchas, P.-O., Kalemli-Özcan, Ṣebnem, Penciakova, V., & Sander, N. (2020). *COVID-19 and SME Failures* (Working Paper Series, Issue 27877). https://doi.org/10.3386/w27877

Gu, L., Hackbarth, D., & Johnson, T. (2018). Inflexibility and stock returns. *The Review of Financial Studies*, *31*(1), 278–321. https://doi.org/10.1093/rfs/hhx092

Hyytinen, A., & Pajarinen, M. (2008). Opacity of young businesses: Evidence from rating disagreements. *Journal of Banking & Finance*, *32*(7), 1234–1241. https://doi.org/https://doi.org/10.1016/j.jbankfin.2007.10.006

LELAND, H. E. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *The Journal of Finance*, *49*(4), 1213–1252. https://doi.org/10.1111/j.1540-6261.1994.tb02452.x

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance*, *1*(3), 249–276. https://doi.org/https://doi.org/10.1016/0378-4266(77)90022-X

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.

The Journal of Finance, *29*(2), 449. https://doi.org/10.2307/2978814

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131. http://www.jstor.org/stable/2490395

Reisz, A. S., & Perlich, C. (2007). A market-based framework for bankruptcy prediction. *Journal of Financial Stability*, *3*(2), 85–131. https://doi.org/10.1016/j.jfs.2007.02.001

Rikkers, F., & Thibeault, A. E. (2009). A Structural form Default Prediction Model for SMEs, Evidence from the Dutch Market. *Multinational Finance Journal*, *13*(3/4), 229–264. https://doi.org/10.17578/13-3/4-4

Sarkar, S., & Zapatero, F. (2003). The Trade-off Model with Mean Reverting Earnings: Theory and Empirical Tests. *Economic Journal*, *113*, 834–860. https://doi.org/10.1111/1468-0297.t01-1-00156

Welburn, J., & Strong, A. (2020). Estimating the Impact of COVID-19 on Corporate Default Risk. *Estimating the Impact of COVID-19 on Corporate Default Risk*, *June*. https://doi.org/10.7249/wra173-2

Yang, Z. R., Platt, M. B., & Platt, H. D. (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research*, *44*(2), 67–74. https://doi.org/10.1016/S0148-2963(97)00242-7

Zhou, C. (2001). The Term Structure of Credit Spreads with Jump Risk. *Journal of Banking & Finance*, *25*, 2015–2040. https://doi.org/10.1016/S0378-4266(00)00168-0

Appendix

A.1 Value of equity at time T

Equation A1, Equation 10 constraints

$$\begin{cases} A = 0, \\ Bx^{*\beta_2} + \left[\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{x^*}{r-\mu_Q} - (1-\tau) * \frac{Int_i + Foc_i}{r} \\ = Cx^{*\beta_1} + Bx^{*\beta_2} + (1+\eta)\left\{\left[\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{x^*}{r-\mu_Q} - (1-\tau) * \frac{Int_i + Foc_i}{r}\right\}, \\ \beta_2 Bx^{*\beta_2 - 1} + \left[\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{1}{r-\mu_Q} \\ = \beta_1 Cx^{*\beta_1} + Bx^{*\beta_2} + (1+\eta)\left\{\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right\}\frac{1}{r-\mu_Q}, \\ Cx_d^{\beta_1} + Bx_d^{\beta_2} + (1+\eta)\left\{\left[\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right] * \frac{x_d}{r-\mu_Q} - (1-\tau) * \frac{Int_i + Foc_i}{r}\right\} = 0, \\ \beta_1 Cx_d^{\beta_1} + Bx_d^{\beta_2} + (1+\eta)\left\{\frac{\tau\overline{DSR} - \overline{CSR}}{GMR_{i,t}} + (1-\tau)\right\}\frac{x_d}{r-\mu_Q} = 0 \end{cases}$$

A.2 Cash holding dynamics

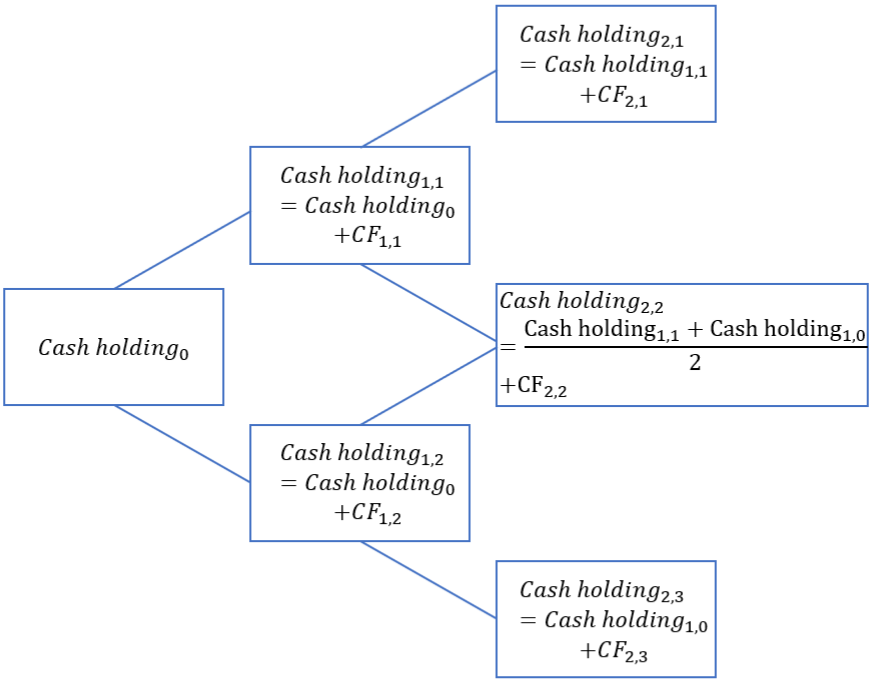
As explain previously, due to the cash holding tree setup, the value of $cash\ holding_{up\ and\ down\ or\ down\ and\ up}$ could be equal to two different values:

$$cash\ holding_{up\ and\ down} = cash\ holding_{up} + CF, \text{ or}$$

$$cash\ holding_{down\ and\ up} = cash\ holding_{down} + CF$$

To simplify the cash holding tree the value of this $cash\ holding$ was set as a simple average of this two cases as showed in figure A2.1

Figure A2.1, cash holding tree dynamic



A.3 Value of equity at time 0

Equation A3.1, value of equity at time 0

$$E_{i,0}(x_{i,0}) = (1-\tau^{Div}) * [(1-\tau^{Corp})(A_{Solv} - A_{Int} - A_{Foc} - A_{Ooc}) + \tau^{Corp} A_{Dep} - A_{Capex},$$

Where,

$$A_{Solv} = A_t - p_d * A_d$$

$$p_d = \left(\frac{A_t}{A_d}\right)^{-a}, \text{ with } a = \frac{1}{\sigma^2}\left[\left(\mu_Q - \frac{\sigma^2}{2}\right) + \sqrt{\left(\mu_Q - \frac{\sigma^2}{2}\right)^2 + 2r\sigma^2}\right]$$

$$A_{Int} = \frac{Int_t}{r}(1-p_d)$$

$$A_{Foc} = \frac{FOC_t}{r}(1-p_d)$$

$$A_{Ooc} = \frac{\overline{OocSalesRatio}}{\overline{GrossMarginRatio}} * A_{Solv}$$

$$A_{Dep} = \frac{\overline{DepSalesRatio}}{\overline{GrossMarginRatio}} * A_{Solv}$$

$$A_{Capex} = \frac{\overline{CapexSalesRatio}}{\overline{GrossMarginRatio}} * A_{Solv}$$

A.4 Default gross margin threshold

1. Re-writing the elements of the equity value, we have that:

$$A_{Solv} = \frac{x_i}{k} - \left(\frac{x_i}{x_d}\right)^{-a} * \frac{x_d}{k}$$

$$A_{Int} = \frac{Int_t}{r}(1 - p_d)$$

$$A_{FOC} = \frac{FOC_t}{r}(1 - p_d)$$

$$A_{Ooc} = \frac{\overline{OOR}}{\overline{GMR}} * \left(\frac{x_i}{k} - \left(\frac{x_i}{x_d}\right)^{-a} * \frac{x_d}{k}\right)$$

$$A_{Dep} = \frac{\overline{DSR}}{\overline{GMR}} * \left(\frac{x_i}{k} - \left(\frac{x_i}{x_d}\right)^{-a} * \frac{x_d}{k}\right)$$

$$A_{CAPEX} = \frac{\overline{CSR}}{\overline{GMR}} * \left(\frac{x_i}{k} - \left(\frac{x_i}{x_d}\right)^{-a} * \frac{x_d}{k}\right)$$

2. Computing the derivatives in order to x_i, we have:

$$\frac{\partial A_{solv}}{\partial x} = \frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}$$

$$\frac{\partial A_{int}}{\partial x} = \frac{Int_t}{r} * \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}}$$

$$\frac{\partial A_{FOC}}{\partial x} = \frac{FOC_t}{r} * \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}}$$

$$\frac{\partial A_{Ooc}}{\partial x} = \frac{\overline{OOR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right]$$

$$\frac{\partial A_{Dep}}{\partial x} = \frac{\overline{DSR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right]$$

$$\frac{\partial A_{CAPEX}}{\partial x} = \frac{\overline{CSR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right]$$

3. Write the derivate of equity in order to x_i and make it equal to zero

$$\frac{\partial E}{\partial x} = \left(1 - \mathcal{T}^{Div}\right) * \left[\begin{array}{c} \left(1 - \mathcal{T}^{Corp}\right) * \left(\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k} - \frac{Int}{r} * \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} - \frac{FOC_t}{r} * \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} - \frac{\overline{OOR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right]\right) \\ + \mathcal{T}^{Corp} * \frac{\overline{DSR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right] - \frac{\overline{CSR}}{\overline{GMR}} * \left[\frac{1}{k} + \frac{a\, x_i{}^{-a-1}}{x_d{}^{-a}} * \frac{x_d}{k}\right] \end{array}\right] = 0$$

4. When equalizing x_i to x_d and isolating x_d, we have that:

$$x_d = \frac{a * k * (1 - t_{corp}) * (Int + Foc)/r}{1 - t_{corp} + a * (1 - t_{corp}) + (1 - a) * [t_{corp} * \overline{DSR} - \overline{CSR} - (1 - t_{corp}) * \overline{OOR}]/\overline{GMR}}$$

B.1 Effective tax rate

Table B1.1-Effective tax rate per sector

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviations
Agriculture	95.21%	-12.92%	7.84%	17.39%	22.32%	15.75%	10.41%
Mining and quarrying	77.82%	-7.06%	77.58%	17.04%	24.55%	16.28%	12.16%
Manufacturing	99.72%	-320.40%	9.28%	18.50%	25.00%	17.49%	12.40%
Electricity and gas	44.39%	-150.80%	14.56%	21.59%	25.04%	19.18%	19.00%
Water supply	90.63%	0.00%	8.91%	18.42%	25.01%	17.31%	12.42%
Construction	97.20%	-372.06%	10.53%	18.50%	25.00%	17.87%	12.52%
Wholsesale and retail trade	100.00%	-194.14%	10.18%	19.25%	26.07%	18.43%	15.58%
Transportation and storage	92.53%	-3.56%	9.26%	17.13%	23.29%	16.38%	11.08%
Accomodation and food services	97.32%	-146.31%	5.10%	13.80%	21.01%	13.21%	16.09%
Information and communication	98.34%	-39.58%	12.33%	21.91%	29.13%	21.44%	14.89%
Real estate	96.46%	-278.18%	8.53%	18.58%	25.17%	17.47%	13.28%
Professional scientific and technical	100.00%	-343.70%	9.06%	19.82%	27.64%	19.36%	15.36%
Education	94.65%	-28.58%	5.95%	14.39%	23.40%	14.78%	13.50%
Human health	99.94%	-2.51%	15.88%	24.37%	30.39%	23.03%	14.00%

B.2 Capital investment by sectors

As mention in chapter 3.3 an alternative way was used to compute the capital investment by each company. By following equation B2.1 and aggregating per company using the median Table B2.2 was created.

Equation B2.1

$$CapexToSales_{i,t} = \frac{TangibleAssets_{i,t} - TangibleAssets_{i,t-1} + Depreciations_{i,t}}{Total\ Sales_{i,t}}$$

Table B2.2-Capital investment per sector

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviations
Agriculture	56.11	-1.78	0.00	0.04	0.14	0.09	101.33%
Mining and quarrying	0.80	-0.45	0.00	0.01	0.06	0.03	9.15%
Manufacturing	6.62	-0.73	0.00	0.00	0.02	0.01	9.84%
Electricity and gas	0.54	-8.11	0.00	0.00	0.04	0.02	8.35%
Water supply	3.37	-0.07	0.00	0.01	0.04	0.02	20.64%
Construction	100.49	-4.31	0.00	0.00	0.01	0.01	79.35%
Wholsesale and retail trade	21.95	-1.76	0.00	0.00	0.01	0.00	11.95%
Transportation and storage	2.23	-4.42	0.00	0.00	0.01	0.01	5.39%
Accomodation and food services	157.51	-2.65	0.00	0.00	0.01	0.01	112.30%
Information and communication	33.27	-0.17	0.00	0.00	0.02	0.01	64.74%
Real estate	88.87	-4.69	0.00	0.00	0.02	0.02	180.80%
Professional scientific and technical	11.20	-2.30	0.00	0.01	0.02	0.01	14.53%
Education	5.17	-3.97	0.00	0.00	0.02	0.01	13.40%
Human health	3.06	-2.55	0.00	0.01	0.03	0.02	6.98%

B.3 Total operating costs

Table B3.1-Total operating costs to revenues

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviations
Agriculture	42.17	-82.31	0.20	0.41	0.67	0.43	149.88%
Mining and quarrying	3.74	0.01	0.62	0.77	0.90	0.77	34.94%
Manufacturing	12.49	-12.00	0.41	0.55	0.72	0.57	35.54%
Electricity and gas	11.94	-0.01	0.19	0.36	0.61	0.42	117.18%
Water supply	3.20	-0.01	0.35	0.56	0.79	0.57	34.38%
Construction	126.41	-80.94	0.49	0.63	0.78	0.64	561.58%
Wholsesale and retail trade	17.11	-225.80	0.19	0.30	0.48	0.34	94.96%
Transportation and storage	61.14	-1.17	0.79	0.87	0.93	0.86	55.70%
Accomodation and food services	23.25	-150.08	0.40	0.50	0.63	0.52	115.35%
Information and communication	40.25	-27.50	0.64	0.83	0.94	0.79	87.45%
Real estate	169.68	-275.82	0.52	0.82	0.96	0.76	528.41%
Professional scientific and technical	108.85	-98.52	0.72	0.86	0.93	0.82	113.82%
Education	8.79	-9.70	0.84	0.92	1.03	0.96	54.89%
Human health	9.63	-1.44	0.60	0.75	0.86	0.73	26.81%

B.4 Growth rates and volatility of the state variable

Due to model limitations, the growth rates under the physical measure, P, cannot be negative. Before using the methodology presented in chapter 3.3 two different methodologies were used. First, by computing a growth rate of the gross margins per company (Table B4.1 presents a summary statistic by sector).

Table B4.1-Long-term growth rate and volatility of the state variable, method 1

Sector of economic activity	Max	Min	Q1	Median	Q3	Average	Standard deviations
Agriculture	3.36	-2.15	-0.06	0.05	0.18	0.06	29.97%
Mining and quarrying	1.75	-2.48	-0.23	-0.05	0.05	-0.09	36.14%
Manufacturing	4.92	-4.11	-0.11	0.00	0.10	-0.01	30.85%
Electricity and gas	1.63	-1.11	-0.01	0.03	0.14	0.06	27.07%
Water supply	2.23	-2.23	-0.12	0.07	0.24	0.06	41.43%
Construction	4.20	-4.15	-0.18	-0.01	0.11	-0.03	34.95%
Wholsesale and retail trade	7.52	-4.15	-0.13	-0.01	0.10	-0.02	32.29%
Transportation and storage	1.77	-3.89	-0.08	0.00	0.07	-0.01	25.47%
Accomodation and food services	2.86	-3.84	-0.12	0.00	0.10	-0.01	30.13%
Information and communication	2.98	-3.38	-0.11	0.03	0.20	0.04	35.49%
Real estate	3.28	-3.72	-0.12	0.03	0.20	0.04	40.66%
Professional scientific and technical	2.59	-4.06	-0.09	0.01	0.14	0.02	30.22%
Education	1.92	-1.96	-0.07	0.03	0.19	0.05	30.10%
Human health	3.85	-2.60	-0.05	0.02	0.12	0.04	24.68%

Second, by applying the regression B4.2 to each company and use the betas as the growth rate and the volatility of the error term as volatility. Table B4.3 shows a summary statistic by sector.

Equation B4.2

$$Log\left(\frac{GrossMargin_{i,t}}{GrossMargin_{i,t-1}}\right) = \beta_i + e_{i,t}$$

Table B4.3-Regression B4.2 summary statistics per sector

Sector of economic activity	Growth rate	Standard deviation	Skewness	Kurtosis	p-value
Agriculture	0.05	86.91%	0.11	11.74	0.00
Mining and quarrying	-0.09	67.03%	-0.36	15.74	0.00
Manufacturing	-0.01	59.25%	-0.32	42.77	0.00
Electricity and gas	0.05	67.32%	-0.46	13.23	0.00
Water supply	0.05	71.58%	0.19	18.96	0.00
Construction	-0.03	88.65%	-0.12	16.14	0.00
Wholsesale and retail trade	-0.02	62.39%	-0.42	64.01	0.00
Transportation and storage	-0.01	44.53%	-0.71	76.17	0.00
Accomodation and food services	-0.02	63.14%	-0.35	22.07	0.00
Information and communication	0.03	70.18%	-0.01	15.98	0.00
Real estate	0.03	94.75%	0.06	16.48	0.00
Professional scientific and technical	0.01	59.77%	-0.12	22.83	0.00
Education	0.03	58.40%	0.02	21.00	0.00
Human health	0.03	41.68%	0.59	101.76	0.00