UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

MODELOS DE DURAÇÃO APLICADOS À SOBREVIVÊNCIA DAS EMPRESAS PAULISTAS ENTRE 2003 E 2007.

André Luís Pavão

Orientador: Prof. Dr. Hélio Nogueira da Cruz

São Paulo

2013

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Reinaldo Guerreiro
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Joaquim José Martins Guilhoto
Chefe do Departamento de Economia

Prof. Dr. Pedro Garcia Duarte
Coordenador do Programa de Pós-Graduação em Economia

ANDRÉ LUÍS PAVÃO

MODELOS DE DURAÇÃO APLICADOS À SOBREVIVÊNCIA DAS EMPRESAS PAULISTAS ENTRE 2003 E 2007.

Tese apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo IPE/USP.

Orientador: Hélio Nogueira da Cruz

São Paulo

2013

FICHA CATALOGRÁFICA
Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Pavão, André Luís
Modelos de duração aplicados à sobrevivência das empresas paulistas entre 2003 e 2007 / André Luís Pavão. – São Paulo, 2012.
119 p.

Tese (Doutorado) – Universidade de São Paulo, 2013.
Orientador: Hélio Nogueira da Cruz.

1. Econometria 2. Modelos de riscos proporcionais 3. Mortalidade 4. Inferência paramétrica 5. Inferência não paramétrica I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade II. Título.

CDD – 330.015195

Aos meus pais e esposa.

AGRADECIMENTOS

Agradeço, primeiramente, à Deus pela dom da vida.

Não tenho como não mencioná-la na sequência, pois sem ela eu não teria conseguido chegar nem mesmo à metade do percurso para concluir essa tese. À minha esposa, Janaína, dedico esse trabalho que, ao final de contas, é dela também.

Aos meus pais, José e Maria, simples assim, cujo papel fraternal e os habituais incentivos foram essenciais durante toda a minha caminhada.

Ao meu irmão Leandro e minha cunhada, Juliana, que além dos incentivos, foram os revisores de partes desse trabalho.

Aos companheiros de pós-graduação: Daniel Coelho e Ricardo Fasti, pelos preciosos comentários e sugestões.

Aos professores Milton Campanário e Luque, cujas sugestões contribuíram para a continua melhoria deste trabalho até sua versão final.

Agradeço a toda equipe econômica do SEBRAE-SP, em especial ao Pedro João Gonçalves, que foi muito solícito desde nosso primeiro contato para firmarmos a parceria que possibilitou a realização deste trabalho, além dos comentários feitos durante todo o processo de elaboração da tese.

Por fim, agradeço ao Prof. Hélio, pela dedicação, contribuição e sensibilidade nas horas mais difíceis dessa nossa caminhada, principalmente na fase em que buscávamos os primeiros resultados. Além de professor e orientador, é um grande amigo.

Não podia deixar de dizer o quanto agradeço à Dirce, pela atenção que sempre me dispensou em todos os meus contatos e solicitações.

Por fim, aos meus amigos que sempre me apoiaram ao longo de todo o período do doutorado, sem citar nomes para não fazer nenhuma injustiça, sou muito grato a vocês.

Que Deus nos abençoe.

RESUMO

Este trabalho apresenta as principais causas para a mortalidade das empresas paulistas criadas entre 2003 e 2007 a partir de base de dados cedida pelo SEBRAE-SP para o desenvolvimento dessa pesquisa. A amostra final, construída a partir de dados disponibilizados pela primeira vez para estudos desta natureza, contou com 662 empresas e 33 variáveis coletadas por meio de questionário aplicado diretamente às próprias empresas.

A análise consistiu no teste de modelos econométricos, baseados na literatura dos modelos de duração, de forma a traduzir quais fatores são mais críticos para a sobrevivência das empresas a ponto de distingui-las em dois grupos: o das empresas vencedoras, cuja longevidade está pautada em ações que promovem ganhos de produtividade e eficiência, e aquelas desprovidas dessas ações e que muito provavelmente deixarão o mercado.

Os três tipos de modelos abordados neste trabalho - não paramétrico, semi-paramétrico (riscos proporcionais) e paramétrico - apresentaram resultados similares, sendo que na abordagem de riscos proporcionais os resultados foram segmentados por tamanho e setor de atuação das empresas. Para as micro empresas, a idade do empreendedor e a iniciativa em investir na qualificação da mão de obra dos funcionários mostraram-se importantes mitigadores do risco de falha desse grupo de empresa, enquanto que para as pequenas empresas, a inovação em processos e a elaboração de um plano de negócios se destacaram dentre o conjunto de variáveis.

Entre empresas dos setores de comércio e serviços, as empresas do primeiro grupo que faziam o acompanhamento das finanças (fluxo de caixa) apresentaram menor risco de falhar. Para aquelas do setor de serviços, a idade do empreendedor, o investimento em qualificação dos funcionários e o tamanho da empresa ao nascer foram importantes para reduzir o risco de falha no tempo.

Outro resultado encontrado, por meio do modelo paramétrico utilizando distribuição Weibull, foi que o risco de a empresa deixar o mercado mostrou-se crescente, pelo menos nos cinco primeiros anos de existência da empresa. Entretanto, esse resultado não deve ser generalizado para períodos de tempo maiores que cinco anos.

ABSTRACT

This thesis presents the main results that determined the bankruptcy of enterprises located in the São Paulo State from 2003 to 2007. The models used in this work were possible due to the partnership with SEBRAE, Small Business Service Supporting, located in the State of São Paulo. This institution provided the data basis for this research and its final version was compound by 662 enterprises and 33 variables, which were collected from a survey done by SEBRAE and the related enterprise. For first time available for research like this

The research was supported by econometrics models, more precisely duration models, which identified the most important factors regarding enterprises survival. Two enterprise groups were distinguished: that one that will survive and grow and another will fail.

In this work, three models were used: parametric, non-parametric and proportional risk with all of them presenting similar results. The proportional risk approach was applied for economic sectors and enterprises size. For the micro size business, the entrepreneurship's age and the resources applied on the employee's qualification were important to reduce the risk to fail in the time, whereas for small enterprises, variables like innovation and business plan building were the most important variables.

For the commerce and service sectors, the enterprises related to the first one, the enterprises which kept attention on financial results (cash flow) presented lower risk to fail. For service sector, variables such as: entrepreneur's age, investment on the employee's qualification and enterprise's size were the most important variables to explain the difference the risk to fail between the enterprises.

Another result presented was the risk to fail, which indicates the likelihood of an enterprise to leave its business activity. In this case, the parametric model using Weibull distribution concluded that the risk grows in the first five years. However, this result must be carefully evaluated since it would be necessary a longer term data to ensure this result.

SUMÁRIO

Capítulo 1 - Aspectos econômicos, institucionais e a sobrevivência das empresas 1
1. Introdução 1
1.1. O problema de pesquisa 4
1.2. Aspectos gerais do estudo e da base de dados da pesquisa 4
1.3. Contexto econômico 5
1.4. Contexto institucional 10
1.5. Revisão da literatura 13
1.6. Conclusão 20

Capítulo 2 – Abordagens paramétrica e não-paramétrica para as empresas 23
2. Introdução 23
2.1. Tratamento e análise da base de dados 24
2.2. Análise da função de risco 33
2.3. Metodologia Empírica 35
2.4. Modelagem não-paramétrica 39
2.5. Teste para as distribuições 43
2.5.1. Distribuição Exponencial 43
2.5.2. Distribuição Weibull 44
2.5.3. Distribuição log-normal 45
2.5.4. Resultados dos testes 45
2.6. Modelagem paramétrica 50
2.6.1. Modelo exponencial 53
2.6.1.1. Ajuste geral do modelo 54
2.6.1.2. Cox-Snell 54
2.6.2. Modelo Weibull 55
2.6.2.1. Ajuste geral do modelo 57
2.6.2.2. Cox-Snell 57
2.7. Conclusão 58

Capítulo 3 - Modelo de Riscos Proporcionais 64
3. Introdução 64
3.1. O modelo de riscos proporcionais 64
3.2. Especificação do modelo 66
3.3. Dados 68
3.4. Resultados 70
3.4.1. Resultados para base sem segmentação 70
3.4.1.1. Ajuste geral do modelo 73
3.4.1.2. Cox-Snell 73
3.4.1.3. Resíduos Martingal 74
3.4.1.4. Resíduos deviance 75
3.4.2. Microempresa 79
3.4.3. Pequena e média empresa 83
3.4.4. Setor de comércio 87
3.4.5. Setor de serviços 91
3.5. Conclusão 95

Capitulo 4 - Conclusão 114

Anexo 1 22
Anexo 2 60
Anexo 3 97
Referências Bibliográficas 117

Lista de Gráficos

Gráfico 1 - Taxa de desemprego – RMSP (%PEA) 6
Gráfico 2 - Evolução do Crédito – Público e Privado 8
Gráfico 3 - Evolução do IGP-M (% a.m) 9
Gráfico 4 - Evolução da taxa de Câmbio (R$/US$) 9
Gráfico 5 - KM para variável "principal motivação 25
Gráfico 6 - Curva de risco para empresas que falharam 34
Gráfico 7 - Curva de risco para todas as empresas 34
Gráfico 8 - Tempo de observação por grupos de empresas 35
Gráfico 9 - Testes de distribuição para a variável "duração 36
Gráfico 10 - Curva de sobrevivência de Kaplan-Meier 40
Gráfico 11 - Modelos versus Kaplan-Meier 48
Gráfico 12 - Modelos linearizados 49
Gráfico 13 - Modelos versus Kaplan-Meier 50
Gráfico 14 - Ajuste geral do modelo – Cox-Snell 55
Gráfico 15 - Kaplan-Meier versus Weibull de risco acumulado 57
Gráfico 16 - Empresas por quantidade de funcionários 68
Gráfico 17 - Sobrevivência projetada 70
Gráfico 18 - Função sobrevivência Kaplan-Meier 71
Gráfico 19 - Ajuste geral do modelo – Cox-Snell 74
Gráfico 20 - Resíduos martingal para variável idade 75
Gráfico 21 - Análise dos resíduos *deviance* 76
Gráfico 22 - Teste de Schoenfeld - experiência 77
Gráfico 23 - Teste proporcionalidade - qualificação M.D.O. 78
Gráfico 24 - Teste de proporcionalidade *versus* Klapan-Meier - qualificação 79
Gráfico 25 - Qualificação mão de obra 81
Gráfico 26 - Ajuste Cox-Snell 81
Gráfico 27 - Martingal idade 82
Gráfico 28 - Martingal tempo planej. 82
Gráfico 29 - Teste Schoenfeld - experiência 83
Gráfico 30 - Teste Schoenfeld - tempo planejam. 83
Gráfico 31 - Inova em processos 85
Gráfico 32 - Plano de negócio 85
Gráfico 33 - Ajuste Cox-Snell 85
Gráfico 34 - Martingal inov. processo 86
Gráfico 35 - Martingal plano negócio 86
Gráfico 36 - Teste Schoenfeld - plano negócio 86
Gráfico 37 - Teste Schoenfeld –inova processo 86
Gráfico 38 - Faturamento 89
Gráfico 39 - Ajuste Cox-Snell 89
Gráfico 40 - Martingal "Faturamento 90
Gráfico 41 - Teste Schoenfeld – "Faturamento 90
Gráfico 42 - Qualificação da MDO 93
Gráfico 43 - Ajuste Cox-Snell 93
Gráfico 44 - Martingal quantidade de funcionários 94

Gráfico 45 - Martingal idade 94
Gráfico 46 - Martingal qualificação da MDO 94
Gráfico 47 - Teste Schoenfeld – idade 95
Gráfico 48 - Teste Schoenfeld –quantidade de funcionários 95
Gráfico 49 - Teste Schoenfeld - qualificação da MDO 95

Lista de Tabelas

Tabela 1 - Pessoal Ocupado por tamanho da empresa 7
Tabela 2 - Empresas por macrorregiões 11
Tabela 3 - Distribuição das empresas por macrorregiões 11
Tabela 4 - Critérios de classificação de empresas por receita bruta 12
Tabela 5 - Critérios de classificação de empresas por número de funcionários 12
Tabela 6 - Teste de logrank para variável “principal motivação 25
Tabela 7 - Resultado do teste de Wilcoxon 26
Tabela 8 - Teste de Cox para a variável “idade do empreendedor 26
Tabela 9 - Resultados testes de logrank, Wilcoxon e Cox 27
Tabela 10 - Variáveis selecionadas para os modelos 28
Tabela 11 - Estatísticas das variáveis selecionadas 29
Tabela 12 - Tempo de planejamento x Principal motivação 30
Tabela 13 - Percentis para a variável duração 31
Tabela 14 - Quadro geral de variáveis 32
Tabela 15 - Função sobrevivência no período de estudo 41
Tabela 16 - Sobrevivência das empresas por setor 42
Tabela 17 - Comparação entre distribuições 47
Tabela 18 - Resultado para modelo exponencial 53
Tabela 19 - Resultado para modelo Weibull 56
Tabela 20 - Resultados para os modelos exponencial e Weibull 58
Tabela 21 - Estatísticas descritivas 68
Tabela 22 - Estatísticas sob perspectiva setorial 69
Tabela 23 - Estatísticas por tamanho 69
Tabela 24 - Decomposição da função risco e sobrevivência 71
Tabela 25 - Resultados para modelo de riscos proporcionais 72
Tabela 26 - Schoenfeld para resíduos padronizados 77
Tabela 27 - Resultados da regressão para microempresas 80
Tabela 28 - Schoenfeld para resíduos padronizados 82
Tabela 29 - Resultados da regressão para PME 84
Tabela 30 - Schoenfeld para resíduos padronizados 86
Tabela 31 - Resultados da regressão para setor de comércio 88
Tabela 32 - Schoenfeld para resíduos padronizados 90
Tabela 33 - Resultados da regressão para setor de serviços 92
Tabela 34 - Schoenfeld para resíduos padronizados 94
Tabela 35 - Resultados dos modelos 96

Lista de Figuras

Figura 1 - Histograma para a variável duração 33

Capítulo 1 - Aspectos econômicos, institucionais e a sobrevivência das empresas

1. Introdução

De acordo com a corrente neoclássica o sistema econômico e seus agentes, responsáveis pelas transações e trocas, são caracterizados pela maximização dos lucros e da utilidade baseados no eixo central da análise que é o próprio mercado. Nesse arcabouço, a atividade econômica rege as inter-relações entre firmas e unidades familiares formando um fluxo circular que não sofre alterações relevantes com o passar do tempo. A partir da abordagem estática do fluxo circular, a teoria Schumpeteriana propõe que o advento do lucro propriamente dito, tem origem nas inovações que provocam falhas no fluxo circular e não na exploração do trabalho ou dos ganhos de capital.

A ocorrência dessas falhas, às quais Schumpeter denominou de “revolucionárias”, constituem a questão central do processo de desenvolvimento econômico. O autor observa: “entenderemos por desenvolvimento, portanto, apenas as mudanças da vida econômica que não lhe forem impostas de fora, mas que surjam de dentro, por sua própria iniciativa” Schumpeter (1985, p. 47)[1].

Nesse sentido, o capitalismo é um sistema que estabelece interação entre os agentes econômicos, famílias e firmas, e as inovações transformam o que já se encontra estabelecido, dando forma a um mecanismo capaz de promover, de maneira contínua, um processo constante de desenvolvimento. Esse sistema econômico estimula a inovação, que se constitui, por excelência, a fonte dos lucros. As inovações, por outro lado, também provocam a chamada "destruição criadora" que representa os prejuízos associados aos processos produtivos e produtos tornando-os obsoletos pelas inovações.

No livro A Teoria do Desenvolvimento Econômico de 1911, Schumpeter associa o aparecimento das inovações à criação de novas empresas. Evidentemente, apenas uma pequena fração das novas empresas na realidade de qualquer economia real é inovadora no sentido desse autor. Mais tarde, no trabalho intitulado Capitalismo, Socialismo e Democracia (1942), Schumpeter reconhece que as médias e grandes empresas também podem gerar inovações.

---

[1] Posteriormente, os schumpeterianos admitiram que as inovações menores, não tão revolucionárias, também tinham os mesmos efeitos analíticos.

Essa forma de inserção de novos agentes na atividade econômica pode ser ratificada pela quantidade de empresas criadas todos os anos. Para pontuar, somente no Estado de São Paulo, cerca de 134 mil empresas são criadas todos os anos. Entretanto, boa parte delas não completa sequer o primeiro ano de vida, segundo o SEBRAE-SP (2008).

Evidentemente, o processo de fechamento de parte dessas empresas acarreta em um custo social tanto em geração de postos de trabalho e renda como na redução da arrecadação. Embora seja "uma porta aberta" aos novos empreendedores dispostos a enfrentar os riscos inerentes a qualquer negócio, o sistema de mercado se encarrega de disparar, continuamente, um processo de seleção entre as unidades produtivas, que no final das contas, premiará as vencedoras e punirá, sem nenhuma hesitação, as empresas sem condições de sobreviver.

A punição ou premiação às empresas faz sentido, na visão Schumpeteriana, pois o crescimento econômico, que consiste na ruptura do fluxo circular, na perturbação do equilíbrio do sistema econômico, acontece por conta das empresas, as quais promoverão "novas combinações" de meios produtivos e que vão definir o processo de desenvolvimento.

Além disso, os novos métodos de produção, os novos mercados e as novas formas de organização da indústria impulsionarão um processo de mutação industrial que revolucionará incessantemente a estrutura econômica a partir de dentro, destruindo incessantemente o antigo e criando elementos novos. Segundo Schumpeter (1961, p. 105-106) "este processo de destruição criadora é básico para se entender o capitalismo e é a ele que toda empresa deve se adaptar para sobreviver".

Posto que algumas empresas conseguirão crescer e permanecer na atividade econômica, e outras, entretanto, serão forçadas a deixar o mercado, este trabalho se propõe a compreender os mecanismos e fatores pelos quais as empresas deixam o mercado identificando, especialmente, para as empresas paulistas que compõem a base de dados, quais fatores contribuem para prolongar a longevidade das empresas e quais aceleram seu fracasso, doravante, falha.

A literatura associada à questão da sobrevivência de empresas é bem estabelecida no cenário internacional, mas pouco aplicada no Brasil, principalmente em trabalhos empíricos utilizando modelos de duração. Dentre alguns dos trabalhos mais conhecidos nessa literatura, tem-se o de Dunne, Roberts e Samuelson (1988). Esses autores analisaram o padrão de entrada e saída de firmas da manufatura americana observando, principalmente, a importância

relativa dos diferentes tipos de entrantes, a correlação da entrada e saída entre indústrias no tempo, o tamanho das entrantes ao nascer e o padrão de saída.

Esses autores encontraram que o tamanho das firmas entrantes determina um padrão de saída diferente entre elas. Como exemplo, eles destacam que uma empresa já constituída em determinado setor e que diversifica sua atuação entrando em outro segmento, tem suas chances de falhar, ou seja, de deixar o mercado, reduzidas quando comparada às demais empresas recém-criadas. Concluíram ainda que, em indústrias com altas taxas de entrada há também, contemporaneamente, altas taxas de saída e que o *turnover* das empresas difere não somente em relação ao tamanho, mas também entre os tipos de indústria.

Embora cada país possa apresentar características peculiares[2] que modelam o comportamento de entrada e saída das empresas do mercado, de forma geral, os resultados empíricos encontrados revelam comportamentos similares entre as empresas entrantes, como veremos na apresentação da literatura.

No contexto internacional, algumas instituições se dedicam a analisar o processo de criação de empresas e ambiente de negócios em vários países. Dentre essas entidades podemos citar o GEM - *Global Entrepreneurship Monitor*. Em seu relatório *Annual Report GEM* (2011), a instituição compara 54 países em temas diretamente ligados ao empreendedorismo. E dentre as análises, ressalta que a crise econômica da zona do euro reduziu a percepção de oportunidade do empreendedorismo em países como Grécia, Hungria, Portugal e Espanha. Constatou, também, que o Brasil é o segundo país em que o empreendedorismo é encarado como uma boa escolha de carreira, atrás da Colômbia, ainda no contexto dos 54 países que compõem o estudo. Outro resultado bastante positivo para o Brasil é o fato de a taxa de participação da mulher no empreendedorismo quase se igualar à masculina, característica que ocorre somente em outros sete países da amostra, a saber: Jamaica, Panamá, Venezuela, Guatemala, Suíça, Singapura e Tailândia. Do lado oposto, o relatório cita o leste europeu, principalmente Polônia e Eslováquia, como lugares em que a participação da mulher é bastante reduzida. Em outro enfoque, esse mais negativo para as empresas do país, o estudo do GEM revela o baixo grau de internacionalização das novas empresas brasileiras.

Além dessa introdução, este capítulo apresenta, na seqüência, o problema de pesquisa deste trabalho, os aspectos gerais que caracterizam a análise, o contexto econômico e

[2] Conforme a legislação vigente e estrutura tributária, por exemplo.

institucional no qual o presente estudo está inserido, além da revisão da literatura de modelos de duração aplicados à sobrevivência de empresas.

1.1. O problema de pesquisa

No Brasil, pouco se estudou sobre o padrão de saída/falha das empresas. Em vista disso, procuramos, por meio da aplicação dos modelos de duração, compreender os mecanismos que aceleram a falha de uma empresa e quais permitem a ela prolongar sua sobrevivência na atividade econômica.

1.2. Aspectos gerais do estudo e da base de dados da pesquisa

Os resultados encontrados nesse trabalho são baseados na utilização dos modelos de duração aplicados à sobrevivência de empresas. Esses modelos podem ser classificados em três tipos: i) não paramétrico, ii) semi paramétrico e iii) paramétrico, sendo que a diferença entre eles, nessa ordem, está no grau crescente de especificação, como será detalhado nos próximos capítulos.

A esses modelos aplicamos base de dados fornecida pelo SEBRAE-SP (2008), proveniente de projeto da referida instituição, intitulado "10 Anos de Monitoramento da Sobrevivência e Mortalidade de Empresas", sendo que o projeto original do SEBRAE-SP compreendeu os anos de 1998 a 2008. No presente estudo, as informações disponibilizadas pela referida instituição compreendeu os anos de 2003 a 2007[3].

Entre os anos de 2003 a 2007 a pesquisa do SEBRAE-SP selecionou, para cada ano, amostra de quatrocentas empresas e aplicou o questionário, sob o qual o levantamento foi baseado, em dezembro de 2008. Essa amostra de quatrocentas empresas a cada ano foi selecionada, aleatoriamente, por meio de metodologia desenvolvida pelo próprio SEBRAE-SP, em meio a um universo, como já citado na introdução, ao redor de 134 mil novas empresas registradas todos os anos na Junta Comercial do Estado de São Paulo (JUCESP).

A base completa do levantamento do SEBRAE-SP, disponibilizada para este estudo totalizou 1.998 empresas, dos mais diversos tamanhos e setores de atuação. Como um dos objetivos desse trabalho, conforme mencionado no problema de pesquisa, é verificar a relação das características individuais sobre a mortalidade das empresas, excluímos da análise as empresas que não possuíam respostas para variáveis apropriadas para realizar o mapeamento

[3] O anexo deste capítulo apresenta a nota metodológica que detalha os critérios de seleção das empresas da amostra.

de perfil individual, tais como escolaridade do empreendedor, quantidade de funcionários, dentre outras. Ainda no contexto das variáveis definidas para análise, excluímos, também, as empresas que não responderam à pesquisa ou não foram localizadas para aplicação do questionário.

O questionário elaborado pelo SEBRAE-SP com perguntas, na maioria delas, de respostas fechadas (objetivas) propiciou a compilação, em forma de base de dados, de uma série de informações inerentes ao setor de atuação de cada empresa, ao perfil de cada empreendedor e a fatores mercadológicos de cada negócio, como veremos próximo capítulo. Além das 33 variáveis que compõem a base de dados desse estudo, tem-se a data da constituição das empresas na JUCESP e a data do fechamento, caso a empresa tenha sido classificada como encerrada[4]. Tem-se, portanto, o tempo de nascimento e morte das empresas bem definidos ao longo da janela de observação, que vai de janeiro de 2003 a dezembro de 2008, totalizando 71 meses.

Dentre os critérios pautados nos possíveis resultados estatísticos a serem obtidos, seguimos a recomendação de Mata e Portugal (1994) e selecionamos apenas os setores com pelo menos dez empresas entrantes. Nesse sentido, foi possível aplicar os testes não paramétricos (Kaplan-Meier) aos grupos de empresas, por setor e CNAE, e obter a função sobrevivência para cada um deles, como mostrado no tabela 12 do capítulo 2.

Após a aplicação desses filtros definidos *a priori*, permaneceram no estudo 662 empresas. Desse total, 128 deixaram o mercado (falharam ou morreram) de acordo com a metodologia dos modelos de sobrevivência e 535 foram censuradas à direita, ou seja, não sabemos quanto tempo, além do final do período de observação, essas empresas continuaram em atividade.

A próxima seção utiliza como pano de fundo, o cenário econômico do período da amostra analisada, 2003 a 2007, além da contextualização da atual crise econômica, procurando mostrar a importância dos pequenos negócios na geração de postos de trabalho e como o ambiente econômico pode influenciar na sobrevivência das empresas.

1.3. Contexto econômico

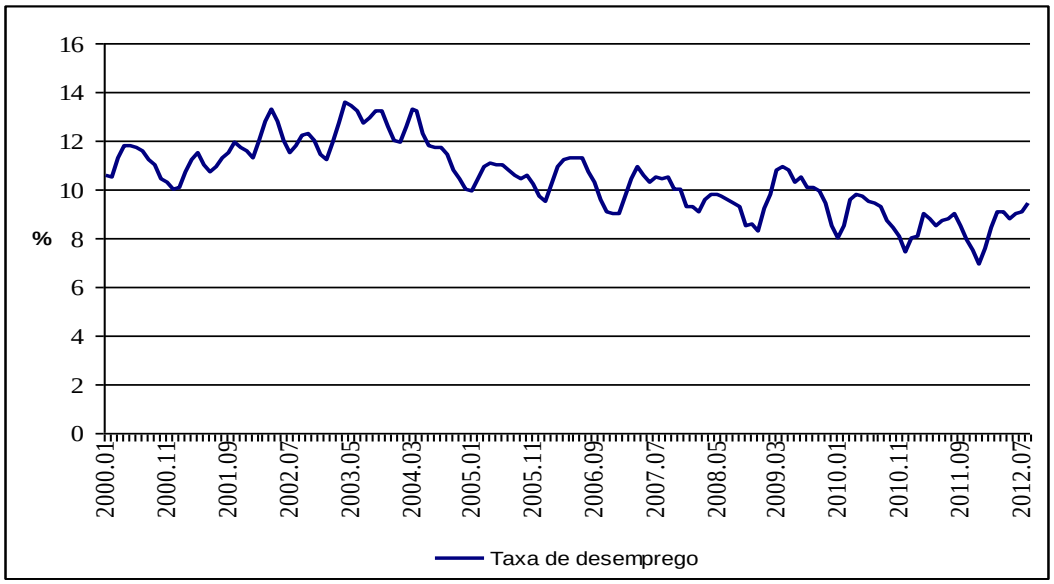
Como destacado na literatura Schumpeteriana, o sistema econômico é acometido por ciclos econômicos que de tempos em tempos perturbam o equilíbrio do fluxo circular entre os agentes econômicos e causam picos de atividade ou depressão econômica, bem como

[4] O item 2.4 do capítulo 2 apresenta a metodologia utilizada pelo SEBRAE-SP para classificar as empresas como encerradas.

possíveis rupturas que promovem quebras de paradigmas que culminam e novo ciclo de crescimento econômico. Nos últimos dez anos, entretanto, os maiores sobressaltos vividos pelo sistema econômico aparentemente não tiveram origem nas revoluções nos meios de produção, mas sim do sistema financeiro. Dentre eles, o mundo experimentou a crise financeira de 2008 e atualmente convive com a crise da zona do Euro. Em ambos os casos, a economia mundial sofreu forte turbulência e os efeitos mais perversos de uma crise econômica mostraram sua face mais cruel: estagnação econômica, desemprego elevado, aumento da dívida pública, etc. Note-se, entretanto, que o período de análise entre 2003 e 2007 é anterior a ambas as crises, como abordaremos adiantes, em mais detalhes. O *approach* utilizado neste trabalho também, entre outras variáveis, procura relacionar as conseqüências dessas crises para a sobrevivência das empresas, principalmente as micro e pequenas empresas brasileiras (MPEs).

A representatividade das micro e pequenas empresas na geração de postos de trabalho, com quase 52%[5] das posições ocupadas, vem permitindo ao país e, mais especificamente, ao Estado de São Paulo, apresentar ao longo dos últimos anos, taxa de desemprego majoritariamente inferior a 10% da população economicamente ativa (PEA), como mostra o gráfico 1.

Gráfico 1 – Taxa de desemprego – RMSP (%PEA)



Fonte: SEADE/SP e elaboração do autor

Nesse sentido, a representatividade das pequenas e médias empresas em relação à geração de postos de trabalho pode ser observada na tabela 1, que retrata a quantidade de

[5] SEBRAE-SP. Anuário do Trabalho na Micro e Pequena Empresa 2010/2011 (vide bibliografia).

pessoal ocupado (com registro formal) por tamanho do estabelecimento. Podemos observar alguns fatos, não surpreendentes, e que revelam as características do país. À medida que as empresas aumentam de tamanho, a formalização torna-se quesito quase que obrigatório. E conforme a empresa cresce, a formalização e a remuneração média se elevam. Além disso, a correlação tamanho/salário sinaliza maior tempo de sobrevivência para as empresas, como confirma alguns dos estudos que serão discutidos na revisão da literatura.

Tabela 1 – Pessoal Ocupado por tamanho da empresa

Faixas de pessoal ocupado total	Pessoal ocupado em 31.12		Salário médio mensal (salários mínimos)	Idade média das empresas
	Total	Assalariado		
0 a 4	5.940.887	1.653.313	1,6	8,8
5 a 9	3.896.664	2.834.501	1,6	11,2
10 a 19	3.968.423	3.365.741	1,8	12,5
20 a 29	2.035.112	1.872.411	2,0	13,4
30 a 49	2.298.543	2.188.807	2,1	14,1
50 a 99	2.709.287	2.635.541	2,4	15,4
100 a 249	3.086.899	3.047.339	2,9	17,8
250 a 499	2.316.365	2.303.132	3,2	21,3
500 e mais	10.932.236	10.920.338	4,1	24,6

Fonte: Cadastro Central de empresas/IBGE/2010 e elaboração do autor

O desempenho das empresas muitas vezes não recai somente na competência individual do empreendedor ou na gestão do negócio em si. É fato que algumas variáveis macroeconômicas interferem, diretamente, nos resultados obtidos pelas empresas, sobretudo em cenários de incerteza. Assim, observamos como pano de fundo da análise os aspectos macroeconômicos ligados a variáveis como crédito, taxas de juros, câmbio, dentre outras.

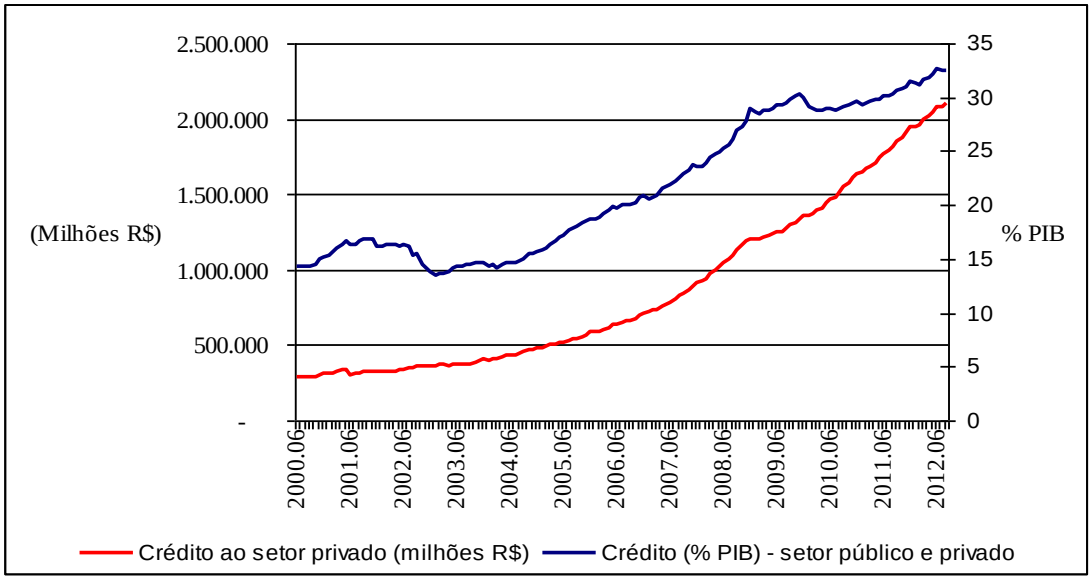
Ademais, a análise recai sobre o período em que os dados do questionário foram coletados, ou seja, entre 2003 a 2008. De forma geral, esse período pode ser considerado de relativa estabilidade para as variáveis destacadas.

Como mencionado logo acima, o crédito desempenha papel crucial em relação ao dinamismo e capacidade de investimento das empresas. Segundo Schumpeter (1982, p.74), o crédito é essencial ao processo econômico, pois parte das inovações são financiadas com recursos de terceiros. Para o empresário, “o concessão de crédito opera como uma ordem para o sistema econômico acomodar os propósitos do empresário, como um comando sobre os bens de que necessita: significa confiar-lhe forças produtivas”.

De forma geral, o crédito tanto para o setor público quanto privado tem aumentado de forma consistente ao longo dos últimos anos. O gráfico 2 fornece a dimensão da expansão do

crédito em percentual do PIB e em montante financeiro para os setores público e privado, de forma que é possível observar o aumento da oferta de crédito, de forma consistente, entre junho de 2000 a junho de 2012.

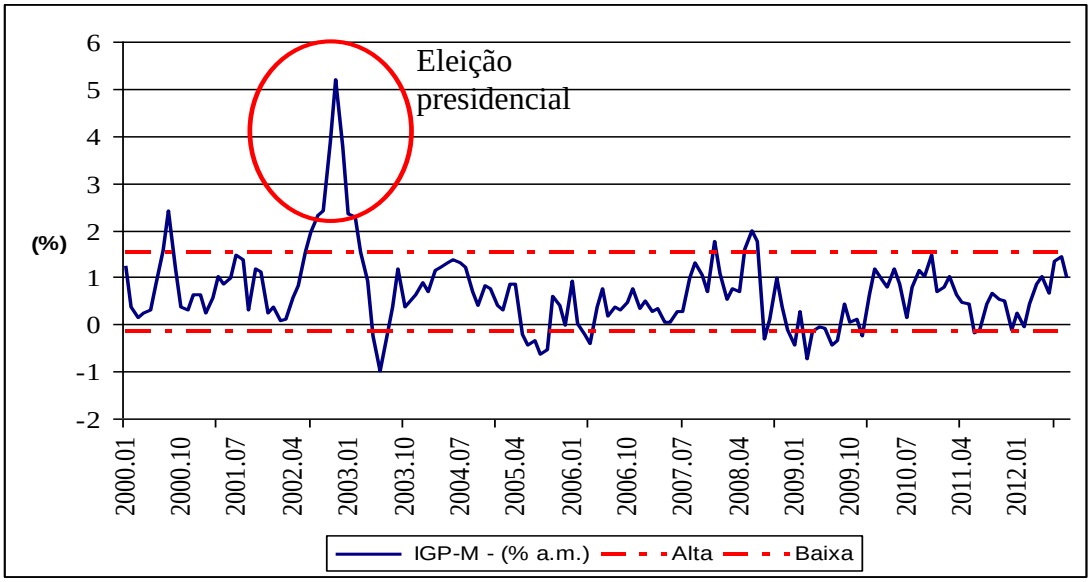
Gráfico 2 – Evolução do Crédito – Público e Privado



Fonte: BACEN e elaboração do autor

Outra variável de destaque é a variação do índice de preços medida pelo Índice Geral de Preços de Mercado (IGP-M). Esse indicador tem impacto direto sobre o preço dos insumos e diversos tipos de contratos que impactam diretamente as MPEs. Conforme destacado no gráfico 3, os efeitos mais significativos de alta foram registrados na eleição presidencial de 2002 e no pico de elevada atividade da economia brasileira no pré-crise das hipotecas norte americanas, ao final de 2008. Entretanto, os efeitos dessa crise chegaram à economia brasileira com certa defasagem temporal, de forma que nos permite afirmar que as empresas da amostra, no período observado, não sofreram diretamente seus efeitos.

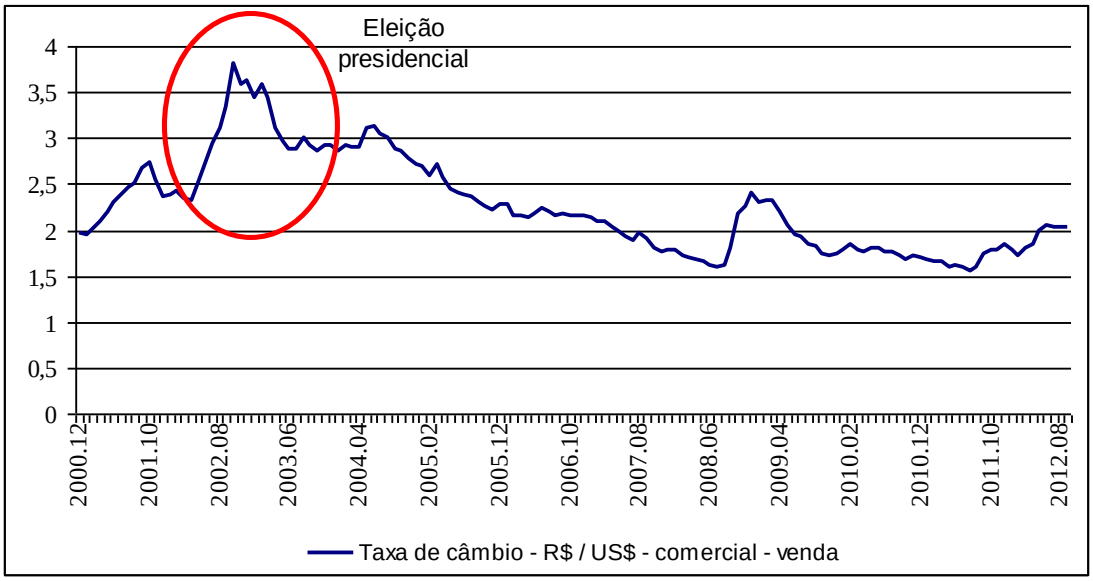
Gráfico 3 – Evolução do IGP-M (% a.m)



Fonte: FGV

O estudo do GEM, mencionado na introdução deste capítulo, revela a baixa internacionalização das novas empresas brasileiras. Essa característica em particular, não recai, exclusivamente, sobre a taxa cambial, (R$/US$), uma vez que o câmbio se manteve desvalorizado, acima de dois reais por dólar de 2003 ao final de 2006[6]. A análise mais detalhada desse comportamento "voltado para dentro" das MPEs brasileiras deve considerar, principalmente, fatores como produtividade e escala. O gráfico 4, mostra a evolução da taxa de câmbio.

Gráfico 4 – Evolução da taxa de Câmbio (R$/US$)



Fonte: BACEN e elaboração do autor.

[6] Os insumos produtivos para as MPES são, salvo raras exceções, de origem nacional. Dessa forma, o custo de importação pode ser desprezado.

Além dos aspectos econômicos, o ambiente institucional pode ser um indutor das atividades empreendedoras ou desestimular, sobremaneira, quaisquer iniciativas nesse sentido. Para tanto, detalhamos algumas características em relação às leis existentes, tanto da esfera federal como estadual, que tenham como premissa promover a atividade das MPEs. Apresentamos, também, a distribuição das empresas brasileiras por tamanho e distribuição geográfica, além dos dois métodos mais utilizados para classificar as empresas quanto ao porte. Esse último ponto tem relação direta com os coortes na base de dados quanto analisamos, no capítulo 3, a sobrevivências das empresas em relação ao tamanho das mesmas.

A próxima seção detalha esses tópicos.

1.4. Contexto institucional

Além das medidas de viés econômico, o governo, em todas as suas esferas, tem promovido medidas institucionais para prover um ambiente mais propicio à atividade empreendedora, principalmente para as MPEs, dada a importância dessas categorias de empresas na geração de novos postos de trabalho e renda[7].

Como medidas institucionais, o governo Federal implantou o SIMPLES (Sistema Integrado de Impostos e Contribuições das Microempresas e das Empresas de Pequeno Porte, instituído pela Lei Federal 9.317 de 5/16/1996 e posteriormente substituída pela Lei Complementar nº 123 de 14/12/2006), regime de recolhimento de impostos simplificados e com alíquotas menores para MPEs com faturamento bruto anual de até R$ 3.600.000 (três milhões e seiscentos mil reais). Entre as vantagens estão redução da carga tributária e a desburocratização.

Na esfera Estadual, mais especificamente no Estado de São Paulo, cujas empresas compõem nossa amostra, as MPEs contribuintes do ICMS puderam se beneficiar do SIMPLES Paulista, que previa a isenção de imposto estadual (ICMS) para empresas com faturamento até R$ 150 mil/ano e alíquotas progressivas para empresas com faturamento até R$ 1,2 milhão/ano, instituído pela Lei Estadual 10.086 de 19/11/1998.

Essas medidas legais, principalmente para o Estado de São Paulo, que concentra a maior quantidade de empresas entre os estado brasileiros, vai de encontro aos números que

[7] No âmbito federal, pode-se citar a criação do Estatuto da Microempresa e da Empresa de Pequeno Porte (Lei Federal n° 9.841, de 5 de outubro de 1999, também revogada pela Lei complementar n° 123, de 14/12/2006). Na esfera Estadual, tem-se a atuação dos bancos estaduais de fomento e também do Banco do Povo.

apresentamos a seguir. Em relação à quantidade de empresas, segundo dados da RAIS/MTE[8] o Brasil possuía, em 2010, mais de três milhões de empresas, distribuídas entre as cinco macrorregiões do país, conforme a tabela 2.

Tabela 2 – Empresas por macrorregiões

Região	Total	%
Norte	131.932	4%
Nordeste	503.321	15%
Sudeste	1.700.339	50%
Sul	748.552	22%
Centro-Oeste	319.304	9%
Total	3.403.448	100%

Fonte: RAIS e elaboração do autor

O maior contraste entre as macrorregiões é o fato de a região Sudeste, a mais rica e populosa do país, concentrar metade das empresas do Brasil enquanto a região norte apenas 4%.

Na tabela 3 podemos observar a distribuição das empresas, por faixas de tamanho, considerando a quantidade de funcionários entre as macrorregiões do país.

Tabela 3 – Distribuição das empresas por macrorregiões

Região	0 Empregado	De 1 a 4	De 5 a 9	De 10 a 19	De 20 a 49
Norte	14.421	68.935	22.403	13.280	7.740
Nordeste	49.855	280.624	84.286	44.735	26.362
Sudeste	181.524	938.160	277.197	155.905	92.452
Sul	88.758	435.343	112.598	60.702	32.692
Centro-Oeste	41.505	184.253	47.135	25.224	13.670
Total	376.063	1.907.315	543.619	299.846	172.916

Região	De 50 a 99	De 100 a 249	De 250 a 499	De 500 a 999	1000 ou Mais
Norte	2.371	1.488	664	370	260
Nordeste	8.216	4.957	2.117	1.265	904
Sudeste	28.855	16.159	5.676	2.532	1.879
Sul	9.701	5.530	1.849	813	566
Centro-Oeste	3.850	2.100	809	426	332
Total	52.993	30.234	11.115	5.406	3.941

Fonte: RAIS e elaboração do autor

[8] Relatório Anual de Informações Sociais do Ministério do Trabalho e Emprego.

A tabela 3 mostra a relevância das pequenas empresas, sendo que o grupo de empresas com até quatro funcionários representam 67% do universo das empresas existentes. E se considerarmos aquelas com até nove funcionários, esse percentual salta para 83%.

A metodologia para classificação das empresas quanto ao tamanho, torna-se peça-chave para determinação das principais características de um grupo de empresas, seja em relação à sobrevivência ou a algum tipo de efeito exógeno que atua sobre cada faixa de tamanho. Há duas formas, no Brasil, mais usadas para se classificar as empresas quanto ao porte, sendo a primeira por faixas de faturamento (receita bruta) e, a segunda, por quantidade de funcionários[9].

No Estatuto da Micro e Pequena Empresa, de 1999, o critério adotado para conceituar micro e pequena empresa é a receita bruta anual, cujos valores foram atualizados pelo Decreto nº 5.028/2004, de 31 de março de 2004[10].

O SEBRAE, por exemplo, adota a quantidade de funcionários como parâmetro para definição do porte. As tabelas 4 e 5 referem-se às duas classificação:

Tabela 4 – Critérios de classificação de empresas por receita bruta

Esfera[11]	**Microempresa**	**Empresa de Pequeno Porte**
Federal	Receita bruta anual igual ou inferior a R$ 360.000,00.	Superior a R$ 360.000,00 e igual ou inferior a R$ 3.600.000,00.

Fonte: Lei complementar nº 123, de 14/12/2006.

Tabela 5 – Critérios de classificação de empresas por número de funcionários

	Microempresa	**Pequena Empresa**
Indústria	De 1 a 19 funcionários	De 20 a 99 funcionários
Comércio Serviços	De 1 a 9 funcionários	De 10 a 49 funcionários

Fonte: SEBRAE

No próximo item, apresentamos a revisão da literatura aplicada aos modelos de duração de empresas. Evidentemente, algumas características interessantes são encontradas, por

[9] Neste trabalho adotamos o número de funcionários no tocante à classificação das empresas por tamanho.

[10] O regime simplificado de tributação - SIMPLES, que é uma lei de cunho estritamente tributário, adota um critério diferente para enquadrar micro e pequena empresa. Os valores foram definidos pela Medida Provisória 275/05 e atualizados pelo Projeto de Lei da Câmara (PLC) 77/11 que ajusta a Lei Geral da Micro e Pequena Empresa (Lei Complementar 123/06).

[11] A maioria dos Estados possui critérios de classificação próprios, dependendo da legislação estadual e situação fiscal.

exemplo, que o tamanho ao nascer, seja ele quanto maior, reduz a probabilidade de falha da empresa ao longo do tempo. Nas conclusões de cada capítulo, cruzamos nossos resultados com aqueles já apresentados pela literatura no sentido de corroborar ou não os resultados encontrados.

1.5. Revisão da literatura

De forma geral, podemos dividir a revisão da bibliografia de modelos de duração aplicados à sobrevivência de empresas em cinco tópicos[12], os quais consideram:

a) Tamanho e idade das firmas com probabilidade de sobrevivência da firma, ao passo que quanto maior o tamanho e tempo de vida, maior a longevidade.
b) Setores capital-intensivo podem ou não determinar barreira à entrada.
c) Meios de produção (bens de capital) do setor como determinante para a sobrevivência.
d) Ambiente macroeconômico e institucional em que as firmas estão inseridas como determinantes para a sobrevivência das empresas.
e) Características como estrutura jurídico-legal; canais de publicidade, etc.

De acordo com a visão tradicional da microeconomia da firma, o excesso de lucratividade em determinado mercado induz à entrada de novas empresas. Essas novas entrantes funcionam como um componente de ajuste, promovendo o equilíbrio de mercado, restaurando o preço e o lucro de longo prazo.

Em relação à dinâmica de mercado, a entrada de novas firmas provoca dois fenômenos: o primeiro estabelece um comportamento mais competitivo das empresas, inclusive das incumbentes, tornando os processos de produção mais eficientes; o segundo provoca um constante processo de inovação no sistema produtivo. E a inovação, seja ela em produtos ou processos, aumenta a produtividade da economia e também promove o crescimento econômico no longo prazo.

Dada a importância desse processo de entrada e saída das empresas para o processo econômico, muitos autores têm buscado respostas para essa dinâmica quase que autônoma de auto-regulação e seleção natural. Os trabalhos precursores que tratam sobre o tema “duração de empresas” têm como foco de análise o fluxo de entrada e saída de firmas do mercado e

[12] No anexo deste capítulo apresentamos uma tabela síntese da literatura abordada neste capítulo destacando, principalmente, os resultados obtidos.

aborda, também, os fatos estilizados característicos do comportamento das firmas recém-criadas[13]. Da literatura, ao menos quatro proposições são firmadas: i) a verificação de altas taxas de entrada na maioria dos países já estudados[14], ii) altas taxas de entrada são acompanhadas contemporaneamente por elevadas taxas de saída, iii) tanto entrantes quanto empresas que deixam o mercado, geralmente são pequenas e iv) somente uma pequena parcela das entrantes conseguirá sobreviver e conquistar uma parcela do mercado.

Saber o que acontece com as empresas recém-criadas possibilita a compreensão dos fenômenos que afetam o desempenho dessas novas empresas, seja pelas características associadas ao setor de atuação ou intrínsecas a cada empresa. Embora o tema *causas mortis* das empresas desperte interesse de planejadores e governos, em todas as esferas, e pelos motivos econômicos já abordados, pouco se estudou, por meio de estudos aplicados, os padrões de falha (saída do mercado) das empresas brasileiras[15].

Alguns trabalhos clássicos, relacionados ao item (a), têm apontado a relação direta entre tamanho e sobrevivência, sendo que a taxa de falha de uma empresa, ou seja, a probabilidade desta em deixar o mercado se reduz conforme aumenta seu tamanho[16], como propõe o trabalho de Dunne, Roberts e Samuelson (1988).

O argumento de que o tamanho da firma importa para a redução da probabilidade de falha também é encontrado em Evans (1987a), que identificou a existência de forte e positiva correlação entre a probabilidade de sobrevivência e tamanho da firma para 81 setores dos 100 que foram analisados em seu trabalho. A principal explicação para essa relação é que grandes empresas têm, mais provavelmente, sua produção convergindo para a escala mínima de eficiência (EME), enquanto as menores firmas apresentam desvantagens que já são inerentes até pelo próprio tamanho, tais como: escala, estrutura de custos, desempenho gerencial e logística.

Em trabalho do mesmo ano, Evans (1987b) concluiu que o crescimento das firmas decresce com a idade, ou seja, com o tempo em que a empresa está estabelecida e com o tamanho da empresa. A explicação para a primeira constatação é baseada na teoria do

---

[13] Para um *survey* sobre todos os fatos estilizados, teóricos e empíricos, ver Geroski (1995).

[14] Dentre os estudos citam-se, por exemplo: Dunne, Roberts e Samuelson (1989) e Audrestch (1991) para os Estados Unidos; Audrestch, Santarelli e Vivarelli (1999) para a Itália; Mata e Portugal (1994) para Portugal. Há também uma edição especial do *International Journal of Industrial Organization*, editado por Mata e Audretsch (1995).

[15] O Sebrae – Serviço de Apoio à Micro e Pequena Empresa – publica, periodicamente, o relatório de mortalidade das novas empresas, mas a abordagem metodológica de modelos de duração não foi empregada.

[16] O tamanho da firma ao nascer leva em conta o número de empregados, em todos os artigos citados.

learning-by-doing, desenvolvida por Jovanovic (1982) e que será discutida mais a frente. A segunda vai na contramão do que é conhecido como *Gibrat Law*[17], a qual diz que o crescimento da firma é independente de seu tamanho.

Em outro artigo, Phillips e Kirchhoff (1989), usando a mesma base de dados de Evans (1987a), também encontraram que a sobrevivência das firmas tende a crescer à medida que a idade das empresas avança, ou seja, quanto mais tempo estabelecidas no mercado, menor é o risco de que venham a deixá-lo.

Outros trabalhos como o de Mata, Portugal e Guimarães (1995) usaram dados em painel para relacionar o tamanho das empresas e a probabilidade de sobrevivência. Eles encontraram que o tamanho após a entrada é mais eficiente ao se analisar as taxas de falha do que o tamanho das empresas no momento da entrada. Uma das hipóteses para essa conclusão é o processo de aprendizado por qual passa a empresa recém-criada. Caso a empresa tenha sucesso no decorrer desse período é provável que ela cresça e se desenvolva, inclusive com provável aumento de porte por conta do aumento dos postos de trabalho.

Os trabalhos discutidos até aqui observam a questão da taxa de risco e da duração média das empresas, mas sem levar em conta possíveis diferenças para essas duas medidas caso as empresas sejam segmentadas por portes. Esse critério para análise foi construído em Holmes, Stone e Braidford (2010), que analisaram a sobrevivência de 931 empresas do nordeste da Inglaterra entre 1973 e 1994 e segmentaram a amostra em micro, pequena e média empresas[18]. Os autores concluíram que o tamanho inicial é relevante somente para as pequenas e médias empresas, cujo impacto sobre a sobrevivência é positivo, ou seja, conforme aumenta o tamanho da firma, a probabilidade de falha ao longo do tempo é menor.

Outra linha de pesquisa sobre duração de empresa aponta a intensidade de capital no setor como sendo determinante para a sobrevivência das empresas. Dentre os trabalhos dessa linha, destacam-se o de Phillips e Kirchhoff (1989), no qual enfatizaram, também, que a maioria das empresas entrantes tende a operar em uma escala de produção sub ótima (em que o custo médio mínimo não se iguala ao custo marginal), pelo menos nas indústrias intensivas em capital. A explicação disso é que as entrantes passam por um período de aprendizagem até conseguirem se tornar eficientes. Nesse sentido, segundo os autores, quanto maior a EME exigida em uma indústria, menor será a probabilidade das entrantes sobreviverem, a menos

[17] Para detalhes ver Simon and Bonini (1958).
[18] A classificação ocorreu de acordo com a quantidade de empregados nos dois primeiros anos.

que elas atinjam um patamar de crescimento que lhes permita alcançar a EME exigida pelo setor em que atuam.

A idéia de que empresas entrantes e que operem fora da EME possam entrar, aprender e se manter no mercado foi proposta por Jovanovic (1982), que introduziu o modelo de *learning-by-doing*. Esse modelo sugere uma espécie de seleção de empresas, em que as mais aptas a aprender e, portanto, atingir a EME, conseguirão permanecer em atividade e, eventualmente, ganhar mercado. Já as que não se tornarem eficientes, deixarão o mercado.

O trabalho de Audrestch (1991) aprofunda as conclusões de Jovanovic (1982) e conclui que a intensidade de capital utilizada em certo setor não representa, por si só, um entrave à entrada de novas empresas, embora o *sunk costs*[19] eventualmente seja. Também destacou que indústrias intensivas em capital não são barreiras à entrada, entretanto, afirma ele, que a taxa de saída é mais expressiva nesses setores, pois caso a empresa não atinja a EME e se adeque às particularidades do setor, será forçada a deixar o mercado. Ainda segundo o autor, o horizonte de tempo em que se analisa a taxa de sobrevivência é bastante relevante para determinar a duração média das empresas, uma vez que há grande variação entre os coortes de tempo analisados.

Ainda nessa linha de pesquisa, Mata (1992) discute as diferenças entre os tipos de entrantes. Segundo o autor, as barreiras à entrada representam um desafio maior para as novas empresas que ainda não possuam outra atividade pré-existente, em relação àquelas que estão entrando como forma de expansão de sua atuação.

Uma terceira corrente de pesquisa sobre duração de empresa destaca nível de tecnologia dado em um setor como determinante para a longevidade das empresas. Nessa vertente, autores como Audrestch (1991), Audretsch e Mahmood (1995) e Dunne, Roberts e Doms (1988) tem destaque.

O trabalho de Audrestsch (1991) foi basilar para a literatura e modelos de duração. Dentre muitas contribuições, o autor constatou que a taxa de sobrevivência das firmas entrantes é positivamente relacionada com a capacidade de inovação da nova empresa, ou seja, se ela for suficientemente eficaz em absorver uma nova tecnologia rapidamente, tenderá a permanecer por mais tempo em atividade. Além disso, apontou que a taxa de sobrevivência dos negócios varia através dos tipos de indústrias e está relacionada às condições da

[19] Custos incorridos pela empresa caso ela realmente decida entrar no mercado, geralmente não recuperáveis, pelo menos em sua totalidade.

tecnologia disponível e da função demanda de cada segmento. Esse resultado foi importante para apontar pelo menos alguma reconciliação entre o aparente paradoxo surgido dos estudos empíricos[20] e que estão ligados à linha de pesquisa (b) que discutimos anteriormente.

Entretanto, as novas empresas que atuam em setores capital-intensivos apresentam probabilidade de sobrevivência menor em relação àquelas de setores em que economias de escala não são preponderantes. Esse resultado também confirma a conclusão de Phillips e Kirchhoff (1989).

Uma limitação do trabalho de Audretsch (1991) é a agregação dos dados utilizados, uma vez que a taxa de sobrevivência encontrada foi basicamente setorial e não em nível dos estabelecimentos com suas variáveis específicas. Isso não possibilitou identificar o desempenho dos novos negócios, conjuntamente com as características particulares de cada empreendimento, característica que o presente estudo engloba.

Vale destacar, também, o trabalho de Dunne, Roberts e Doms (1988), que também examinou o relacionamento entre intensidade de capital aplicado no processo produtivo e a probabilidade de sobrevivência, bem como as taxas de crescimento em nível de firmas. Os autores concluíram que empresas mais intensivas em capital, assim como as que se utilizam de tecnologias mais avançada, não somente tem maior probabilidade de sobreviver como também maiores taxas de crescimento após a entrada.

Como destacado no item (d), uma das vertentes dá mais ênfase ao ambiente de negócios, em que as instituições e cenário econômico, mais especificamente, o macroeconômico impactam diretamente e são determinantes para a sobrevivência das empresas. Nesse contexto, Audretsch e Mahmood (1995) também contribuem e apontam evidências de que as condições macroeconômicas exercem influência sobre a taxa de risco de falha para os novos empreendimentos.

Ainda no cenário econômico, Holmes, Stone e Braidford (2010) mostraram que a taxa de desemprego da economia, tomada como *proxy* para a função demanda agregada, tem impacto sobre a taxa de sobrevivência das micro empresas, mas não das pequenas e médias, enquanto que a taxa de câmbio se mostrou impactante somente para as duas últimas. A justificativa apontada pelos autores é que, como termômetro da demanda, uma alta taxa de desemprego indicaria retração econômica e, portanto, mais dificuldades para os pequenos negócios. Quanto à taxa de câmbio, as oscilações cambiais poderiam dificultar as exportações,

[20] Detalhes em Geroski (1995).

no caso de uma valorização, ou elevar os custos dos insumos importados para as pequenas e médias empresas, no caso de uma desvalorização.

Já o trabalho de Mata, Portugal e Guimarães (1995) concluiu também que é mais provável para uma entrante sobreviver por um período maior em setores que apresentam maior crescimento do que em setores com fraca expansão. A explicação para esse último ponto vem da lógica do próprio mercado, haja vista que em atividades em expansão, as empresas incumbentes não atendem toda a demanda ou por limitação da produção no curto prazo ou mesmo por seleção da própria demanda, fazendo com que as entrantes ocupem o espaço deixado pelas incumbentes.

Por fim, alguns outros artigos apontam outras características que são relevantes para explicar a sobrevivência das empresas. Nessa linha, destacam-se alguns trabalhos, dentre eles o de Mata (1992), que também destaca que as empresas especializadas em determinado produto ou segmento de mercado estão mais sujeitas a economias de escala e a publicidade, enquanto que as empresas mais diversificadas não sofrem muita influência desses dois fatores. O autor ainda aponta que, além da restrição econômica imposta pela necessidade da empresa apresentar lucro em suas operações, outros três fatores também são determinantes para a entrada: tecnologia, fatores ligados a *marketing* e aqueles vinculados a questão financeira.

Tipicamente, as novas firmas atuam abaixo da EME do setor, como dito anteriormente. Isso proporciona à entrante desvantagem em termos de estrutura de custos frente aos competidores já estabelecidos, tornando ainda mais difícil a sobrevivência dos pequenos negócios. Entretanto, como discutido em Audretsch (1995), para aqueles que conseguem sobreviver, a necessidade em reduzir custos acaba por oferecer um forte incentivo para crescer.

Mata, Portugal e Guimarães (1995), também chamam a atenção ao fato de que os novos competidores são, geralmente, pequenos e quando iniciam suas atividades, acabam por disputar mercado com outras empresas entrantes em estágio pouco anterior, e que também são, na maioria das vezes, pequenos também. Esse comportamento quase que predatório reduz ainda mais o tempo médio de vida dos micro e pequenos empreendimentos. Essa constatação faz com que as pequenas empresas saiam do mercado antes mesmo de se tornar uma ameaça para as empresas líderes. Com uma amostra de empresas de Portugal, os autores verificaram que 20% das novas plantas (empresas) morreram durante o primeiro ano e que mais de 50% não sobreviveram até o quarto ano. No sétimo e último ano de observação e

análise, somente 30% das empresas ainda permaneciam em atividade. Das 30% que conseguiram sobreviver após o sétimo ano, tinham seu tamanho médio duas vezes maior que na fase *startup*.

Outros trabalhos miram na constituição legal da empresas, por exemplo, se o capital é nacional ou estrangeiro. Nesse sentido, Audretsch e Mahmood (1995), concluíram que a probabilidade de um novo negócio sobreviver leva em conta não somente as condições da tecnologia dada e as possíveis economias de escala setoriais, mas também a estrutura patrimonial (se o capital é nacional ou estrangeiro).

Mata e Portugal (2004) corroboram a conclusão de Audretsch e Mahmood (1985) acima e observam a constituição legal da empresas, ou seja, se é limitada, societária ou corporativa. Os autores concluíram que as entrantes domésticas são, geralmente, menores que as estrangeiras, pagam menores salários e empregam mão-de-obra menos qualificada e, na maioria das vezes, adotam a forma de constituição legal mais simples, a limitada. Além disso, as firmas domésticas se comportam diferentemente das estrangeiras em relação à entrada. As estrangeiras, segundo os autores, são mais freqüentes em setores que já contam com a presença de empresas do exterior que, em geral, apresentam barreiras à entrada mais elevadas (com maior economia de escala e maior concentração) que em setores mais procurados pelas entrantes domésticas.

No tocante às ações pré-negócio, a hipótese comumente defendida a esse respeito é de que a elaboração de um plano de negócios reduz a probabilidade de falha da empresa. O estudo de Castrogiovanni (1996) questiona essa correlação e aponta que das 220 empresas que fizeram parte de sua pesquisa, somente 51% daquelas que sobreviveram por mais de três anos tinham algum planejamento formal quando iniciaram suas atividades e os resultados desse planejamento dependeram, em grande parte, do tipo de negócio. Do restante, 70% afirmaram que elaboraram um plano de negócios somente para conseguir financiamentos. Reconhece, entretanto, que a confecção do plano de negócios fornece à empresa, melhores condições para pleitear capital de investimento além de munir o futuro empreendedor com informações mais detalhadas sobre o mercado em que irá atuar.

No próximo capítulo, apresentamos a análise descritiva dos dados e os resultados aplicados para os modelos de duração do tipo paramétrico e não paramétrico.

1.6. Conclusão

Neste capítulo abordamos a importância da atitude empreendedora como geradora de novas oportunidades de emprego e renda, destacando em números a percepção do empreendedorismo em alguns países e, em especial, a posição do Brasil como um dos países mais empreendedores. No entanto, destacou-se que as empresas brasileiras não possuem, em sua grande maioria, orientação ao mercado externo.

Usando o cenário macroeconômico como pano de fundo, procuramos mostrar como se comportaram algumas variáveis no tocante ao ambiente macroeconômico, as quais podem afetar, sobremaneira, a sobrevivência das empresas. Dentro de um contexto de "normalidade" econômica, ressaltamos que os efeitos mais significativos para a economia se fizeram sentir em fases posteriores à janela de observação abordada por este estudo.

Destacamos que o Brasil possuía, em 2010, de acordo com levantamento da RAIS, mais de três milhões de empreendimentos. E, como região mais rica do país, o Sudeste detinha metade das empresas brasileiras enquanto que a região norte, apenas 4%. Ressaltou-se, também, a importância das MPEs na geração de postos de trabalho, sinalizando que à medida que as empresas aumentam de tamanho, a formalização do trabalho apresenta correlação positiva.

O ambiente institucional, tanto em aspectos jurídicos como normativos, vêm evoluindo ao longo das últimas duas décadas, propiciando um ambiente mais dinâmico e propício à atividade empreendedora. Dentre as medidas adotadas pelo governo federal, citamos a criação do SIMPLES (Sistema Integrado de Impostos e Contribuições das Microempresas e das Empresas de Pequeno Porte) proporcionando redução da carga tributária e a desburocratização para empresas com faturamento bruto anual de até R$ 1.200.000. No Estado de São Paulo, especialmente, tivemos a instituição do simples SIMPLES Paulista, que previa a isenção de imposto estadual (ICMS) para empresas com faturamento até R$ 150 mil/ano e alíquotas progressivas para empresas com faturamento até R$ 1,2 milhão/ano.

A relevância do assunto, mostrada pelos inúmeros trabalhos abordados na revisão da literatura, em sua grande maioria empíricos, revelam não só o dinamismo desse campo de pesquisa, mas a possibilidade de se obter resultados a partir de dados reais coletados diretamente em sua origem: as próprias empresas. Esses trabalhos revelaram, considerando os países em que as empresas foram estudadas, características impactam na sobrevivência das empresas nascentes.

Esses estudos abrangem cinco principais resultados como: tamanho e idade das firmas importam para a sobrevivência; acesso a bens de capital (meios de produção) não se como barreira ou não à entrada; Setores intensivos em capital influenciam a sobrevivência; ambiente macroeconômico e institucional como fatores influentes; e características pontuais e específicas como publicidade também se tornam relevantes para explicar o padrão de falha das empresas.

Anexo 1

Resumo da bibliografia utilizada sobre modelos de duração

Resultados/Conclusão	Autores
1. Tamanho e idade das empresas importam para a sobrevivência, ao passo que quanto maior o tamanho, maior a longevidade.	• Dunne Roberts e Samuelson (1988). • Evans (1987a) • Evans (1987b) • Jovanovic (1982) • Phillips e Kirchhoff (1989) • Mata, Portugal e Guimarães (1995) • Holmes, Stone e Braidford (2010)
2. Intensidade de capital no setor como sendo ou não uma barreira à entrada.	• Phillips e Kirchhoff (1989) • Mata (1992) • Jovanovic (1982) • Audrestch (1991)
3. Nível de tecnologia dado como determinante para a sobrevivência.	• Audrestch (1991) • Audretsch e Mahmood (1995) • Dunne, Roberts e Doms (1988)
4. Ambiente macroeconômico e institucional como determinantes para a sobrevivência das empresas.	• Audretsch e Mahmood (1995) • Holmes, Stone e Braidford (2010) • Mata, Portugal e Guimarães (1995)
5. Características como instituição legal, publicidade, dentre outras.	• Mata (1992) • Audrestch (1995) • Mata, Portugal e Guimarães (1995) • Mata e Portugal (2004) • Audretsch e Mahmood (1985) • Castrogiovanni (1996)

Nota Metodologia sobre processo de seleção das empresas da amostra

Esta pesquisa foi realizada a partir do rastreamento (busca) de 3.000 empresas com registro de abertura entre 2003 e 2007. O rastreamento ocorreu entre outubro/2008 e maio/2009. A amostra planejada envolveu dois grupos de empresas: (i) empresas registradas na Junta Comercial do Estado de São Paulo (JUCESP); e (ii) empresas clientes do SEBRAE-SP.

O primeiro grupo foi representado por uma amostra, por conglomerados, em 28 municípios do Estado de São Paulo, tendo sido obtida por sorteio com base no conjunto de registros de abertura de empresas, dos tipos sociedade limitada e empresários (individuais), de acordo com as bases de dados da JUCESP.

O segundo grupo foi representado por uma amostra de empresas obtida por sorteio entre as pessoas jurídicas, com registro de abertura entre 2003 e 2007, que constam no cadastro de clientes do SEBRAE-SP.

O processo de rastreamento envolveu diversas alternativas de busca dos sócios-proprietários, por exemplo, identificação do CNPJ das empresas na Secretaria da Receita Federal, visita ao endereço original registrado na JUCESP, visita a novo endereço, consulta a vizinhos, ao atual inquilino do imóvel ou à imobiliária, contato por telefone, visita à residência dos proprietários, consulta a antigo contador ou advogado da empresa, consultas diversas (p.ex., associação comercial, moradores do bairro, sindicatos, etc.).

Concomitantemente ao rastreamento, procurou-se realizar entrevistas com todos os sócios-proprietários ou ex-sócios-proprietários encontrados, resultando na obtenção de 2.008 entrevistas (1.569 sócios-proprietários de empresas em atividade e 439 de empresas encerradas).

Capítulo 2 – Abordagens paramétrica e não-paramétrica para as empresas

2. Introdução

Neste capítulo utilizamos os modelos paramétrico e não paramétrico para analisar a duração das empresas e seus determinantes. O ponto de análise será o comportamento da probabilidade de falha da empresa com relação à sua duração, ou seja, desde o registro na junta comercial do Estado de São Paulo até o momento em que se verifica a saída dessa empresa do mercado. Enfatizamos, também, os determinantes da duração das empresas à luz da importância das características particulares a cada uma delas, tais como a idade do empreendedor, experiência do empreendedor no ramo de atuação e escolaridade, dentre outras que serão detalhadas ao longo desse capítulo.

O grupo de empresas que compõem a amostra é heterogêneo e é formado por empresas de várias partes do Estado de São Paulo, além, evidentemente, da região metropolitana da capital paulista[21]. Nesse sentido, podemos auferir com mais precisão, uma variação importante nas probabilidades de saída das empresas do mercado de acordo com as características individuais. A incorporação dessas características na análise é uma importante contribuição desse estudo para a literatura sobre mortalidade de empresas no Brasil.

Os modelos dedicados à análise de questões de duração temporal já têm tradição estabelecida dentro da estatística e da econometria[22]. De forma geral, esses modelos são classificados pela literatura em três categorias: não-paramétricos, semi-paramétricos e paramétricos. As nuances entre eles se caracterizam pelas especificações relacionadas à forma funcional e distribuição de probabilidade associada, de forma que, nessa seqüência, os modelos apresentam um grau crescente de hipóteses de especificação, mas também de poder de explicação.

Para os três tipos de modelo, o objetivo central é sempre o mesmo: considerando que a variável t representa a duração observada da transição de um estado para o outro, as duas informações fundamentais são dadas pela função de risco e pela função de sobrevivência.

Na próxima seção, apresenta-se a análise descritiva dos dados bem como expõe os resultados preliminares.

---

[21] As regiões paulistas que foram pontos de coleta dos dados estão disponíveis no relatório SEBRAE/SP (2008).

[22] Wooldridge (2010), Greene (2011), Cox e Oakes (1984), Kalbfleisch e Prentice (1980) e Lancaster (1990) são algumas das referências clássicas. Na literatura nacional, citamos Colosimo (2006).

2.1. Tratamento e análise da base de dados

O primeiro teste aplicado à base de dados foi checar se as respostas do questionário estavam consistentes ou, mais especificamente, estimar a confiabilidade das respostas. Utilizou-se, para este fim, o alpha de Conbrach[23], que mede a correlação entre as respostas em um questionário por meio da análise do perfil das respostas dadas pelos empreendedores entrevistados e tem como respostas, valores dentro do intervalo zero e 1. Destacam-se, entretanto, algumas premissas: *i*) todas as variáveis devem possuir a mesma escala de medição, *ii*) os dados do questionários devem ser agrupados por assunto e, *iii*) o questionário deve ser aplicado a uma amostra significativa. Os dados deste estudo violam duas premissas, pois as variáveis que compõem a base de dados são discretas e contínuas e não, necessariamente, estão dispostas em grupo comuns. Para contornar essas violações, selecionamos somente as variáveis discretas[24] e elencamos, seqüencialmente, as variáveis como dispostas no questionário.

O resultado encontrado para o Alpha foi de 0,73. De forma geral, a literatura não define um valor mínimo para aceitação do teste, mas alguns trabalhos, dentre eles o de Urdan (2001), apontam para um mínimo igual a 0,7. Isto posto, o resultado pode ser interpretado como indicativo de integridade das respostas obtidas pelo levantamento.

Testada a consistência dos dados, o próximo passo consistiu em avaliar quais das 33 variáveis da base de dados serão utilizadas no modelo de sobrevivência. Como teste inicial, construímos os gráficos de Kaplan-Meier (KM) para todas as variáveis dicotômicas selecionadas e aplicamos testes de *logrank* e Wilcoxon para avaliar se há ou não diferenças entre as curvas de sobrevivência[25]. Com o mesmo intuito, para as variáveis continuas, aplicamos o teste univariado de Cox e, em ambos os casos, selecionamos, para compor o modelo final, as variáveis que apresentaram nível de significância menor ou igual a 0,25 ($\lambda^2 <= 0{,}25$)[26].

---

[23] A fórmula para cálculo do Alpha de Conbrach está no anexo deste capítulo.
[24] Exceto as variáveis, idade, quantidade de funcionários, tempo de planejamento e educação, todas as outras são discretas, do tipo 0 e 1.
[25] O anexo deste capítulo detalha os conceitos dos testes logrank e Wilcoxon.
[26] O valor da estatística *logrank*, sob a hipótese de igualdade das curvas de sobrevivência, tem uma distribuição qui-quadrado com dois graus de liberdade. O nível de significância para o p-valor segue a recomendação da literatura para modelos de duração.

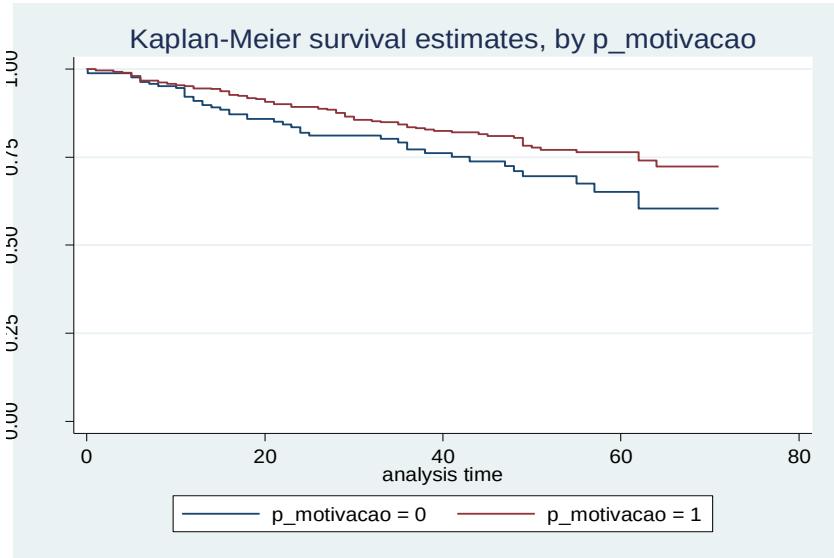
A tabela 6 mostra o valor da estatística do teste de *logrank* para a variável "principal motivação", que foi igual a 3,85, com p-valor (0,0497). Portanto, essa variável deve ser inicialmente considerada nos modelos semi paramétrico e paramétrico.

Tabela 6 – teste de *logrank* para variável "principal motivação"

Principal motivação	Eventos	
	Observados	Esperados
0	40	30.59
1	86	95.41
Total	126	126.00
	chi^2	3.85
	Pr>chi^2	0.0497

As curvas de KM, observadas no gráfico 5, confirmam a estatística do teste para a variável "principal motivação", mostrando que há diferença para a mortalidade das empresas para aqueles que abriram o negócio por se tratar de uma oportunidade de mercado e aqueles que o fizeram por questão de necessidade.

Gráfico 05 – KM para variável "principal motivação"



Outro teste aplicado às variáveis dicotômicas foi o de Wilcoxon, que segue a mesma estrutura formal do teste de *logrank*, com o fato de que os pesos que especificam o teste sejam iguais aos indivíduos que estejam sob risco em um tempo imediatamente inferior a t_j. Para efeito de comparação, tomemos a variável "principal motivação". O resultado do teste é:

Tabela 7 – Resultado do teste de Wilcoxon

Principal motivação	Eventos	
	Observados	Esperados
0	40	30.59
1	86	95.41
Total	126	126.00
	chi^2	3.23
	Pr>chi^2	0.0723

Da tabela 7, temos que a estatística do teste calculada foi de 3.23 com p-valor de (0,0723), sendo, também, pelo método de Wilcoxon, considerada na construção do modelo.

Para as variáveis continuas como idade e número de funcionários, aplicamos o teste univariado de Cox. Tomando a variável idade como exemplo, encontramos estatística do teste de 2.33 e p-valor igual a 0.1266, resultado que também a qualifica para o modelo final. A tabela 8 mostra o resultado da estatística do teste de Cox.

Tabela 8 - teste de Cox para a variável “idade do empreendedor”

Idade	Coeficiente	Desvio padrao	p-valor	chi^2
	-0,01236	0,008189	0,13	2,33

Após a realização dos testes de significância, selecionamos as variáveis que atenderam aos requisitos dos critérios de seleção. Para efeito da modelagem, a variável que atendeu ao menos a um dos testes de seleção, no caso das variáveis dicotômicas, foi considerada, *a priori*, na composição dos modelos. Vale destacar, entretanto, que as variáveis que foram “aprovadas” pelo teste de *logrank* também “passaram” pelo crivo do teste de Wilcoxon. Os resultados estão na tabela 09.

Tabela 09 – Resultados testes de logrank, Wilcoxon e Cox

Variável	Logrank	Wilcoxon	Cox
	teste (p-valor)	teste (p-valor)	teste (p-valor)
Qtde. Funcionários	-	-	0.39 (0.53)
Idade do empreendedor*	-	-	2.33 (0.13)
Escolaridade do empreendedor	2.68 (0.61)	3.3 (0,51)	-
Tempo de planejamento até abrir a empresa*	-	-	9.54 (0,00)
Possuía experiência no ramo*	3.37 (0.06)	3.66 (0.05)	-
Principal motivação (oportunidade/necessidade)*	3.85 (0.05)	3.23 (0.07)	-
Conhecia os concorrentes*	1.61 (0.20)	1.09 (0.29)	-
Conhecia os investimentos e custos iniciais*	1.56 (0.21)	0.27 (0.60)	-
Elaborou plano de negócios*	2.37 (0.12)	0.58 (0.44)	-
Buscava melhoria produtos*	10.06 (0,01)	7.17 (0.00)	-
Acompanhava as finanças*	17.61 (0.00)	12.00 (0.00)	-
Conhecia custos dos produtos*	6.05 (0.01)	4.01 (0.04)	-
Investia na qualificação MDO*	19.86 (0.00)	17.09 (0.00)	-
Conhecia a estratégia dos concorrentes*	6.91 (0.01)	2.76 (0.09)	-
Buscava qualidade e preço*	1.84 (0.17)	2.09 (0.14)	-
Buscava novas tecnologias*	8.48 (0.00)	2.01 (0.15)	-
Inovava em processos*	13.42 (0.00)	5.79 (0.0.1)	-
Tipo constituição empresarial	0.39 (0.53)	0.34 (0.56)	
Sociedade com PJ	0.63 (0.43)	1.15 (0.28)	
Empreendedor (homem ou mulher)	0.15 (0.70)	0.6 (0.44)	
Empreendedor desempregado antes de abrir a empresa	0.55 (0.46)	0.15 (0.70)	
Empreendedor possuía parentes empresários	0.04 (0.84)	0.01 (0.93)	
Estimou o potencial de mercado	0.15 (0.69)	0.14 (0.71)	
Conhecia os aspectos legais inerentes ao negócio	0.35 (0.55)	0.44 (0.50)	
Sabia sobre a melhor localização para a empresa	0.84 (0.36)	0.01 (0.91)	
Havia identificado público-alvo antes abrir empresa	0.05 (0.82)	0.27 (0.60)	
Buscou orientação de instituições antes da abertura	0.83 (0.36)	1.20 (0.27)	
Havia estipulado o lucro pretendido	1.41 (0.26)	1.43 (0.26)	
Calculado o nível de vendas necessário	0.06 (0.81)	0.00 (0.95)	
Havia identificado as principais tarefas	0.05 (0.83)	0.55 (046)	
Planejado o funcionamento na ausência do empreendedor	0.25 (0.61)	0.08 (0.77)	
Definido estratégia para evitar capacidade ociosa	0.25 (0.61)	0.07 (0.79)	
Investia em publicidade	0.15 (0.69)	0.00 (0.99)	

* Variáveis aceitas pelos testes

Após a seleção das variáveis, detalhamos os parâmetros para as variáveis selecionadas, conforme tabela 10. Embora as variáveis “quantidade de funcionários” e “escolaridade” tenham sido rejeitadas nos testes, decidimos mantê-las para efeito de comparação com os resultados obtidos em outros trabalhos destacados na revisão da literatura, no capítulo 1.

Das 33 variáveis testadas inicialmente, o conjunto de informações ficou reduzido a 17 regressores em função da duração das empresas da amostra.

Tabela 10 – Variáveis selecionadas para os modelos

Variável	Descrição da variável	Tipo*
Id	ID	Contínua
Quantidade de funcionários	Quantidade funcionários	Contínua
Tempo de planejamento	Tempo planejamento para abrir o negócio	Contínua
Idade do empreendedor	Idade do empresário	Contínua
Escolaridade do empreendedor	=1 se 1º grau incompleto; =2 se 1º grau completo; =3 se 2º grau completo; =4 se superior completo; =5 se pós-graduado	Escala
Experiência do empreendedor	=1 se tinha experiência anterior; 0	Binária
Principal motivação (oport./necessidade)	=1 se oportunidade; =0 se necessidade	Binária
Conhecia os concorrentes	=1 se conhecia os concorrentes; 0	Binária
Conhecia investimentos e custos iniciais	=1 se conhecia custos e investimentos iniciais; 0	Binária
Elaborou plano de negócios	=1 se elaborou plano de negócios; 0	Binária
Buscava melhoria produtos	=1 se praticava melhoria produto; 0	Binária
Acompanhava as finanças	=1 se acompanhava finanças; 0	Binária
Conhecia custos dos produtos	=1 se acompanhava custo; 0	Binária
Investia na qualificação MDO	=1 se investia na qualificação da mão de obra; 0	Binária
Conhecia a estratégia dos concorrentes	=1 se conhecia estratégia concorrentes; 0	Binária
Buscava qualidade e preço	=1 se buscava qualidade e preço; 0	Binária
Buscava novas tecnologias	=1 se atualizava em relação às tecnologias; 0	Binária
Inovava em processos	=1 se inovava em processos de trabalho; 0	Binária
Duração da empresa	=tempo de observação no estudo (em meses)	Contínua
Censura	=1 se falha; =0 se censurado	Binária

Definida as variáveis que comporão os modelos, analisamos as principais estatísticas descritivas, conforme mostra a tabela 11. Podemos observar que, com exceção das variáveis “quantidade de funcionários”, tempo de planejamento e idade, que são contínuas, as demais são do tipo 1/0. Além disso, a média da variável duração só faz sentido caso admitamos a ausência de censura.

Tabela 11 - Estatísticas das variáveis selecionadas

Variável	Obs.	Média	Desvio padrão	Mínimo	Máximo
ID	662	331,50	191,247	1	662
Qtde. Funcionários	659	5,76	21,564	1	501
Idade do empreendedor	661	39,03	11,241	19	72
Escolaridade do empreendedor	662	2,93	1,010	1	5
Tempo de planejamento até abrir a empresa	662	8,42	15,171	0	120
Possuía experiência no ramo	662	0,75	0,495	0	1
Principal motivação (oportunidade/necessidade)	661	0,74	0,433	0	1
Conhecia os concorrentes	662	0,66	0,470	0	1
Conhecia os investimentos e custos iniciais	661	0,71	0,452	0	1
Elaborou plano de negócios	662	0,54	0,498	0	1
Buscava melhoria em produtos	662	0,86	0,346	0	1
Acompanhava as finanças	662	0,71	0,451	0	1
Conhecia custos dos produtos	661	0,79	0,407	0	1
Investia na qualificação da MDO	662	0,57	0,495	0	1
Estratégia dos concorrentes	661	0,54	0,498	0	1
Buscava qualidade e preço	662	0,82	0,380	0	1
Buscava novas tecnologias	662	0,77	0,421	0	1
Inovava em processos	662	0,80	0,396	0	1
Duração	662	35,98	18,224	0,1	71
Censura	662	0,19	0,395	0	1

O tempo de planejamento observado revela que, considerando todas as empresas da amostra, o tempo médio de planejamento foi de 8,4 meses. Se separarmos as empresas entre aquelas que falharam e não falharam durante o período de observação, o tempo médio de planejamento foi de 5,4 meses e 9,1, respectivamente. Esse resultado era esperado, considerando que um tempo maior para se planejar, na média, pode indicar um maior embasamento em relação às principais características do mercado em que se irá atuar, reduzindo as chances de falha.

Além da análise descritiva, construímos a matriz de correlação entre as 17 variáveis selecionadas. A matriz está no anexo deste capítulo e, destacamos com asterisco os valores das correlações significantes a 5%. Alguns coeficientes, entretanto devem ser observados com cautela, pois o resultado apesar de estatisticamente significante, sua interpretação prática é inócua. Para exemplificar, encontramos que a correlação entre idade e educação se mostrou negativamente correlacionada, ou seja, poderíamos inferir que quanto maior a idade, menor a escolaridade do empreendedor. Por outro lado, todavia, outros resultados se mostraram significativos e positivamente correlacionados como, o acompanhamento das finanças e

conhecimento dos custos iniciais por parte do empreendedor, investimento em qualificação da mão de obra e escolaridade do empreendedor, dentre outros.

Embora a correlação esteja presente entre algumas variáveis, o método de estimação por máxima verossimilhança utilizado nos modelos de duração permite aos estimadores serem consistentes e assintoticamente eficientes, como veremos adiante[27].

Ainda no contexto das análises preliminares, outro resultado interessante vem do cruzamento das variáveis "tempo de planejamento" e "principal motivação". A tabela 12 mostra, por tempo de planejamento, o percentual em relação ao total das empresas que foram abertas por necessidade ou por oportunidade. O resultado dessa tabela deve ser interpretado da seguinte forma: 87,35% das empresas que foram abertas por necessidade, foram planejadas por até 12 meses, ao passo que aquelas que foram constituídas por oportunidade representam 86,9% do total desse grupo.

Nesse sentido, o resultado mostra que as empresas abertas por necessidade tiveram menos tempo para se planejar do que aquelas resultantes de oportunidade de mercado. Isso pode estar ligado à própria condição de dificuldade do futuro empreendedor em encontrar nova colocação no mercado de trabalho. Cabe ressaltar, entretanto, que a diferença entre os dois grupos cai à medida que o tempo de planejamento aumenta. Esse resultado chama a atenção, pois a partir de determinado tempo de planejamento, a duração total do planejamento deixa de ser relevante.

Tabela 12 – Tempo de planejamento x Principal motivação

Tempo de planejamento	Necessidade	Oportunidade
Até 3 meses	54,8%	45,2%
Até 6 meses	78,92%	70,16%
Até 12 meses	87,35%	86,90%

Das 128 empresas da amostra que falharam durante o período de estudo, 67,2% foram abertas sob a justificativa de uma oportunidade de mercado e 32,8% por necessidade. Esse resultado revela um comportamento interessante, pois, apesar de menor tempo de planejamento, aqueles empreendedores que iniciaram sua atividade por necessidade têm, proporcionalmente, menor probabilidade de encerrar o negócio. Uma das possíveis

[27] Uma abordagem detalhada sobre o método de máxima verossimilhança pode ser encontrado em Cramer (1986).

explicações para esse resultado é o fato de o negócio representar a única via de subsistência para o empreendedor e, portanto, a relutância em fechar o negócio tende a ser maior.

A variável “duração”, que aufere o tempo de vida das empresas desde o nascimento até a sua saída do mercado ou até a censura, tem seus percentis apresentados na tabela 13. Se ignorássemos a censura, os percentis revelariam, por exemplo, que 34 meses explica o tempo de vida de 50% das empresas da amostra. Nesse sentido, 62 meses seria a duração de 90% das empresas da amostra.

Tabela 13 – percentis para a variável duração

Percentis				
10%	25%	50%	75%	90%
14	20	34	51	62

Uma visão geral acerca de todas as variáveis consideradas no estudo em relação a proporção das empresas que falharam e não falharam é mostrada na tabela 14**.** É possível observar as nuances que diferem as empresas que falharam e aquelas que ainda continuam em atividade. De acordo com parte da teoria apresentada, podemos verificar que as empresas que nascem com maior número de funcionários tendem a sobreviver por mais tempo.

Cabe destacar que o percentual de respostas positivas das empresas que ainda estavam em atividade ao final do período de observação era, para todas as variáveis, superior ao das empresas que haviam falhado. Tomando como mais relevante uma diferença igual ou maior que 10% entre as respostas, destacam-se as variáveis relacionadas ao investimento em qualificação da mão de obra, à elaboração de plano de negócios, ao acompanhamento da estratégia da concorrência, ao acompanhamento do resultado financeiro da empresa (faturamento e custos), à busca por novas tecnologias e, por fim, à melhoria em processos. Variáveis como idade e escolaridade não foram discrepantes entre os dois grupos.

Tabela 14 – Quadro geral de variáveis

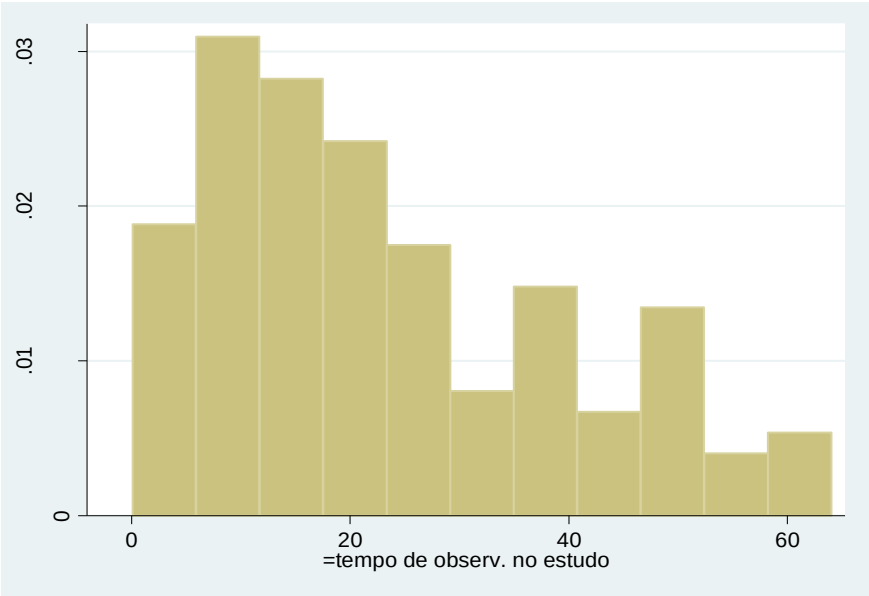
Variáveis	Resposta	Empresas que	
		falharam	Não falharam
Número de funcionários	média	4,7	6,0
Idade do empreendedor	média	39	39
Conhecia investimentos e custos iniciais?	sim	66%	73%
	não	34%	27%
Investe(ia) na qualificação da MDO?	sim	41%	62%
	não	59%	38%
Elaborou Plano de Negócios?	sim	46%	56%
	não	54%	44%
Acompanhava estratégia dos concorrentes?	sim	43%	57%
	não	57%	43%
Buscava melhoria produto?	sim	79%	88%
	não	21%	12%
Buscava fornecedor com melhor preço/qualidade?	sim	77%	84%
	não	23%	16%
Tinha experiência anterior?	sim	51%	59%
	não	49%	41%
Acompanhava as finanças?	sim	55%	75%
	não	45%	25%
Buscava novas tecnologias?	sim	66%	80%
	não	34%	20%
Conhecia os concorrentes?	sim	62%	68%
	não	38%	32%
Acompanhava custos?	sim	70%	81%
	não	30%	19%
Inovava em processos?	sim	68%	84%
	não	32%	16%
Escolaridade	1º grau inc.	9%	12%
	1º grau comp.	18%	15%
	2º grau comp.	49%	46%
	Sup. completo	21%	22%
	Pós-graduação	3%	5%

No item a seguir exploramos um pouco mais os dados e a variável duração a fim de obtermos a forma da taxa de risco para as empresas da amostra, destacando a diferença para aquelas que falharam.

2.2. Análise da função de risco

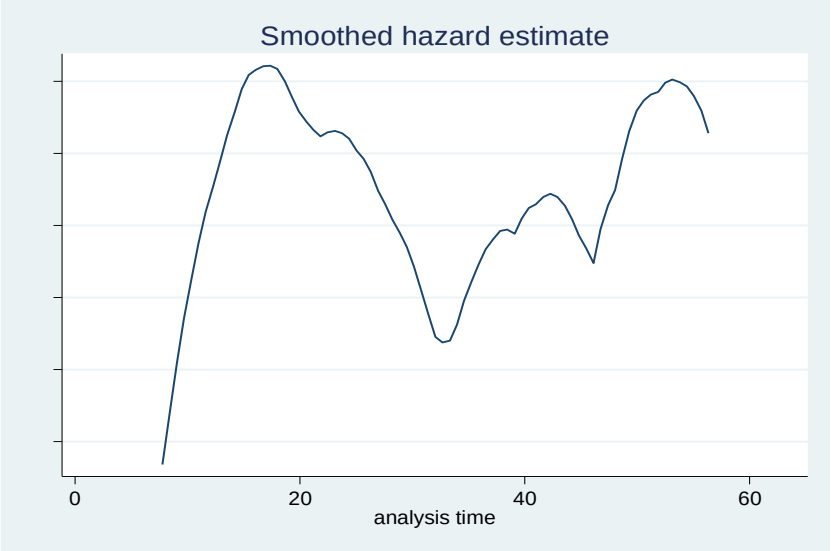
Para checarmos a incidência de falha para o grupo de empresas que deixaram o mercado ao longo do período de estudo, construímos o histograma 1.0, no qual observamos a distribuição da duração dessas empresas. Podemos observar que a incidência de falha é crescente a partir do quinto mês após a constituição da empresa e atinge o ápice ao final do primeiro ano e meio, sendo menos incidente a partir desse ponto. Embora o risco de falha de uma empresa dependa também de condições externas, como as macroeconômicas e regulatórias, conforme exposto no capítulo 1, podemos observar que a taxa de falha se reduz a partir do trigésimo mês. Os picos verificados entre os períodos de 35-40 e 45-55 (mês) podem refletir algum tipo de sazonalidade setorial.

Figura 1 – histograma para a variável duração



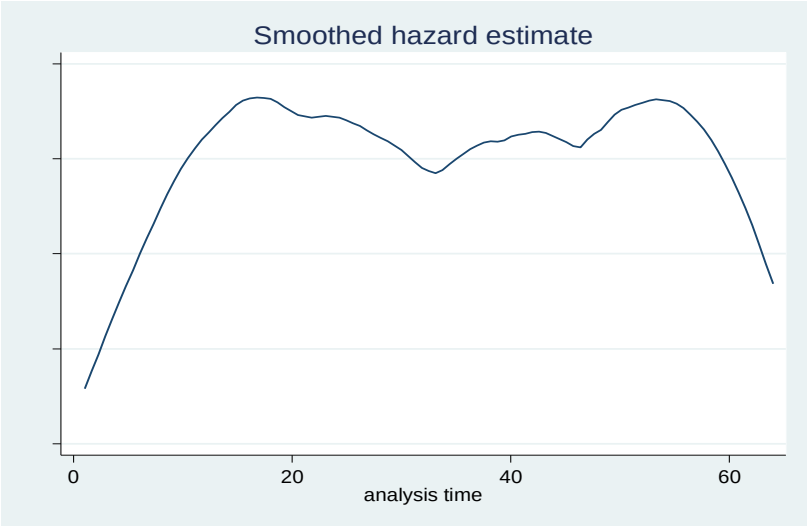
De certa forma, uma aproximação do comportamento da função risco para a taxa de falha das empresas pode ser fornecido pela curva do histograma da figura 1. O gráfico 06 revela o comportamento da taxa de risco para as empresas que falharam no período do estudo.

Gráfico 06 – Curva de risco para empresas que falharam



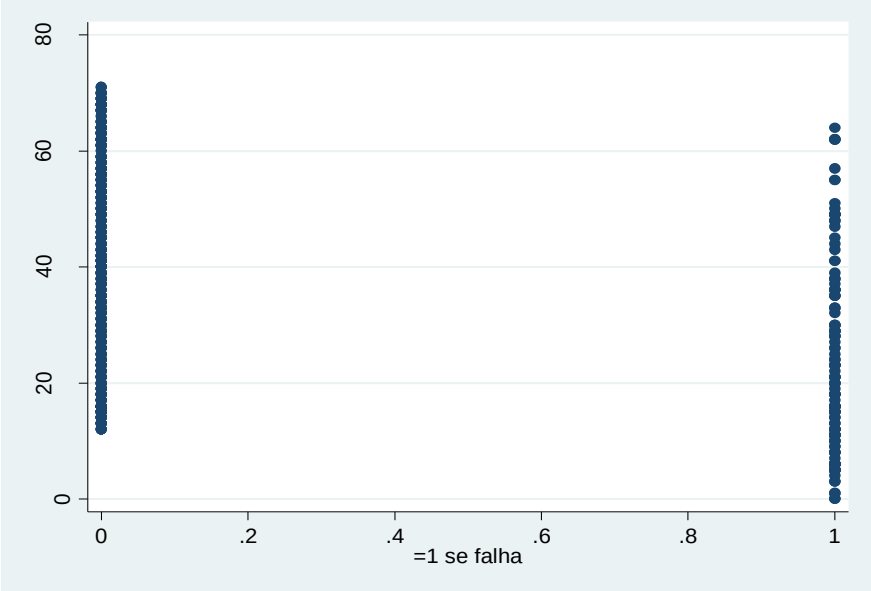
Se tomarmos o conjunto de todas as 662 empresas da amostra e traçarmos a curva de risco, como mostrado no gráfico 07, veremos que a curva de risco se comporta de forma "mais suave" quando comparada à curva das empresas que falharam, mas o resultado é o mesmo, ou seja, a taxa de risco se mostra crescente no início, até o primeiro ano e meio, torna-se relativamente constante entre o vigésimo e qüinquagésimo quinto mês. Entretanto, não podemos afirmar, *a priori*, se a taxa de risco se apresenta mais próxima de uma distribuição exponencial, a qual possui risco constante ao longo do tempo, ou se é mais adequada a uma distribuição de weibull, oscilando entre picos e vales ou ainda um ponto de pico a partir do qual a taxa de risco passa a decrescer em função do tempo, como poderia mostrar uma distribuição log-normal. Esse ponto será objeto de estudo no item 2.6.

Gráfico 07 – Curva de risco para todas as empresas



Das empresas que sobreviveram até o final do período de observação ou foram censuradas, o *range* de tempo de observação vai de 12 a 71 meses. Já para as empresas que falharam, de 0,1 a 64 meses. O gráfico 08 evidencia o comportamento dos dois grupos de empresas, conforme abaixo:

Gráfico 08 – Tempo de observação por grupos de empresas



No próximo item apresentamos a metodologia utilizada.

2.3. Metodologia Empírica

Os primeiros trabalhos utilizando a metodologia empírica dos modelos de duração em economia foram relatados em Kiefer (1988). Dentre as aplicações desse método em economia, destacam-se, por exemplo, o tempo para inserção de uma nova tecnologia, a duração de ativos financeiros, tempo de permanência do aluno na escola e o tempo de vida de empresas.

Quando se trabalha com dados de duração, alguns pré-requisitos precisam ser levados em conta. Um deles se refere à questão da definição precisa da duração, a qual necessita do tempo de origem da observação, a escala de tempo adotada e uma definição para o término do evento.

No caso deste trabalho, novas empresas entram e outras deixam suas atividades todos os anos. Entretanto, de acordo com SEBRAE (2004), 58% das empresas encerram suas

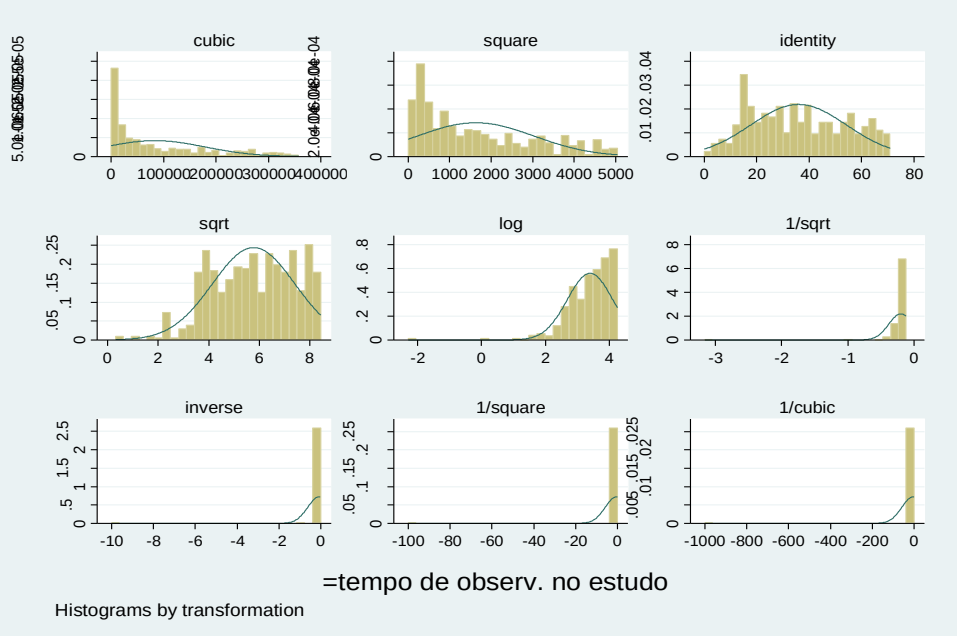
atividades sem dar baixa na JUCESP (Junta Comercial do Estado de São Paulo). Nesse sentido, a metodologia para definir uma empresa como encerrada é:

i. Cujo(a) proprietário(a) foi localizado(a) e informou que a empresa estava em atividade (encerrada formalmente ou não) ou que não chegou a operar desde a sua constituição.
ii. Cujo(a) proprietário(a) foi localizado(a) e informou que a empresa não utilizava mais seu CNPJ (mesmo não oficialmente cancelado) em nenhuma transação comercial.
iii. Cujo CNPJ foi oficialmente cancelado na Secretaria da Receita Federal e/ou
iv. Que não foi localizada, bem como nenhum(a) dos(as) proprietários(as), mesmo após rastreamento exaustivo.

Outras características desse tipo de tratamento é que os modelos de sobrevivência apresentam características como censura e não-normalidade dos resíduos, inviabilizando o uso dos modelos econométricos mais tradicionais, como o de mínimos quadrados ordinários.

Para confirmarmos a impossibilidade de trabalharmos com uma distribuição normal ou gaussiana para os dados de duração, construímos os gráficos de algumas distribuições para a variável “duração”. O gráfico 09 revela que mesmo com uma transformação logarítmica para os dados, não seria possível uma adaptação à normal.

Gráfico 09 – testes de distribuição para a variável “duração”



A idéia central dos conceitos estatísticos presentes nos modelos de duração é a probabilidade condicional, ou seja, se estamos falando em duração de empresas, seria correto indagar não a probabilidade de uma empresa sobreviver exatos 12 meses, mas sim, a probabilidade de determinada empresa sobreviver 12 meses dado que ela sobreviveu os 11 meses anteriores.

A modelagem da probabilidade condicional associada ao momento que uma firma cessará suas atividades envolve o conceito de taxa de risco. A taxa de risco pode ser considerada como a taxa com que as empresas morrem após vigência t, dado que elas estão em operação, pelo menos até o tempo t.

Formalmente, podemos mostrar que o tempo de duração é uma variável aleatória com densidade de probabilidade $f(t)$ e distribuição acumulada $F(t)$, assim temos a função risco da empresa definida como:

$$\lambda(t) = \lim_{h \to 0} \frac{\Pr(t \leq T < t + h) / T \geq t)}{dh} = \frac{f(t)}{1 - F(t)} \tag{2.1}$$

O significado da expressão acima é a probabilidade condicional mensurada em cada instante de tempo t, ou seja, qual a probabilidade da empresa encerrar suas atividades operacionais em certo período, dado que ela sobreviveu até t. Em conseqüência, a função de distribuição acumulada é definida como a probabilidade de uma empresa não sobreviver ao tempo t, que é denominador do lado direto da função risco (2.1):

$$\begin{aligned} S(t) &= 1 - F(t) \\ &= \Pr(T \geq t) \end{aligned} \tag{2.2}$$

Sua interpretação é a probabilidade da empresa continuar em atividade no período t. Formalmente, as funções risco e sobrevivência são diretamente relacionadas e representam duas interpretações possíveis para a mesma informação[28]. Outro conceito importante para

[28] Cabe ressaltar que a função risco é mais informativa que a função sobrevivência, pois diferentes funções de sobrevivência podem ter formas semelhantes, enquanto as respectivas funções de taxa de falha podem diferir drasticamente. Assim, a modelagem da função risco é um importante método para dados de sobrevivência.

dados de sobrevivência é a função de risco acumulada, que fornece a taxa de risco acumulada da empresa e é definida como:

$$\Lambda(t) = \int_0^t \lambda(u)du \tag{2.3}$$

A função acumulada da taxa de risco não tem interpretação direta, mas é útil na avaliação da função de maior interesse que é a taxa de risco $\lambda(t)$[29].

Calculada a função risco $\lambda(t)$, podemos encontrar $F(t)$ e $f(t)$ uma vez que, pela equação da função risco, a derivada de $S(t)$ é $-f(t)$, temos:

$$\lambda(t) = -\frac{d \log S(t)}{dt} \tag{2.4}$$

Dado que $F(0) = 0$ e integrando a função acima, temos:

$$F(t) = 1 - \exp\left[-\int_0^t \lambda(t)ds\right], t \geq 0 \tag{2.5}$$

Diferenciando a função acima, encontramos $f(t)$, dado que

$$f(t) = \lambda(t)\exp\left[-\int_0^t \lambda(t)ds\right] = F^{'}(t) \tag{2.6}$$

A que se considerar, entretanto, que o maior problema encontrado quando se trabalha com dados de duração é a questão da censura. Neste trabalho, a censura se refere àquelas firmas que ainda sobrevivem após o último ano da pesquisa. A solução desse problema vem do conceito de taxa de risco da empresa *i*, uma vez que se considera não somente o momento t em que a produção cessará, mas também o tempo total em que permanecerá em atividade, o qual pode ser obtido por meio da taxa de risco.

---

[29] Isso é mais aparente na estimação não-paramétrica em que $\Lambda(t)$ apresenta um estimador com propriedade ótimas e $\lambda(t)$ é difícil de ser estimada.

O conceito de censura pode ser formalizado tomando T^* como uma variável que represente todo o período de observação na ausência de censura. Seja c o tempo censurado, medido a partir da origem da observação. Assim, tem-se que a variável aleatória a ser observada é menor que T^* ou $T = \min(T^*, c)$. Outra indicação importante é distinguir as observações censuradas, em que $d = 1$ indica que a observação é censurada $(T = c)$, e, $d = 0$ se não há censura $(T = T^*)$. No caso deste trabalho, o tempo até a censura dos dados coincide com o último ponto do tempo levantado na pesquisa, ou seja, dezembro de 2008.

Uma importante verificação a partir da função risco é a questão da dependência da duração. Se a função risco para uma particular distribuição tende para cima (para baixo), então a distribuição tem dependência de duração positiva (negativa). Dependência de duração positiva (negativa) implica que a probabilidade de falha no tempo t, condicionada à duração até t, é crescente (decrescente) em t, ou seja, uma tendência de duração positiva, no caso de empresas, a cada período de tempo, torna-se mais provável que a empresa deixe o mercado. Formalmente, dependência positiva existe para o ponto t^* se $d\lambda(t)/dt > 0$ para $t = t^*$. Para a tendência de duração negativa existe para o ponto t^* se $d\lambda(t)/dt < 0$ para $t = t^*$ e significa que a probabilidade de uma empresa deixar o mercado se reduz ao longo do tempo.

A seguir, apresentamos a análise não-paramétrica deste estudo.

2.4. Modelagem não-paramétrica

Nessa categoria de modelos, a função de sobrevivência é estimada diretamente pela razão entre as empresas ainda existentes no final de cada período t e todas aquelas que estavam em operação no inicio de t. O modelo empregado para o cálculo da função de sobrevivência nesse contexto é o estimador de Kaplan-Meier (K-M). A estimativa de K-M (1958) é uma função *step* com descontinuidades ou saltos em momentos nos quais se observa a interrupção da continuidade no tempo da variável observada. Para os dados com censura, o estimado de K-M assume a forma:

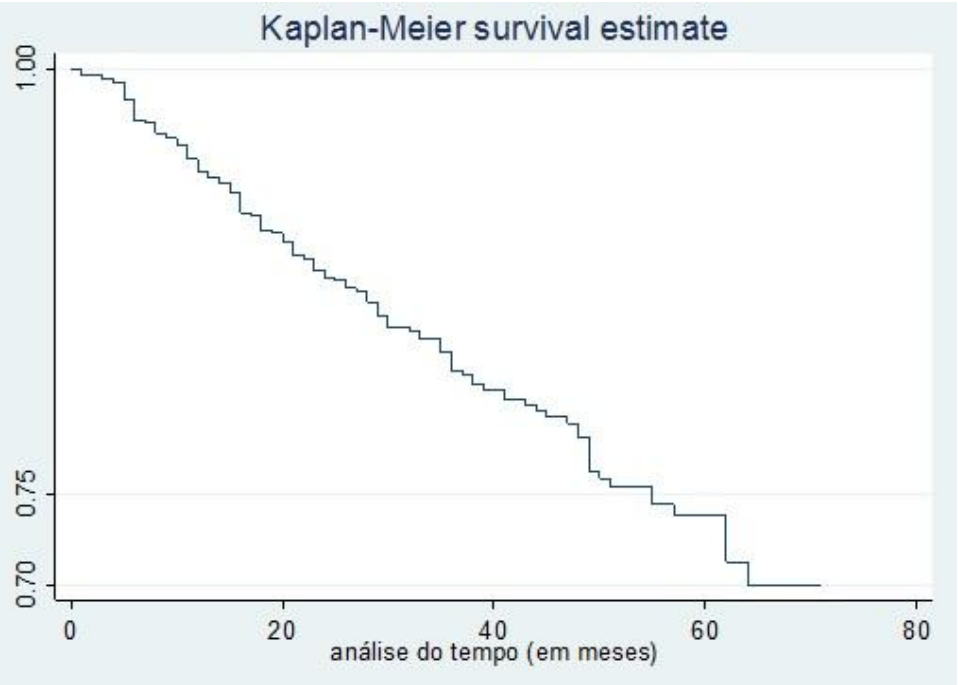
$$\hat{S} = \prod_{i:t_{(i)}<t} \sum_{i=1}^{n} \left(1 - \frac{d_i}{n_i}\right) \tag{2.7}$$

em que d_i é o numero de falhas em $t_i, i = 1,...,k$, e n_i é o numero de indivíduos sob risco em t_i, ou seja, os indivíduos que não falharam e não foram censurados até o instante imediatamente anterior a t_i. A idéia é que as empresas que sobrevivam até t_2 tenham sobrevivido em t_1. Assim, como não há mortalidade entre t_i e t_{i-1}, assume-se que a probabilidade de uma empresa deixar o mercado entre os intervalos é igual zero.

A probabilidade condicional de uma empresa deixar o mercado em t_i, dado que ela estava em operação em t_{i-1} pode ser estimado por d_i / n_i. Assim, a probabilidade condicional para que a empresa sobreviva em t_i é $1 - d_i / n_i$. A probabilidade global incondicional de sobreviver em t é obtida multiplicando as probabilidades condicionais para todos os momentos relevantes até *t*.

Se tomarmos os dados de duração das empresas na amostra sem nenhum tipo de segmentação, teremos um estimador como uma função decrescente em relação ao tempo, haja vista que nem todas as empresas sobreviverão ao final de t. O gráfico 10 mostra a curva de sobrevivência de Kaplan-Meier.

Gráfico 10 – curva de sobrevivência de Kaplan-Meier



O eixo vertical representa o valor da função de sobrevivência, ou seja, a probabilidade de uma empresa continuar no mercado em cada instante do tempo, medido em meses no eixo

horizontal. Podemos observar que ao final do período de observação, aproximadamente, 69% das empresas da amostra ainda estavam em atividade.

Na prática, o resultado do estimador K-M é importante como análise preliminar dos dados. Entretanto, ele não considera possíveis correlações entre as variáveis de controle disponíveis na amostra, como já apontamos na análise dos dados.

As vantagens de se trabalhar com o estimador de K-M é que ele atende a algumas características que são interessantes do ponto de vista econométrico, ou seja, ele é não viciado para grandes amostras, converge assintoticamente para um processo gaussiano[30] e é ainda estimador de máxima verossimilhança.

Se tomarmos o tempo de estudo em escala de cinco meses e auferirmos a taxa de sobrevivência, baseada no estimador de K-M, encontraremos os valores informados na tabela 15.

Tabela 15 – função sobrevivência no período de estudo

Tempo (meses)	Função Sobrevivência (K-M)	Desvio-Padrão	[95% int. confiança	
.1	0.9970	0.0021	0.9880	0.9992
5	0.9789	0.0056	0.9646	0.9874
10	0.9517	0.0083	0.9324	0.9656
15	0.9242	0.0103	0.9012	0.9420
20	0.8953	0.0121	0.8688	0.9167
25	0.8727	0.0135	0.8436	0.8967
30	0.8452	0.0151	0.8130	0.8723
35	0.8290	0.0159	0.7966	0.8593
40	0.8091	0.0172	0.7726	0.8404
45	0.7927	0.0184	0.7539	0.8261
50	0.7570	0.0211	0.7127	0.7954
55	0.7417	0.0224	0.6947	0.7827
60	0.7353	0.0231	0.6868	0.7776
65	0.6947	0.0297	0.6322	0.7486
70	0.6947	0.0297	0.6322	0.7486
71	0.6947	0.0297	0.6322	0.7486

Como podemos observar da tabela 15, ao final de 71 meses de observação, temos que 69,47% das empresas da amostra ainda estavam em atividade.

---

[30] A consistência e normalidade assintótica de $\hat{S}(t)$ foram provadas por Breslow e Crowley (1974).

A tabela 16 mostra a distribuição das empresas da amostra por setor de atuação, conforme especificação CNAE 2.0[31]. Ao todo são 25 setores para as 662 empresas da amostra, sendo a coluna N a quantidade de empresas por setor de atividade. A função de sobrevivência de K-M, por setor, foi observada ao final do último mês do período de acompanhamento.

Dentre os setores que se mostraram mais sensíveis frente à sobrevivência, destacam-se: transporte rodoviário coletivo (Cnae 49213), transporte rodoviário de carga (cnae 49302) e comércio varejista de cosméticos (Cnae 47725). Por outro lado, o estimador de K-M não apontou nenhuma falha para mortalidade nos setores de consultoria em gestão empresarial (cnae 70204) e comércio varejista especializado de móveis, colchoaria e artigos de iluminação (cnae 47457).

Tabela 16 – sobrevivência das empresas por setor

Setor	CNAE (2.0)	Atividade	N	Função (K-M)
Comércio	47890	Com. varejista de outros produtos novos não especificados anteriormente	71	0,8575
	47814	Com. varejista de artigos do vestuário e acessórios	54	0,4903
	45307	Com. de peças e acessórios para veículos automotores	53	0,7790
	47440	Com. varejista de ferragens, madeira e materiais de construção	52	0,8835
	47512	Com. varejista especializado de equipamentos e suprimentos de informática	39	0,6242
	47211	Com. varejista de produtos de padaria, laticínio, doces, balas e semelhantes	27	0,7168
	47121	Com. varej. merc. em geral, com pred. aliment. - minimercados, mercearias e armazéns	26	0,6281
	47610	Com. varejista de livros, jornais, revistas e papelaria	24	0,8217
	47725	Com. varejista de cosméticos, produtos de perfumaria e de higiene pessoal	19	0,4176
	47296	Com. varejo prod. aliment. em geral ou espec. não especificados ant.; prod. do fumo	14	0,7714
	45111	Com. varejo e por atacado de veículos automotores	12	0,6111
	47539	Com. varejista especializado de eletrodomésticos e equipamentos de áudio e vídeo	13	0,7404
	47237	Com. varejista de bebidas	12	0,6500
	47636	Com. varejista de artigos recreativos e esportivos	11	0,8889
	47849	Com. varejista de gás liquefeito de petróleo (GLP)	10	0,9000
	47547	Com. varejista especializado de móveis, colchoaria e artigos de iluminação	10	1
Serviços	56112	Restaurantes e outros estabelecimentos de serviços de alimentação e bebidas	94	0,6774
	49302	Transporte rodoviário de carga	33	0,3464
	70204	Atividades de consultoria em gestão empresarial	16	1
	45200	Manutenção e reparação de veículos automotores	15	0,8667
	63119	Tratamento dados, provedores de serv. de aplicação e serv. hospedagem na internet	13	0,7385
	56201	Serviços de catering, bufê e outros serviços de comida preparada	12	0,6250
	49213	Transporte Rod. coletivo de passag. com itinerário fixo, mun. e em região metropol.	11	0,2386
	62015	Desenvolvimento de programas de computador sob encomenda	10	0,9000
Ind.	14126	Confecção de peças do vestuário, exceto roupas íntimas	11	0,9091
Total			**662**	

[31] Maiores detalhes podem ser obtidos no sítio do IBGE (Instituto Brasileiro de Geografia e Estatística).

Na próxima seção, realizamos os testes de distribuição para ajustar a distribuição que melhor se adéqua aos dados.

2.5. Teste para as distribuições

2.5.1. Distribuição Exponencial

Como dissemos anteriormente, a distribuição normal ou gaussiana se mostra adequada para descrever inúmeras variáveis em economia, biologia e na indústria. Entretanto, quando observamos até o tempo de falha, outras distribuições se mostram mais precisas. Nesse tópico vamos analisar dentre três tipos de distribuição: exponencial, weibull e log-normal, aquela que melhor se adéqua aos dados de distribuição do estudo.

A distribuição exponencial apresenta um único parâmetro e possui função de taxa de falha (ou risco) constante. A função densidade de probabilidade para a variável aleatória tempo de falha *T* com distribuição exponencial é expressa em 2.8:

$$f(t) = \frac{1}{\alpha}\exp\{-(t/\alpha)\}, t \geq 0, \tag{2.8}$$

em que $\alpha > 0$ é o tempo de vida. O parâmetro α tem a mesma unidade do tempo de falha *t*. Neste trabalho, por exemplo, *t* é medido em meses, então α também será disposto em mês.

As funções de sobrevivência $S(t)$ e de taxa de falha $\lambda(t)$ são dadas, respectivamente, por:

$$S(t) = \exp\{-(t/\alpha)\} \tag{2.9}$$

$$\lambda(t) = \frac{1}{\alpha}; t \geq 0 \tag{2.10}$$

O fato de a taxa de risco ser constante significa que uma empresa que foi constituída em 2003 e outra constituída em 2005 tem o mesmo risco de falhar em um dado intervalo futuro. Essa propriedade é conhecida como falta de memória da distribuição exponencial. A média e variância dessa distribuição são, respectivamente, α e α^2.

2.5.2. Distribuição Weibull

A aplicação da distribuição de Weibull se deve à característica da taxa de risco ser monótona, ou seja, é crescente, decrescente ou constante.

Para a variável aleatória *T* com distribuição de Weibull, a função densidade de probabilidade é dada por:

$$f(t) = \frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1} \exp\{-(t/\alpha)^{\gamma}\}, t \geq 0, \tag{2.11}$$

em que γ é o parâmetro de forma e α o de escala, ambos positivos. O parâmetro α tem a mesma unidade de medida de *t* e γ não tem unidade.

Para a distribuição de Weibull, as funções sobrevivência e de risco, são respectivamente:

$$S(t) = \exp\{-(t/\alpha)^{\gamma}\} \tag{2.12}$$

$$\lambda(t) = \frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1}; t \geq 0 \tag{2.13}$$

Para $t \geq 0$, α e $\gamma > 0$. Quando $\gamma = 1$, tem-se a distribuição exponencial e, assim, a distribuição exponencial é um caso particular da distribuição de Weibull. Nessa distribuição, a taxa de falha $\lambda(t)$ é estritamente crescente para $\gamma > 1$, estritamente decrescente para $\gamma < 1$ e constante para $\gamma = 1$.

A média e variância da distribuição incluem o uso da função gama, de modo que:

$$E(T) = \alpha\Gamma[1 + (1/\gamma)] \tag{2.14}$$

$$Var(T) = \alpha^2[\Gamma[1 + (2/\gamma)] - \Gamma[1 + (1/\gamma)]^2]^{32} \tag{2.15}$$

---

[32] Sendo a função gama, $\Gamma(k)$, definida por $\Gamma(k) = \int_0^{\alpha} x^{k-1} \exp\{-x\} dx$

2.5.3. Distribuição log-normal

Para a distribuição log-normal, a função densidade de uma variável aleatória *T* é igual a:

$$f(t) = \frac{1}{(2\pi t\sigma)} \exp\{-\frac{1}{2}(\log(t) - \mu)/\sigma)^2\}; t > 0, \tag{2.16}$$

em que μ é a média do logaritmo do tempo de falha e σ é o desvio-padrão. O logaritmo de ma distribuição log-normal de parâmetros μ e σ tem uma distribuição normal com média μ e desvio-padrão σ. Assim, as informações de uma distribuição log-normal podem ser interpretadas como sendo de uma distribuição normal, desde que se considere o logaritmo dos dados em vez dos valores originais.

As funções de sobrevivência e taxa de falha da distribuição log-normal são, respectivamente,

$$S(t) = \Phi\{(-\log(t) - \mu)/\sigma\} \tag{2.17}$$

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{2.18}$$

em que $\Phi(.)$ é a função de distribuição acumulada de uma normal padrão. A função log-normal caracteriza-se por não ser monótona como a Weibull. Elas atingem um máximo e depois decrescem.

A média e variância para a distribuição log-normal são:

$$E(T) = \exp\{\mu + \sigma^2/2\} \tag{2.19}$$

$$Var(T) = \exp\{2\mu + \sigma^2\}(\exp\{\sigma^2\} - 1) \tag{2.20}$$

2.5.4. Resultados dos testes

Os testes foram feitos para as três distribuições e, primeiramente, elaborou-se a função sobrevivência para a variável *T* em um *t* qualquer. Para a distribuição exponencial, encontramos para a variável α valor igual a 186.36. Se tomarmos *t* igual a 10 meses, temos que a função sobrevivência para a distribuição exponencial é:

$$\hat{S}(t)_e = \exp\{-t/186{,}36) = 0.9477 \qquad (2.21)$$

ou seja, a estimativa da sobrevivência das empresas em 10 meses após a data da constituição é igual a 94,77%.

Na distribuição Weibull, encontramos α igual a 170.3448 γ (gama) igual a e 1.063689. Na função de sobrevivência para a distribuição Weibull, considerando *t* igual a 10 meses, teríamos:

$$S(t)_w = \exp\{-(10/170.3448)^{1.063689}\} = 0.9517 \qquad (2.22)$$

Concluímos então, para a função Weibull que, após 10 meses de vida, a estimativa de sobrevivência é de 95,17%.

No caso da distribuição log-normal encontramos μ, média do logaritmo do tempo de falha, igual a 5.267446 e σ, desvio-padrão, igual a 1.834314. A função sobrevivência para a distribuição log-normal fica:

$$S(t) = \Phi\{[-\log(10) - 5.267446/1.834314]\} = 0.9475 \qquad (2.23)$$

Conforme distribuição log-normal, a expectativa de sobrevivência após o décimo mês de vida é de 94,75%.

A estimativa de Kaplan-Meier para o mesmo período, *t* igual a 10 meses, é 0.9517 ou 95,17%. Pela análise das funções sobrevivência, a distribuição de Weibull é a que mais se aproxima da distribuição de Kaplan-Meier, ou seja, por esse critério é a que melhor se ajusta aos dados. Entretanto, não podemos descartar as outras distribuições sem o auxilio da análise gráfica das distribuições e do cruzamento feito via distribuição de Kaplan-Meier.

Abaixo, a tabela 17 mostra o comparativo entre as quatro distribuições, considerando até a sétima casa decimal.

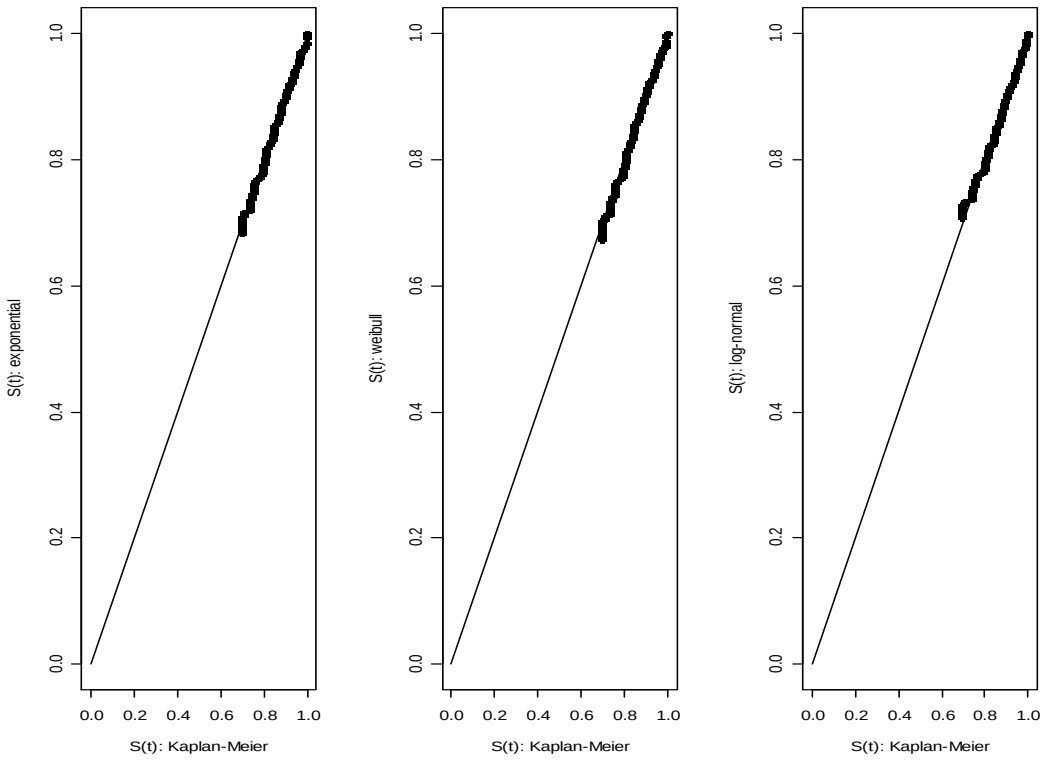
Tabela 17 – comparação entre distribuições

Tempo (em meses)	Kaplan-Meier	Exponencial	Weibull	Log-normal
0,1	0.9969834	0.9994635	0.9996244	0.9999825
1	0.9939668	0.9946484	0.9956960	0.9980102
5	0.9788839	0.9735269	0.9765272	0.9772664
10	0.9517345	0.9477547	0.9516835	0.9475489
15	0.9241798	0.9226647	0.9267111	0.9192392
20	0.8953197	0.8982389	0.9019002	0.8930236
25	0.8726816	0.8744598	0.8773898	0.8688205
30	0.8452078	0.8513102	0.8532598	0.8464210
35	0.8305440	0.8287734	0.8295598	0.8256126
40	0.8091014	0.8068332	0.8063219	0.8062061
45	0.7927193	0.7854738	0.7835669	0.7880397
50	0.7569890	0.7646799	0.7613080	0.7709757
55	0.7417339	0.7444365	0.7395530	0.7548972
60	0.7353397	0.7247290	0.7183055	0.7397045
65	0.6946577	0.7055432	0.6975664	0.7253120
70	0.6946577	0.6868653	0.6773342	0.7116459
71	0.6946577	0.6831895	0.6733484	0.7089940

Podemos observar também, da tabela 17, que as distribuições de Weibull e exponencial possuem estimativas mais próximas enquanto que a distribuição log-normal apresenta um valor estimado ligeiramente diferente dos obtidos pelos outros dois modelos.

Outro método para checar os ajustes dos modelos é por meio do método gráfico. A análise consiste no cruzamento das funções sobrevivência dos modelos propostos com o de Kaplan-Meier $\hat{S}(t)$. O modelo adequado é aquele em que sua curva de sobrevivência se aproximar desse último. O resultado $\hat{S}(t)$ versus $\hat{S}_e(t)$, $\hat{S}(t)$ versus $\hat{S}_w(t)$ e $\hat{S}(t)$ versus $\hat{S}_{\ln}(t)$ estão no gráfico 11.

Gráfico 11 – modelos versus Kaplan-Meier



A partir dos resultados apresentados no gráfico 11, podemos observar que o modelo log-normal parece menos adequado para os dados, pois a curva se apresenta levemente afastada da reta $y = x$. Já os modelos exponencial e Weibull acompanham mais fielmente a reta, indicando melhor adequação aos dados.

Para confirmar os resultados testamos para os modelos linearizados. Para a distribuição exponencial, temos:

$$-\log[S(t)] = (1/\alpha)t \tag{2.24}$$

Já o modelo Weibull linearizado tem a forma expressa como:

$$-\log[S(t)] = (1/\alpha)^{\gamma} \tag{2.25}$$

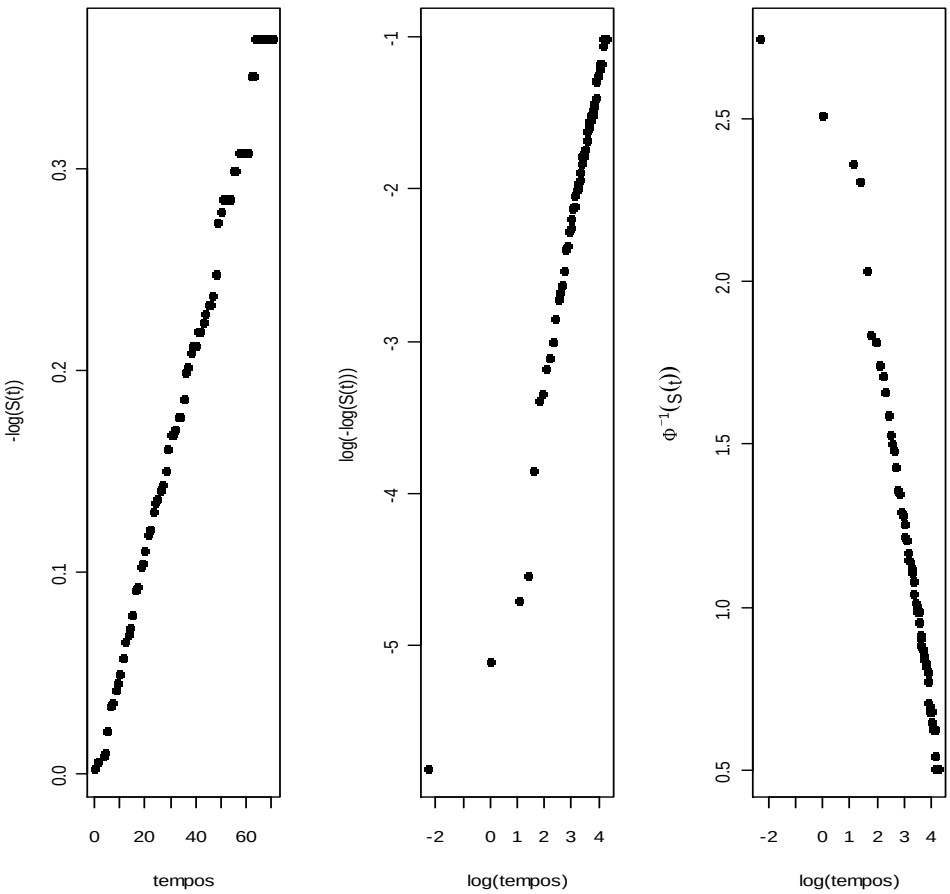
$$\log[-\log[S(t)] = -\gamma \log(\alpha) + \gamma \log(t) \tag{2.26}$$

Por fim, a linearização do modelo log-normal é formalizado conforme equação 2.27

$$\Phi^{-1}(s(t)) = (-\log t + u)/\sigma \tag{2.27}$$

Os resultados dos modelos linearizados estão representados no gráfico 12.

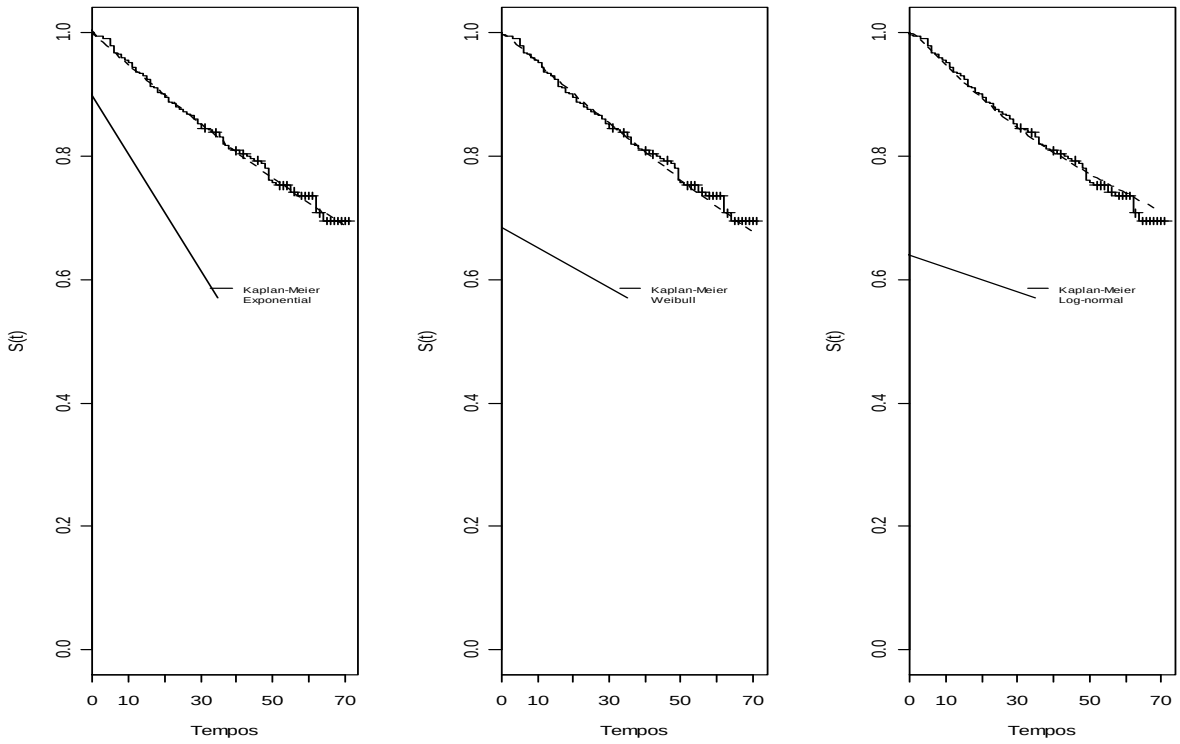
Gráfico 12 – modelos linearizados



Os resultados linearizados deveriam se aproximar de uma reta, formato esse mais próximo das distribuições exponencial e Weibull.

O gráfico 13 mostra as curvas de sobrevivências dos modelos em relação à curva de Kaplan-Meier. Podemos novamente observar que os modelos exponencial e Weibull apresentam ajuste mais adequado para os dados.

Gráfico 13 – modelos versus Kaplan-Meier



Após a escolha do(s) modelo(s) probabilístico(s) mais adequado(s) para os dados, podemos realizar os testes paramétricos com o intuito de verificar como as variáveis influenciam na duração das empresas.

No próximo item, abordamos esse modelo de análise.

2.6. Modelagem paramétrica

A modelagem paramétrica consiste em especificar a forma funcional do modelo e a distribuição de probabilidade para os erros. Há algumas vantagens na especificação do modelo completo, dentre elas: o uso da máxima verossimilhança para estimar os parâmetros, valores ajustados dos modelos podem oferecem tempos de sobrevivência estimados e os resíduos podem ser computados como a diferença entre os valores de tempo observados e projetados.

A modelagem paramétrica também é conhecida como modelos de tempo de vida acelerado, dado que as covariadas podem acelerar o reduzir o tempo de sobrevivência de determinado produto, individuo ou empresa, pois o efeito da covariada é multiplicado sobre a

escala de tempo, como mostrado na equação 2.28. A literatura sobre modelagem paramétrica é discutida em textos como os de Cox e Oakes (1984) e Collett (2003), dentre outros.

Se tomarmos o exemplo mais simples, ou seja, o modelo com apenas uma covariada, teríamos que *T* seria representado pela equação,

$$T = e^{\beta_0 + \phi_1} . \varepsilon \tag{2.28}$$

e como o tempo deve ser positivo, podemos expressar (2.28) como produto do componente $\exp(\beta_0 + \beta_1 x)$ e ε, componente de erro, que também deve apresentar somente valores positivos. Esse modelo pode ser linearizável conforme (2.29),

$$\ln(T) = \beta_0 + \beta_1 x + \varepsilon^* \tag{2.29}$$

em que $\varepsilon^* = \ln(\varepsilon)$. Caso o componente de erro ε siga uma distribuição exponencial, então a distribuição do erro segue uma distribuição do tipo *G*(0,1). O modelo 2.29 é o modelo de regressão exponencial. Esse modelo pode ser generalizado ao permitir que o parâmetro de formato da curva de dependência seja diferente de 1, de tal forma que G(0, σ), cujo modelo é expresso em 2.30,

$$\ln(T) = \beta_0 + \beta_1 x + \varepsilon^* . \sigma \tag{2.30}$$

que é o modelo de regressão de Weibull.

Como os resultados dos testes de distribuição apontaram os modelos exponencial e Weibull como os mais adequados para a variável *T*, aplicaremos esses dois modelos nas regressões.

Para o modelo exponencial, uma variável dicotômica 0/1 tem a razão média do tempo de sobrevivência para o grupo $x = 1$ e para $x = 0$ TR($x = 1$, $x = 0$) é:

$$TR(x = 1, x = 0) = \frac{t_{50}(x = 1, \beta)}{t_{50}(x = 0, \beta)} = \frac{-e^{\beta_0 + \beta_1} . \ln(0.5)}{-e^{\beta_0} . \ln(0.5)} = e^{\beta_t} \tag{2.31}$$

em que TR é a razão de tempo. A função de risco é,

$$h(t, x, \beta) = e^{-(\beta_0 + \beta_1)} \tag{2.32}$$

A função risco em 2.32 depende somente dos coeficientes e do valor das covariadas, os quais são constantes no tempo.

Como dito anteriormente, os dados que compõem a base deste trabalho são censurados pela direta, ou seja, as observações se findam em 31/12/2008. Assim, o modelo de máxima verossimilhança apresenta características adequadas para a estimação do modelo final, sob a hipótese de que há *n* observações independentes no tempo e, *p* e um indicador de censura do tipo:

$$L(\beta) = \sum \{c_i \ln[f(t_i, \beta, x_i)] + (1 - c_i) \ln[S(t_i, \beta, x_i)]\} \tag{2.33}$$

sob a hipótese de que os erros ε são distribuídos $G(0,1)$, temos que,

$$f(y, \beta, x) = e^{[(v - x\ \beta') - \exp(v - x'\beta)} \tag{2.34}$$

$$S(y, \beta, x) = e^{-\exp(v - x'\beta)} \tag{2.35}$$

em que $y = \ln(t)$. Substituindo as equações 2.34 e 2.35 em 2.33 e simplificando, a função de verossimilhança fica como em 2.36:

$$L(\beta) = \sum c_i z_i - e^{z_i} \tag{2.36}$$

sendo que $z_i = y_i - x_{i'}\beta$, $y = \ln(t)$, $x_i' = (x_{i0}, x_{i1} \ldots x_{ip})$ e $x_{i0} = 1$. As equações de probabilidades são obtidas diferenciando a função de verossimilhança em relação aos parâmetros desconhecidos e igualando a zero.

Como o modelo final é linearizado, conforme abaixo:

$$\log \lambda(t; x) = x\beta + \log \lambda_0(t) \tag{2.37}$$

temos que os resultados reportados para os coeficientes das covariadas são as semi-elasticidades dos $\hat{\beta}_j$. Para recuperarmos a estrutura original do modelo, os resultados precisam se tratados antes de interpretados, de forma que para $\hat{\beta}_j$ pequeno, multiplicamos o resultado reportado da regressão por 100 para a elasticidade do risco em relação a x_j. Se $\hat{\beta}_j$ for grande, aplica-se a função exponencial e subtrai 1 e multiplica o resultado por 100, a fim de se obter a proporção de variação, conforme 2.38,

$$100[\exp(\beta) - 1] = x\% \tag{2.38}$$

2.6.1. Modelo exponencial

A tabela 18 mostra os resultados da regressão para o modelo exponencial. Entre parênteses estão os desvios-padrão.

Tabela 18 – resultado para modelo exponencial

variável	Coeficiente
Idade**	-0,0141 (0,0082)
Ac. Finanças*	-0,5314 (0,1854)
Tempo planejamento**	-0,0189 (0,0101)
investia na qualificação M.D.O.*	-0,6456 (0,1893)
constante	-3,8885 (0,3578)
log likelihood	-403,24

* significante a 5%
** significante a 10%

O resultado da regressão para o modelo exponencial mostra que um ano adicional na idade do empreendedor reduz o risco de falha da empresa em 1,4%[33]. Além disso, a empresa

[33] Evidentemente, a função taxa para a idade não é linear e nem infinita, ou seja, a um ponto limite para essa regra, que não necessariamente coincide com a maior idade observada na amostra (72 anos).

que acompanhava suas finanças apresentou taxa de risco de falha 41,2% menor do que aquela que não o fazia.

O tempo de planejamento, também, mostrou-se importante nesse modelo, de forma que reduziu o risco de falha no tempo em 1,9%.

Por fim, a diferença mais marcante entre os grupos das que falharam e não falharam é o investimento da empresa na qualificação da mão de obra. As empresas que investiam em na qualificação de seus colaboradores apresentaram risco de falha 47,5% menor que as empresas que não investiam.

2.6.1.1. Ajuste geral do modelo

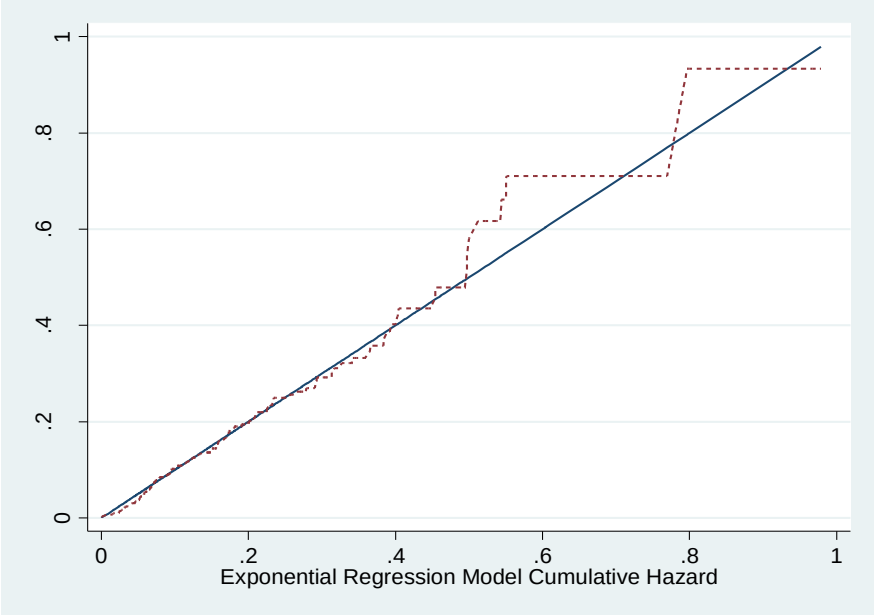
2.6.1.2. Cox-Snell

Após a interpretação das variáveis e como elas afetam a sobrevivência das empresas no tempo, precisamos checar o ajuste global do modelo. Um dos métodos para se verificar o ajuste geral do modelo é o teste do Cox-Snell, em que:

$$e_i = \hat{\Lambda}(t_i / x_i) \tag{2.39}$$

em que $\hat{\Lambda}$ é a função de risco acumulado obtida do modelo ajustado Os resíduos devem seguir uma distribuição exponencial padrão se o modelo for adequado, conforme apontado por (Lawless, 1982). A análise por Cox-Snell consiste em comparar, por meio de gráfico $\hat{e}_i$ *versus* $\hat{\Lambda}(\hat{e}_i)$ que deve ter inclinação igual a 1 (um), quando o modelo for adequado, pois vimos, também, que $\hat{\Lambda}(\hat{e}_i) = -\log(\hat{s}(\hat{e}_i))$, sendo que $\hat{s}(\hat{e}_i)$ é uma função de sobrevivência dos $\hat{e}_i$´s obtida pelo estimador de Kaplan-Meier. Então, o ajuste do modelo se faz comparando a sobreposição das curvas de Kaplan-Meier com os resíduos do modelo, como no gráfico 14:

Gráfico 14 – Ajuste geral do modelo – Cox-Snell



O modelo de regressão exponencial possui poder preditivo relativamente aderente, mas não sobre toda a curva de falha, ou seja, o modelo por si só não é capaz de explicar, totalmente, os fatores que impactam na sobrevivência das empresas da amostra. Provavelmente, fatores externos contribuem nessa questão, tais como: ambiente institucional, econômico, etc.

A significância geral do modelo pode ser testada pela taxa de razão de verossimilhança. A equação 2.40 apresenta o resultado do teste para o modelo exponencial:

$$LR = 2(-403{,}23 - (-398{,}98) = 8{,}5 \qquad (2.40)$$

O valor de 8,5 para razão de verossimilhança é inferior ao valor crítico (22,36) para $(p = 0{,}05)$ e 13 graus de liberdade. Assim, podemos aceitar a hipótese nula, de que as demais covariadas tornam-se insignificantes na presença das demais variáveis relevantes do modelo.

2.6.2. Modelo Weibull

O resultado para o modelo Weibull está na tabela 19.

Tabela 19 – resultado para modelo Weibull

variável	Coeficiente
Idade*	-0,0172 (0,0083)
Ac. Finanças*	-0,5067 (0,1868)
Tempo planejamento**	-0,0182 (0,0099)
Qualificação M.D.O.*	-0,653 (0,1898)
Principal motivação**	-0,3267 (0,1952)
Constante	-4,1969 (0,5044)
log likelihood	-400,04
p	1,1737
	(0,0928)

* significante a 5%

** significante a 10%

Os resultados para o modelo Weibull reportados na tabela 19 apontam, após exponencializá-los, que a adição de mais 1 ano à idade do empreendedor reduz a probabilidade de falha em 1,7%, resultado muito próximo ao obtido pelo modelo exponencial. Em relação ao tempo de planejamento, um mês adicional reduziu o risco de falha em 1,8%.

Assim como no modelo exponencial, no modelo Weibull o investimento na qualificação da mão de obra reduziu o risco de falha em 47,9% em relação à empresa que não investia, resultado muito próximo do obtido pelo modelo exponencial.

A prática de acompanhar as finanças da empresa reduziu o risco de falha em 39,7% ao passo que a empresa constituída por oportunidade de negócio apresentou probabilidade de falha 27,8% menor do que a empresa aberta por necessidade.

Se considerarmos até somente até a primeira casa decimal, o resultado para p de 1,1737 e desvio-padrão de 0.0928 não permite rejeitar a hipótese $H_0 : p = 1$ contra $H_0 : p > 1$, o que não aponta para dependência positiva da duração, ou seja, quanto maior o tempo de vida da empresa, maior o risco de falha. Nesse caso, o resultado obtido pode ser interpretado como dependência constante, sendo, neste caso, a taxa de falha é uma constante ao longo do tempo. Isso quer dizer que, independentemente da idade da empresa, a chance de falha é a mesma em qualquer ponto do tempo.

Outro teste que é realizado para checar o valor de p é regredir o modelo sem as covariadas. Usualmente, o resultado reportado é menor do que no modelo completo. A variação pode confirmar ou refutar a questão da dependência. Entretanto, o resultado reportou p igual 1,1420, com erro padrão 0,0906, resultado que corrobora a análise do parágrafo anterior.

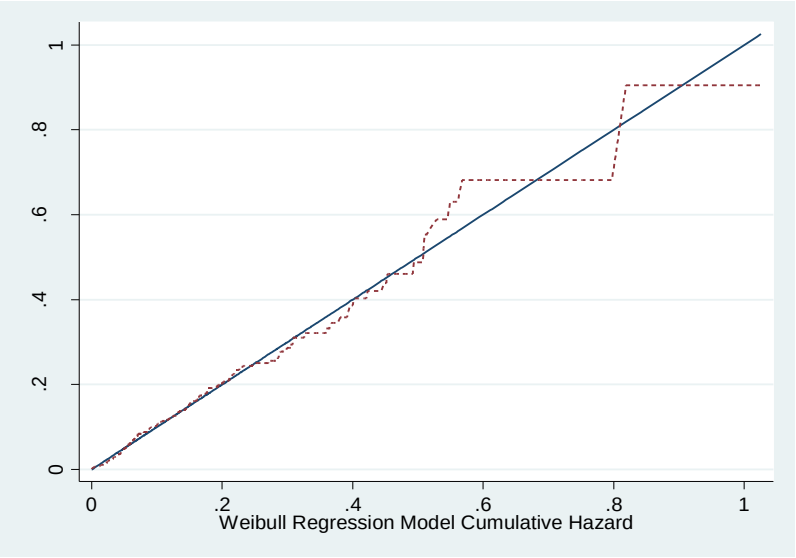
O resultado referente à dependência da duração deve ser observado com cautela, pois vimos tanto na revisão da literatura, como apontado no trabalho Mata, Portugal e Guimarães (1995) como nos resultados reportados pelo SEBRAE-SP (2008), que os 5 primeiros anos são os mais críticos no tocante à probabilidade de morte das empresas. Como nossa amostra tem horizonte de cinco anos, o tamanho do período estudo pode retratar se caracterizar com dependência crescente.

2.6.2.1. Ajuste geral do modelo

2.6.2.2. Cox-Snell

O ajuste geral para o modelo Weibull pelos resíduos de Cox-Snell é mostrado no gráfico 15. O modelo de regressão de Weibull se mostra mais aderente para o diagnóstico dos resíduos de Cox-Snell que o modelo exponencial, embora também sofra efeitos exógenos ao modelo.

Gráfico 15 – Kaplan-Meier versus Weibull de risco acumulado



A significância geral do modelo pode ser testada pela taxa de razão de verossimilhança. A equação 2.41 apresenta o resultado do teste para o modelo exponencial:

$$LR = 2(-400{,}04 - (-396{,}91) = 6{,}26 \tag{2.41}$$

O valor de 6,26 para razão de verossimilhança é inferior ao valor crítico (22,02) para $(p = 0{,}05)$ e 12 graus de liberdade. Assim, podemos aceitar a hipótese nula, de que as demais covariadas tornam-se insignificantes na presença das demais variáveis relevantes do modelo.

2.7. Conclusão

A análise preliminar dos dados identificou alguns resultados anterior ao modelos que podem responder a algumas questões sobre a relevância de determinada variável e seu impacto para a sobrevivência das firmas. Dentre elas, destacamos o investimento na qualificação da mão de obra (capital humano das empresas) e o controle sobre as finanças como pontos cruciais para manutenção da atividade produtiva.

Outro resultado encontrado foi que o tempo de planejamento para os dois grupos de motivação, oportunidade ou necessidade, passa a ser insignificante após 12 meses de planejamento. Uma das explicações possíveis é a perda do *timing* adequado para a abertura e possível redundância de informações acerca do ramo de negócio que se pretende atuar.

Ademais, embora o tempo de planejamento tenha sido menor para aqueles que o fizeram motivado por necessidade e não por oportunidade de mercado, a taxa de mortalidade se mostrou mais elevada para o segundo caso. Uma das explicações para isso pode ser pautada na relutância do empreendedor em encerrar seu negócio se o mesmo representa a única alternativa de renda.

Após os testes das distribuições identificamos que os modelos Exponencial e Weibull são os mais adequados para explicar a distribuição da variável duração. A partir daí, especificamos a forma funcional bem como as duas distribuições associadas, o que nos permitir verificar as variáveis mais significativas para cada modelo, bem como o grau de ajuste de cada um deles. O quadro a seguir resume os resultados obtidos:

Tabela 20 – resultados para os modelos exponencial e Weibull

	Variáveis Significativas por Modelo				
Modelo	Idade	Principal motivação*	Finanças	Qualificação M.D.O.	Tempo de planejamento
Exponencial	-1,4%	-	-41,2%	-47,5%	-1,9%
Weibull	-1,7%	-27,8%	-39,7%	-47,9%	-1,8%

Os valores estão representados em forma negativa para indicar que as variáveis da tabela, ao assumir valor igual a 1 (um), reduzem o risco de falha no tempo conforme percentual apresentado.

Destacamos também, que o efeito da dependência deve ser entendido no contexto da base de dados, cujo período abrange cinco anos. Ademais, o desvio padrão encontrado só permite afirma que a duração é positiva a partir da segunda casa decimal de coeficiente *p*.

No próximo capítulo, aplicamos o modelo de riscos proporcionais em diferentes recortes da base de dados, por setor e tamanho, com objetivo de revelar possíveis diferenças entre grupos mais homogêneos de empresas.

Anexo 2

Teste de logrank

A estatística deste teste é a diferença entre o número observado de falhas em cada grupo e o número esperado de falhas sob a hipótese nula. O teste de logrank tem a mesma expressão do teste *score* para o modelo de regressão de Cox.

Se tomarmos o teste de igualdade de duas funções $S_1(t)$ e $S_2(t)$. Sejam $t_1 < t_2 < ... < t_k$ os tempos de falha distintos da amostra formada pela combinação das duas amostras individuais.

Admitindo que no tempo t_j aconteçam d_j e que n_j indivíduos estejam sob risco em um tempo imediatamente inferior a t_j na amostra combinada e, respectivamente, d_{ij} e n_{ij} na amostra *i*; $i = 1,2$ e $j = 1,...,k$.

Em cada tempo de falha t_j os dados podem ser dispostos em forma de um tabela de contingência 2x2 com d_{ij} falhas e $n_{ij} - d_{ij}$ sobreviventes na coluna *i*

	Grupos 1	Grupos 2	
Falha	d_{ij}	d_{2j}	d_j
Não Falha	$n_{1j} - d_{1j}$	$n_{2j} - d_{2j}$	$n_j - d_j$
	n_{1j}	n_{2j}	n_j

E condicional à experiência de falha e censura até o tempo t_j e ao numero de falhas no tempo t_j a distribuição de d_{2j} segue uma distribuição hipergeométrica,

$$\left(\frac{n_{ij}}{d1j}\right)\left(\frac{n_{2j}}{d_{2j}}\right)/\left(\frac{n_j}{d_j}\right)$$

Com media $w_{2j} = n_{2j} d_j n_j^{-1}$, traduzindo que se não houver diferença entre as duas populações no tempo t_j, o numero total de falhas d_j pode ser dividido entre as duas

amostras de acordo com a razão entre o número de indivíduos sob risco em cada amostra e o numero total sob risco.

A variância de $d_{2j} - w_{2j}$, pela distribuição hipergeométrica fica,

$$(V_j)_2 = n_{2j}(_{n_j} - n_{2j})d_j(n_j - d_j)n_j^{-2}(n_j - 1)^{-1}$$

Por fim a estatística $d_{2j} - w_{2j}$ tem média zero e variância $(V_j)_2$. O teste aproximado para a igualdade das duas funções de sobrevivência é:

$$T = \frac{[\sum_{j=1}^{k}(d_{2j} - w_{2j})]^2}{\sum_{j=1}^{k}(V_j)_2}$$

sob a hipótese nula $H_0 : S_1(t) = S_2(t)$ para todo t no período de observação, tem uma distribuição qui-quadrado com 1 grau de liberdade para grandes amostras. Valendo-se dessa hipótese, quando se tem a generalização da estatística T, temos que a teste fica especificado como:

$$S = \frac{[\sum_{j=1}^{k}(d_{2j} - w_{2j})]^2}{\sum_{j=1}^{k}u_j^2(V_j)_2}$$

em que u_j são os pesos que especificam o teste. O logrank, portanto, é obtido tomando $u_j = 1$ para $j = 1,...,k$.

<u>Teste de Wilcoxon</u>

O teste de Wilcoxon segue a mesma metodologia do teste de logrank e é obtido pela mesma expressão acima, tomando $u_j = n_j$. O teste de Wilcoxon foi adaptado para dados censurados a partir do conhecido teste não-paramétrico de Wilcoxon (Gehan, 1965, Braslow, 1970).

Alpha de Cronbach

O Alpha é calculado a partir da variância dos itens individuais e da variância da soma dos itens de cada empresa entrevistada, conforme expressão abaixo:

$$\alpha = \left(\frac{k}{k-1}\right) x \left(1 - \frac{\sum_{i=1}^{k} S_i^2}{S_t^2}\right)$$

onde,

K é o número de itens do questionário;

S_i^2 corresponde a variância de cada item;

S_t^2 corresponde a variância total do questionário, ou seja, soma das variâncias.

Matriz de correlação (*) indica correlação significativa a 5%

	Func.	Idade	Escol.	T. plan.	Exper.	Op./Nec.	Conc.	custo in.	Pn	M. prod.	Ac. fin.	Custo pr.	I. MDO	Est. conc.	Q. preço	N. tec.	I. proc.
Qtde func.	1,0000																
Idade	-0,0516	1,0000															
	0,1855																
Escolaridade	0,0386	-0,2137*	1,0000														
	0,3227	0,0000															
Tempo plan.	-0,0217	0,0344	-0,0562	1,0000													
	0,5781	0,3773	0,1496														
Expereriencia	-0,0270	0,0238	0,0326	0,1022*	1,0000												
	0,4880	0,5416	0,4025	0,0086													
Oport./Nec.	0,0627	-0,0687	0,2086*	0,0548	-0,0048	1,0000											
	0,1078	0,0780	0,0000	0,1599	0,9022												
Conh. concor.	0,0560	-0,1145*	0,1223	0,0371	0,1414*	0,1006*	1,0000										
	0,1509	0,0033	0,0016	0,3407	0,0003	0,0098											
Custo iniciais	0,0179	-0,1595*	0,1914*	0,0549	0,0745	0,0711	0,3176*	1,0000									
	0,6456	0,0000	0,0000	0,1590	0,0559	0,0683	0,0000										
Pn	-0,0176	-0,0791*	0,0916*	0,0665	0,1266*	0,0817*	0,2742*	0,4016*	1,0000								
	0,6510	0,0424	0,0186	0,0877	0,0011	0,0361	0,0000	0,0000									
M. produtos	0,0503	-0,0426	0,0626	0,0506	-0,0494	0,0821*	0,1530*	0,1528*	0,1823*	1,0000							
	0,1969	0,2746	0,1083	0,1946	0,2053	0,0352	0,0001	0,0001	0,0000								
Ac. finanças	-0,0071	-0,0205	0,0552	0,0904*	0,0278	0,0638	0,1541*	0,1888*	0,2270*	0,2015*	1,0000						
	0,8545	0,5993	0,1566	0,0202	0,4756	0,1015	0,0001	0,0000	0,0000	0,0000							
Custo prod.	0,0484	-0,0721	0,0224	0,1023*	0,0023	0,0805*	0,1662*	0,1279*	0,2635*	0,1996*	0,3759*	1,0000					
	0,2143	0,0644	0,5661	0,0086	0,9527	0,0389	0,0000	0,0010	0,0000	0,0000	0,0000						
Invest. MDO	0,0923	-0,0838*	0,1152*	0,0714	0,0562	0,0307	0,1351*	0,1484*	0,2336*	0,2993*	0,2441*	0,2612*	1,0000				
	0,0177	0,0316	0,0030	0,0669	0,1491	0,4320	0,0005	0,0001	0,0000	0,0000	0,0000	0,0000					
Estrat. conc.	0,0364	-0,0471	0,0104	0,0418	0,0634	0,0122	0,2175*	0,1432*	0,2360*	0,1880*	0,2520*	0,1376*	0,2383*	1,0000			
	0,3508	0,2272	0,7894	0,2835	0,1038	0,7549	0,0000	0,0002	0,0000	0,0000	0,0000	0,0004	0,0000				
Quali. preço	0,0404	-0,0588	-0,0363	0,0604	-0,0442	0,0590	0,1598*	0,0460	0,0658	0,1496*	0,1880*	0,2719*	0,1134*	0,1950*	1,0000		
	0,2997	0,1316	0,3514	0,1211	0,2563	0,1302	0,0000	0,2385	0,0914	0,0001	0,0000	0,0000	0,0035	0,0000			
Novas tec.	0,0523	-0,0800*	0,1781*	0,0753	0,0592	0,0409	0,1848*	0,2495*	0,2937*	0,3491*	0,2658*	0,2795*	0,3911*	0,2772*	0,2209*	1,0000	
	0,1796	0,0401	0,0000	0,0532	0,1287	0,2941	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000		
Inova proces.	0,0503	-0,0652	0,1619*	0,1233*	0,0038	0,0345	0,1233*	0,1711*	0,2283*	0,3863*	0,2561*	0,2511*	0,3251*	0,2721*	0,1967*	0,5532*	1,0000
	0,1968	0,0946	0,0000	0,0015	0,9226	0,3770	0,0015	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	

Capítulo 3 - Modelo de Riscos Proporcionais

3. Introdução

Nesse capítulo utilizamos a abordagem de riscos proporcionais para realizar os testes de sobrevivência das empresas. Também conhecido como modelo de Cox, esse modelo possibilita a estimação dos efeitos das co-variadas sobre a duração das empresas e não é necessário fazer qualquer hipótese sobre a distribuição da função densidade de probabilidade $f(t)$, definida em 1.1.

Como dito anteriormente, os três tipos de modelos de duração são complementares na análise, de tal forma que aplicaremos a modelagem de risco proporcional à base de dados utilizada no capítulo anterior para fins de comparação dos resultados. De modo geral, o presente capítulo testar os dados para o modelo de Cox e, também, segmenta a base de dados em uma perspectiva setorial, observando as principais particularidades entres os setores de comércio e serviços, principalmente àquelas ligadas a inovação[34]. Ademais, consideramos a segmentação das empresas por porte, de forma a observar quais as principais ações, por parte do empreendedor da empresa, contribuem para a sobrevivência das micro e pequenas empresas na medida em que observamos, com o passar do tempo, a possibilidade de crescimento dessas empresas, cuja relevância econômica e metodologia de classificação por tamanho foi discutida no capítulo 1.

A seguir, apresentamos o modelo de risco proporcional.

3.1. O modelo de riscos proporcionais

A diferença do modelo semi-paramétrico para os não-paramétricos é que se especifica, no primeiro, a forma funcional do modelo, ou seja, a relação das covariadas de controle com a variável de interesse (duração). Nesse caso, leva-se em conta como as variáveis ligadas às empresas influenciam na mortalidade.

A menção ao risco proporcional se deve à decomposição da função risco em dois termos: um comum a todas as empresas e outro associado às covariadas consideradas, como na função 3.1:

---

[34] De acordo com a metodologia PINTEC/IBGE inovação em processo e produto são auferíveis para os setores de comércio e industria.

$$h(t/X,\beta) = h_0(t)\exp(X'\beta)$$
$$ou \qquad (3.1)$$
$$\ln h(t) = \ln h_0(t) + X\beta$$

onde $h_0(t)$ é a função de risco *baseline hazard function*, pois $h(t/X,\beta) = h_0 t$ para $X = 0$ e X representa a influencia das covariadas sob a função risco. A forma da função risco em 3.1 leva em conta o fato de que o risco deve ser sempre positivo e as covariadas devem ser consideradas de forma a manter essa propriedade. Assim, a maneira mais intuitiva é tomando a forma exponencial ao passo que a função *baseline* é a própria função risco quando $X = 0$.

Da função 3.1, podemos assumir que $h(t/X,B)$ é a função risco em t para qualquer empresa, X é o conjunto de variáveis que afetam a probabilidade de mortalidade e β, o vetor de parâmetros associado às covariadas. A função $h_0(t)$ reflete as variações para cima ou para baixo dado as variações em X.

No caso do modelo de Cox, $h_0 t$ é a parte não-paramétrica e $\exp(X'\beta)$ é a parte paramétrica. Além disso, $h_0 t$ pode ser considerado como a função de risco para uma empresa "média" da população, em que o efeito das variáveis explicativas atua como fatores multiplicativos da função de risco da empresa "média", $h_0 t$. Ademais, não há imposição de determinar uma forma distributiva para estimar a função *baseline* e β.

O componente paramétrico é usado na forma multiplicativa, em que assume a forma:

$$g(x`\beta) = \exp\{x\beta\} = \{\beta_1 x_1 + \ldots + \beta_p x_p\}, \qquad (3.2)$$

onde $h(t/X,B)$ é sempre não-negativo.

Outra propriedade do modelo de Cox é que a razão de risco é constante. Suponhamos que x_1 e x_0 seja uma variável dicotômica, por exemplo, elaboração ou não do plano de negócios, respectivamente, teríamos que:

$$HR(t,x_1,x_2) = \frac{h(t,x_1,\beta)}{h(t,x_0,\beta)}$$
$$HR(t,x_1,x_2) = \frac{h_0(t)\exp(x_1,\beta)}{h_0(t)\exp(x_0,\beta)} = \exp((x_1 - x_0)\hat{\beta}) = \exp(\hat{\beta}) \qquad (3.3)$$

O modelo em 3.3 tem sua interpretação como razão de risco relativo, com $X = 1$(se elaborou plano de negócios) em relação a não ter elaborado. Ademais, no modelo de risco proporcional, a função sobrevivência fornece a probabilidade de uma empresa sobreviver além de certo período de tempo, geralmente, o último tempo de observação da amostra.

A função sobrevivência também leva em conta o efeito das covariadas e sobre o risco, conforme segue:

$$S(t) = \exp(-H(t)) \tag{3.6}$$

No modelo de Cox, a função sobrevivência em que $H(t, x, \beta)$ é a função de risco acumulada em *t* com covariada *x* e o tempo de sobrevivência é tomado como continuo como expresso em 3.7,

$$\begin{aligned} S(t) &= \exp[-\int_0^t h(u)du] \\ &= \exp[-\exp(X_i\beta)\int_0^t h_0(u)du] \\ &= [\exp(-\int_0^t h_0(u)du)]^{\exp(X_i\beta)} \\ &= [S_0(t)]^{\exp(X,\beta)} \end{aligned} \tag{3.7}$$

A função de sobrevivência expressa em 3.7 mede, portanto, a influência das covariadas na função sobrevivência $S_0(t)$. Temos ainda que forma da expressão em 3.7 para a função de sobrevivência deriva da relação multiplicativa entre função risco *baseline* e a função exponencial que descreve o efeito das covariadas. O valor da função sobrevivência é sempre entre zero e um.

3.2. Especificação do modelo

Nos testes para os modelos paramétricos no capítulo 2 utilizamos as propriedades do estimador de máxima verossimilhança para encontrar as variáveis de interesse β. O procedimento consistiu em maximizar a função de verossimilhança derivando a função em relação à β. Da função em 2.33, temos que:

$$l(\beta) = \prod_{i-l}^{n}\{[f(t, \beta, x)]^{c_i}.[S(t_j, \beta, x_i]^{1-c_i}\} \tag{3.8}$$

Como já sabemos que a função densidade é a razão entre a função de risco e a função de sobrevivência, podemos expressá-la como:

$$f(x,\beta,t) = h(t,x,\beta).S(t,x\beta) \tag{3.9}$$

Substituindo 3.9 em 3.1, temos:

$$l(\beta) = \prod_{i=1}^{n} \{[h(t_i, x_i, \beta)]^{c_i}.[S(t_i, x_i, \beta)]\} \tag{3.10}$$

Maximizando a função 3.10, ou seja, tomando a derivada da função de verossimilhança em relação ao parâmetro β e substituindo a função risco de 3.1 e a função sobrevivência de 3.10 na resultante da derivada, temos:

$$L(\beta) = \sum_{i=1}^{n} \{c_i \ \ln[h_0(t_i)] + c_i x_i \beta + e^{x_i \beta} \ \ln[S_0(t_i)]\} \tag{3.11}$$

Em que c_i é o indicador de falha se $c_i = 1$ ou $c_i = 0$, caso contrário. Na equação de verossimilhança em 3.14, temos a presença do termo não paramétrico h_o, o que torna esse método completo[35] inapropriado.

Em artigo de 1972, Cox propôs e depois acabou por formalizar, em artigo de 1975, o método da verossimilhança parcial[36]. Esse método depende apenas do parâmetro de interesse, β e a idéia é condicionar a função de verossimilhança ao conhecimento da história passada de falhas e censuras no intuito de eliminar o termo não paramétrico da função. Portanto, a forma funcional do modelo de verossimilhança parcial pode ser expresso como:

$$l(\beta) = \prod_{i=1}^{n} \left[\frac{e^{xi\beta}}{\sum_{jER(t_i)} e^{x_j \beta}} \right]^{c_i} \tag{3.12}$$

em que c_i é o indicador de falha conforme especificado em 3.12 e o denominador $R(t_i)$ representa todas as empresas em risco no tempo t_i[37]. Isto é, temos cada *n* (empresa) associada ao respectivo tempo de falha.

A seguir apresentamos os resultados encontrados para o modelo de risco proporcional.

---

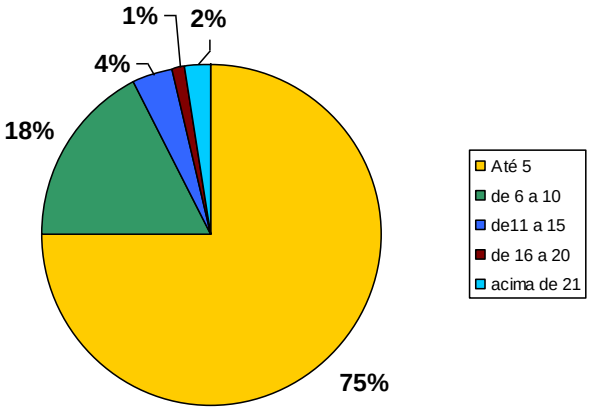
[35] Em oposição ao método da verossimilhança parcial.
[36] As propriedades assintóticas desse estimador são demonstradas em Lin and Wei (1989).
[37] Collett (1994) faz uma apresentação detalhada do modelo.

3.3. Dados

Como dissemos na introdução deste capítulo, além de aplicar a metodologia de riscos proporcionais à mesma base utilizada para os modelos do capítulo anterior, fizemos algumas segmentações na base de dados que subsidiarão o entendimento dos fatores que mais contribuem não somente para a manutenção dos pequenos negócios, mas também aqueles ligados ao setor de atuação da empresa.

O gráfico 16 mostra a divisão das 662 empresas da amostra, por faixa, em relação à quantidade de funcionários. Majoritariamente, ¾ de todas as empresas da amostra possuem atém 5 funcionários e somente 2%, mais de 20.

Gráfico 16 – Empresas por quantidade de funcionários



A tabela 20 mostra as estatísticas descritivas para as variáveis continuas da amostra.

Tabela 21 – estatísticas descritivas

	Qtde Funcionários	Idade	Tempo Planejamento
Média	5,7	39,3	8
Mediana	3	38	4
Mínimo	1	19	0
Máximo	501	72	120

O tempo de planejamento para iniciar o negócio chama a atenção pela amplitude entre os extremos das respostas, de zero a 10 anos.

Quando observamos a base sob a perspectiva setorial, as 662 empresas da amostra dividem em 447 empresas do setor de comércio, 11 da indústria e 204 empresas de serviços.

Da tabela 21 temos que em média, a quantidade de funcionários e o tempo de planejamento são maiores na indústria se comparado aos outros setores.

Tabela 22 – estatísticas sob perspectiva setorial

	Média	Mediana	Mínimo	Máximo
Qtde. funcionários (comércio)	6,2	3	1	501
Qtde. funcionários (indústria)	10,3	7	1	35
Qtde. funcionários (serviços)	4,6	3	1	48
Idade (comércio)	39,2	38	19	72
Idade (indústria)	36,1	35	22	63
Idade (serviços)	39,8	39	19	72
Tempo planej. (comércio)	8,6	4	0	120
Tempo planej. (indústria)	9	6	0	36
Tempo planej. (serviços)	8,1	4	0	120

Outra segmentação analisada nesse capítulo é a de tamanho da empresa. Para a segmentação, tomamos, por convenção, que micro empresas são aquelas com até cinco funcionários (que representam 75% da amostra), e acima disso, pequena e média empresa.

Tabela 23 – estatísticas por tamanho

	Média	Mediana	Mínimo	Máximo
Qtde. funcionários (micro)	3,4	3	1	9
Qtde. funcionários (pequena)	27,7	13	10	501
Idade (micro)	39,6	38	19	72
Idade (pequena)	36,5	35	20	64
Tempo planej. (micro)	8,5	4	0	120
Tempo planej. (pequena)	8,1	3	0	48

A informação da tabela 21 revela que a idade média do empreendedor do grupo pequena empresa é cerca de três anos a menos. Uma das possíveis explicações é que o empreendedor de uma micro empresa inicia o negócio por falta de oportunidades no mercado de trabalho. Outro ponto é que o tempo de planejamento para a pequena empresa revelou-se inferior ao da micro. Esse fato tem como uma das explicações o fato de que uma pequena empresa pode derivar de um outro negócio já estabelecido, o que reduz o tempo de tomada de decisão e mensuração dos possíveis riscos inerentes ao novo negócio por parte do empreendedor.

Na próxima seção apresentamos os resultados para os três coortes e comparamos o resultados obtidos neste capítulo com o anterior, a fim de observar quais variáveis têm maior influencia na sobrevivência das empresas.

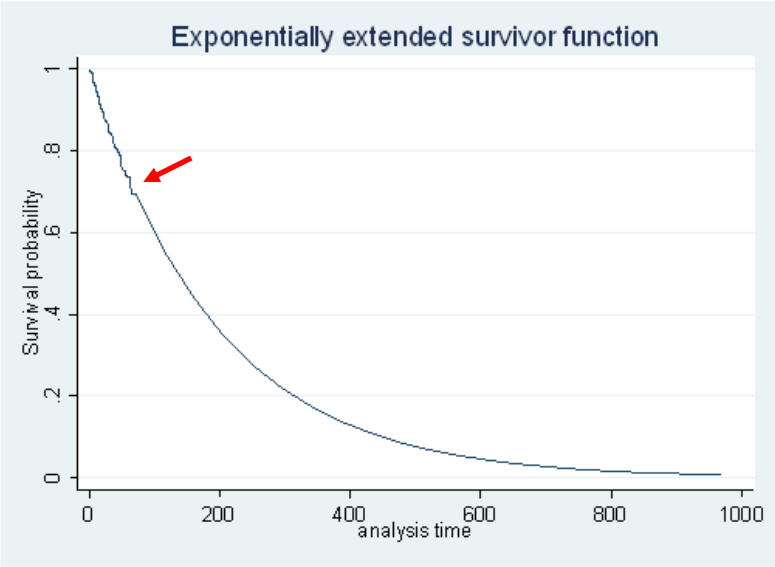
3.4. Resultados

3.4.1. Resultados para base sem segmentação

Em relação ao primeiro modelo, o qual abrange todas as 662 empresas da amostra de maneira homogênea, ou seja, sem qualquer tipo de classificação qualitativa, seja por setor ou porte.

Como mostramos no capítulo anterior, das 662 empresas da amostra, 128 falharam ao longo da janela de observação. Se todas as empresas da amostra não possuíssem qualquer característica particular, ou seja, fossem empresas idênticas, o tempo médio de sobrevivência seria de 194 meses. O gráfico 17 projeta a probabilidade de sobrevivência no tempo.

Gráfico 17 – sobrevivência projetada



Outro ponto observado dentro da análise exploratória dos dados foi a contribuição de cada empresa na composição da função risco e da função sobrevivência. Nesse sentido, o trabalho de Kalbfleisch e (2002, pg. 115) mostra que a contribuição para a função risco baseline é obtida de $h_j 1-\hat{\alpha}_j$ onde $\hat{\alpha}_j$ é a solução da expressão 3.13:

$$\sum_{kED_j} \frac{\exp(x_k \beta)}{1-\hat{\alpha}_j^{\exp(x_k \beta)}} = \sum_{lER_j} \exp(x_l \beta) \tag{3.13}$$

Para obtermos a função sobrevivência, resolvemos a equação 3.13 em que $\hat{\alpha}_o = 1$, como segue:

$$S_0(t_j) = \prod_{h=0}^{j-1} \hat{\alpha}_h \tag{3.14}$$

O resultado para todas as empresas da amostra está no Anexo 3. Entretanto, alguns deles são mostrados na tabela 23:

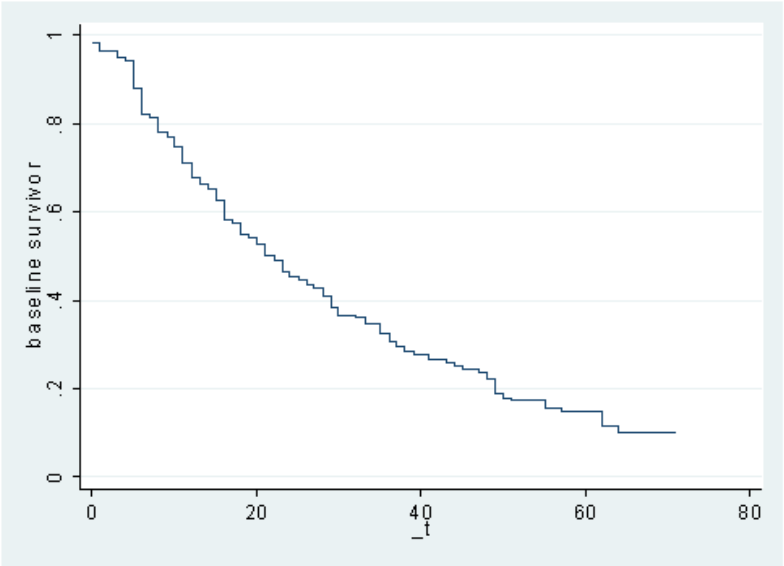
Tabela 24 – decomposição da função risco e sobrevivência

id	to	t	d	h	s
60	0	0,1	1	0,0175124	0,9824876
369	0	0,1	1	0,0175124	0,9824876
20	0	1	1	0,0175124	0,9652427
558	0	1	1	0,0175124	0,9652427
580	0	3	1	0,0176332	0,9482223

A interpretação da tabela acima é indicar que a empresa com "id" 369 falhou (ou deixou o mercado) em $t = 0{,}1$ mês. A contribuição dessa empresa para a função risco é $h369 = 0{,}1751237$ e a função sobrevivência em $t = 0{,}1$ é igual a $S(0{,}1) = 1 - 0{,}1751237 = 0{,}9824876$, como indicado na tabela 23. A próxima falha ocorre em $t = 1$ e a contribuição para a função risco nesse tempo t para a empresa id 20 é $h20 = 0{,}1763321$ e a função sobrevivência é $1 - h20 * S(0{,}1) = 0{,}9652427$. Essa interpretação é repetida para todas as empresas conforme o tempo de falha. Obviamente, caso não se verifique falhas, a função sobrevivência se mantém constante ao longo do tempo.

A 662º empresa da amostra, indicada com "id" 153, no Anexo 3 deste capítulo, apresenta função sobrevivência igual a $S(71) = 0{,}09899117$ e, partindo da tabela 3.3, encontramos o formato da curva de sobrevivência, conforme gráfico 18.

Gráfico 18 – Função sobrevivência Kaplan-Meier



Conforme exposto no item 3.2, especificação do modelo, tomamos o log da função sobrevivência e o efeito de cada co-variada sobre essa função. Adicionalmente, se a empresa falhou (deixou o mercado) a estimação também é ajustada pela função risco, no cálculo dos parâmetros do modelo. Para recuperar o valor do coeficiente, tomamos o exponencial dos resultados apresentados na tabela 24, conforme segue:

Tabela 25 – Resultados para modelo de riscos proporcionais

variável	Coeficiente	variável	Coeficiente
qtde funcionários	-0,0004 (0,0065)	Elaborou plano negócios	0,1713 (0,2093)
idade**	-0,0163 (0,0086)	procurava melhoria em produtos	-0,1128 (0,2614)
educação	0,0564 (0,0971)	acompanhava faturamento*	0,4566 (0,2075)
tempo planejamento	-0,0156 (0,0101)	acompanhava custos produtos	-0,0072 (0,2281)
experiência**	-0,3065 (0,1851)	investia na qualificação M.D.O.*	-0,5756 (0,2097)
principal motivação**	-0,3634 (0,2048)	conhecia estratégia dos concorrentes	-0,1122 (0,2048)
conhecia concorrência	-0,0363 (0,2044)	buscava qualidade e preço	0,0616 (0,2326)
conhecia custos e investimentos iniciais	-0,0572 (0,2213)	buscava atualização tecnologia	0,0352 (0,2507)
inovava em processo de trabalho	-0,2945 (0,2556)		
log likelihood	-731,37		

* significante a 5%
** significante a 10%

Cinco variáveis mostraram-se significativas, sendo elas: experiência, principal motivação, idade, acompanhamento do faturamento e investimento na qualificação, de forma que encontramos os coeficientes tomando o exponencial dos valores da tabela 24.

A interpretação é a mesma do capítulo 2, tanto para as variáveis dicotômicas como para as contínuas. No primeiro caso, tomamos o exponencial do valor da variável, subtraímos 1(um) e multiplicamos o resultado por 100 (cem) e, para o segundo, tomamos simplesmente o percentual do coeficiente calculado diretamente pelo modelo. Portanto, temos que para a variável experiência, o empreendedor que a possuía no ramo de atividade exercida, apresentava risco de falha no tempo 26,4% menor que o empreendedor que não possuía. Para idade, encontramos que 1 ano adicional na idade reduz o risco de falha, também em algum ponto no tempo, em 1,62%.

Para a variável principal motivação, encontramos que se a empresa foi criada por uma oportunidade de mercado, o risco de falha no tempo é 29,18% menor do que se aberta por necessidade. Se o empreendedor acompanhava de perto as finanças da empresa, o risco de falha da empresa no tempo mostrou-se 36,6% menor em relação àqueles que não acompanhavam. Outro resultado bastante interessante foi o investimento em qualificação da mão de obra da empresa. A empresa que investia na qualificação de seus funcionários apresentou risco de falha em algum ponto no tempo 43,8% menor em relação à empresa que não investia. A seguir apresentamos os ajustes para o modelo de riscos proporcionais.

A significância do modelo pode ser verificada por meio do teste de razão de verossimilhança[38]. O teste compara a diferença do máximo do logaritmo natural da função de verossimilhança para o modelo mais parametrizado com a do modelo mais simples. Tomando essa diferença e chamando-a de LR (*likelihood ratio*), temos que se o modelo mais simples for o de melhor ajuste ao conjunto de dados de LR, segue-se esse valor uma distribuição assintótica $\lambda_{\delta,v}{}^{2}$ com δ sendo o parâmetro de não centralidade e v a diferença de graus de liberdade entre os modelos testados.

A hipótese de que o modelo mais simples apresenta melhor ajuste se confirma se $LR > \lambda_{\delta,v}{}^{2}$. Portanto, encontramos para os modelos testados os valores de *ln1* igual a - 731,37 e *ln2* igual a -741,63, e assim temos:

$$LR = 2(741{,}63 - 731{,}37) = 19{,}6 \tag{3.15}$$

cujo valor encontrado é menor que o valor crítico $(p = 0{,}05)$ para 12 graus de liberdade. Assim, podemos aceitar a hipótese nula, ou seja, as demais co-variadas tornam-se insignificantes na presença das cinco relevantes para esse modelo.

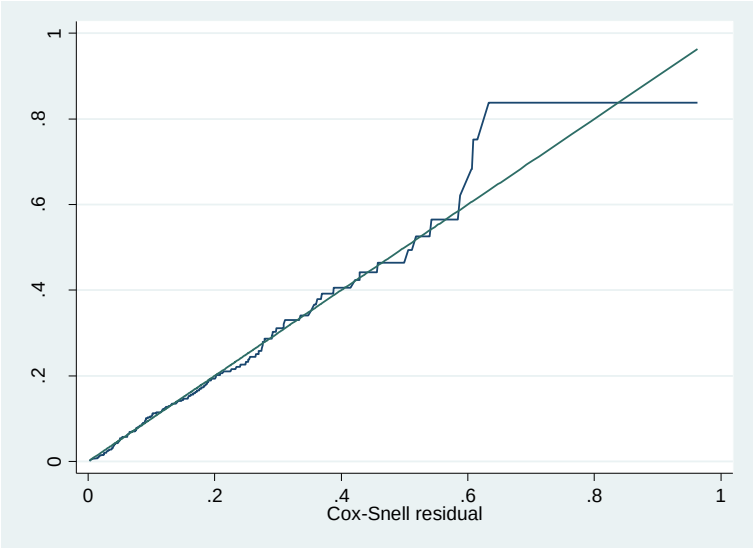
3.4.1.1. Ajuste geral do modelo

3.4.1.2. Cox-Snell

Conforme conceito apresentado no capítulo 2, a análise dos resíduos de Cox-Snell permite auferir o ajuste global do modelo. O resultado da análise está no gráfico 19:

[38] O teste de razão de verossimilhança é apropriado para testar dois modelos, desde que um dos modelos seja um caso particular do outro, também chamados de modelos aninhados.

Gráfico 19 – Ajuste geral do modelo – Cox-Snell



Os resíduos de Cox-Snell mostram que o modelo de Cox apresenta ajuste razoável aos dados de sobrevivência para o grupo de empresas da amostra. Entretanto, fatores externos ao conjunto de variáveis explicativas têm importância para uma melhor acurácia do modelo. Essas variáveis são representadas, predominantemente, por fatores políticos e econômicos.

3.4.1.3. Resíduos Martingal

Outro método de avaliação do ajuste do modelo é por meio da análise dos resíduos martingal[39].

$$rm_i = \delta_i - \hat{e}_i \tag{3.16}$$

sendo rm_i os resíduos martingal, que são resíduos com valores não negativos e que seguem uma distribuição exponencial com media 1, δ_i é um indicador de censura se igual a zero ou falha se igual a 1 e $\hat{e}_i$ são os resíduos de Cox-Snell. Essa categoria de resíduos é mais

[39] De forma mais detalhada, os cálculos dos resíduos martingal obedecem a expressão:

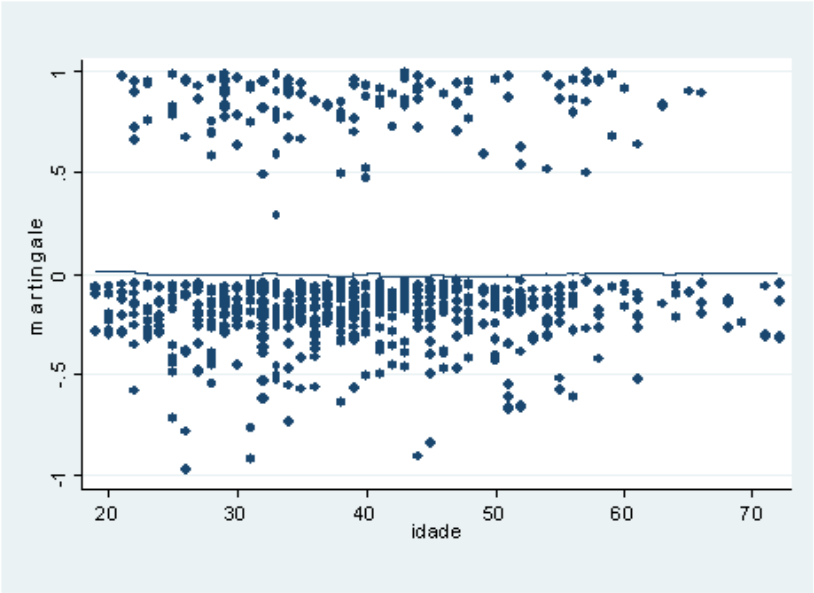
$$r_{M_i} = \delta_i - \sum_{j=1}^{D} I(t_{0i} - t_{(j)}) \frac{d_j \exp(x_i \beta)}{\sum_{lER_j} \exp(x_l \beta +)}$$, e os resíduos estão ente $(-\infty, 1]$.

δ_i é o indicador de falha se igual a 1 ou zero, caso contrário. Usando a aproximação Breslow-Peto (Breslow, 1974) e Peto (1972), o calculo dos resíduos martingal fica:

$$r_{M_i} = 1 - \exp(x_i \beta). \sum_{K:t_K \le t_i}^{D} \frac{w_k \delta_k}{\sum_{lER_j} w_l f_{ijl} \exp(x_l \beta)}.$$

adequada para se verificar a forma funcional do modelo ou até mesmo a inclusão de alguma co-variada na forma quadrática, por exemplo. No gráfico 20, construímos o gráfico dos resíduos martingal para a variável idade[40]. Como a linha plotada se comporta de maneira quase linear, nenhuma transformação nessa co-variada é necessária.

Gráfico 20 – resíduos martingal para variável idade



3.4.1.4. Resíduos deviance

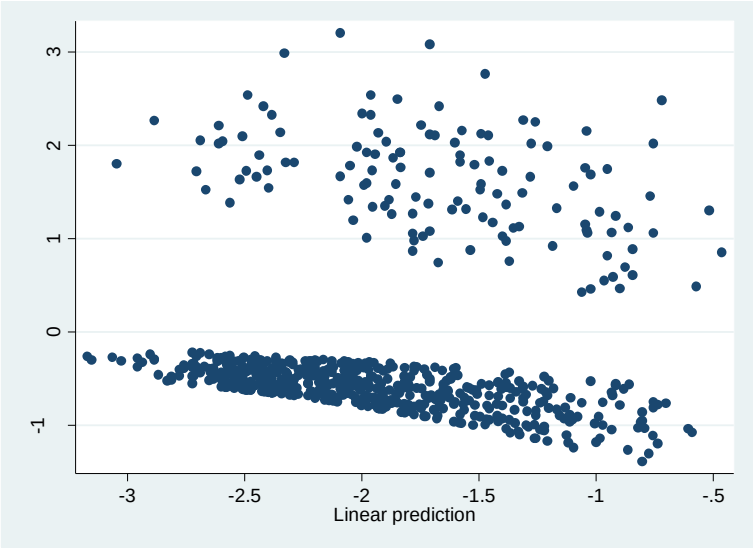
Além dos resíduos martingal, identificamos a possibilidade de pontos atípicos ou *outliers* na amostra através da analise dos resíduos *deviance*. Esses resíduos são calculados por meio da expressão 3.17 e são esperados serem simetricamente distribuídos em torno do zero, mas não necessariamente com soma zero.

$$r_{D_i} = \sin g(r_{M_i}) \left[2\{r_{M_i} + \delta_i . \log(\delta_i - r_{M_i})\} \right]^{1/2} \qquad (3.17)$$

No gráfico 21 observamos que os resíduos são “relativamente bem comportados dentro da simetria (-2,2). Verificamos que os resíduos “bem comportados” dentro do intervalo (-1,1), e alguns *outliers* além do ponto 2.

[40] Para as demais variáveis significativas para o modelo, tais como: experiência, qualificação, faturamento e principal motivação estão no Anexo 3.

Gráfico 21 – análise dos resíduos *deviance*



Outra forma de se avaliar a suposição de riscos proporcionais no modelo de Cox é pela análise dos resíduos de Schoenfeld. Esses resíduos são definidos para cada falha e não são definidos para censura.

Para cada co-variada *p* do modelo, tem-se, para o individuo *i* um correspondente resíduo de Schoenfeld e, como os resíduos são definidos em cada falha, o conjunto de resíduos é uma matriz com *d* linhas e *p* colunas, sendo *d* o número de falhas. Ainda utilizando a aproximação de Breslow-Peto, podemos expressar os resíduos Schoenfeld como:

$$r_U s_i = \sum_{j=1}^{D} \delta_{iJ} \left\{ X_{uI} - \frac{\sum_{lER_j} w_L x_{Ul} \exp(x_l \beta)}{\sum_{lER_j} w_L \exp(x_l \beta)} \right\} \tag{3.18}$$

para cada covariada, $u = 1, \ldots, p$.

Como desejável para resíduos, $\sum_i r_i = 0$ e para permitir que a estrutura de correlação dos resíduos seja considerada, utiliza-se uma forma padronizada dos resíduos de Schoenfeld:

$$s_i^* = [I(\hat{B})]^{-1}.r_i \tag{3.19}$$

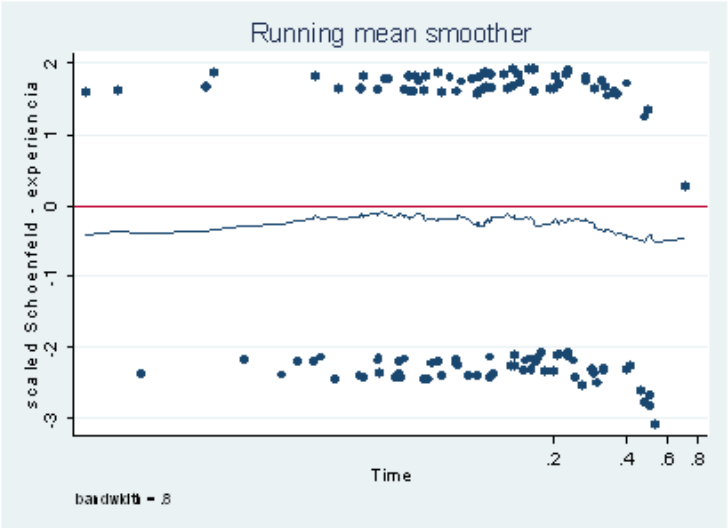
em que $I(\hat{B})$ é um matriz de informação observada. Além disso, Grambsch e Therneau (1994) consideraram o seguinte modelo:

$$\lambda(t) = \lambda_0(t) \exp\{x B(t)\} \tag{3.20}$$

a imposição de que $\beta(t) = \beta$ implica na proporcionalidade dos riscos. Se $\beta(t)$ não é constante, o impacto de uma ou mais co-variáveis no risco pode variar com o tempo. Os autores ainda sugerem que gráfico $s_i^* + \hat{B}_q$ *versus t*, para $q = 1,\ldots,p,$ de forma que a hipótese de riscos proporcionais é mantida se o resultado é uma linha horizontal.

O gráfico 22 mostra que a variável experiência atende à hipótese de proporcionalidade[41]:

Gráfico 22 – Teste de Schoenfeld - experiência



Como a análise gráfica sempre recai em alguma subjetividade, podemos utilizar o coeficiente de correlação de Pearson (*p*) entre os resíduos padronizados de Schoenfeld e *f*(t), uma função do tempo, para cada co-variável. A tabela 25 mostra os resultados para o modelo:

Tabela 26 – Schoenfeld para resíduos padronizados

	rho(*p*)	**chi2**	**df**	**Prob>chi2**
idade	-0,02339	0,08	1	0,7787
principal motivação	0,05971	0,49	1	0,4835
acompanhava faturaramento	-0,08890	0,96	1	0,3283
investia qualificação da m.d.o	0,02797	0,11	1	0,7430
possuia experiência	-0,04463	0,26	1	0,6011
teste global		**1,87**	**5**	**0,8666**

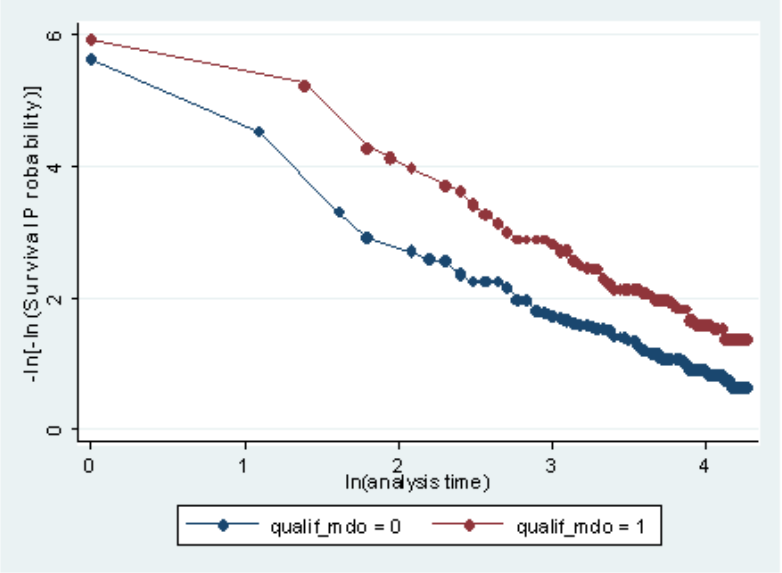
Todos os coeficientes individuais das co-variadas estão próximos de zero, o que indica que não há violação da hipótese nula de riscos proporcionais.

[41] Os testes de Schoenfeld para as outras variáveis do modelo estão no Anexo 3.

Ademais, como dissemos, no modelo de Cox a razão de risco de falha é proporcional ao longo do tempo. Assim, o risco de falha entre empresas deve ser proporcional no período total de observação dos dados. Por exemplo, se o risco de morte das empresas que investem em qualificação da mão de obra é duas vezes menor que aquelas que não investem, essa razão de risco deve ser observada em todo o período do estudo. Caso contrário, a escolha de outro modelo deve ser observada, como os paramétricos ou não paramétricos, discutidos no capítulo 2.

No gráfico 23, plotamos a curva –ln{ln(*survival*) para a variável "qualificação da mão de obra" *versus* *ln*(tempo). Se as linhas são razoavelmente paralelas, a hipótese de proporcionalidade é respeitada.

Gráfico 23 – teste proporcionalidade - qualificação M.D.O.

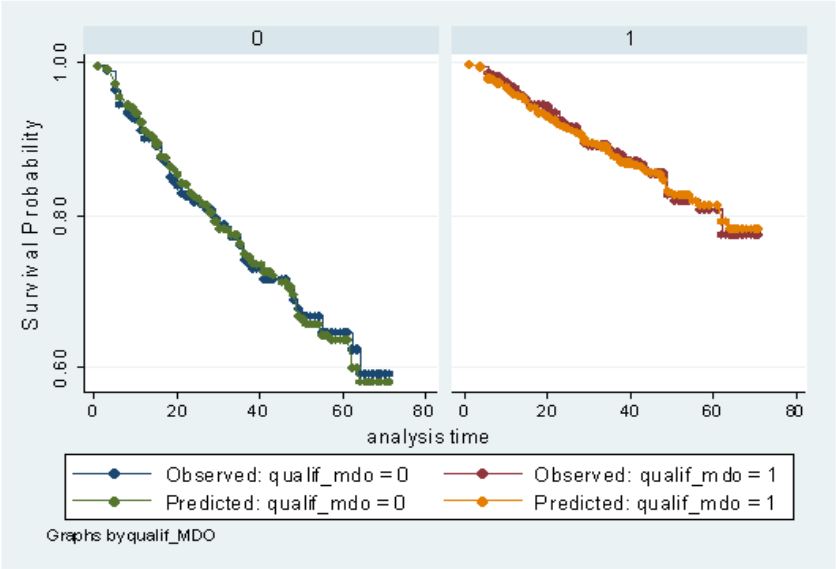


Do gráfico 23, podemos concluir que a hipótese de proporcionalidade se mantém ao longo do período de observação, para a variável mão de obra[42].

A literatura indica, também, outra forma para análise da proporcionalidade que consiste em plotar as curvas de sobrevivência de Kaplan-Meier com as curvas projetadas pelo modelo de Cox para a mesma variável. De acordo com Garret (1997), se a curva observada *versus* a prevista se apresentarem sobrepostas, a hipótese de riscos proporcionais não é violada. No gráfico 24, observamos o resultado do teste:

[42] Os testes de proporcionalidade para as outras co-variadas do modelo estão no Anexo 3.

Gráfico 24 – teste de proporcionalidade *versus* Klapan-Meier - qualificação



Podemos observar que as curvas projetadas pelo modelo de Cox e àquelas produzidas por Kaplan-Meier são, praticamente, sobrepostas. Esse resultado está alinhado com os obtidos para o modelo, confirmando, também, a hipótese da proporcionalidade.

A seguir, apresentamos os resultados para o grupo de microempresas.

3.4.2. Microempresa

Considerando a quantidade de funcionários para classificar as empresas em faixas de tamanho, selecionamos aquelas com até cinco funcionários e as denominamos microempresas[43].

Dentre os resultados descritivos, temos que a média para a variável idade é igual a quarenta anos e tempo médio de planejamento, oito meses. Os resultados para coeficientes do modelo estão na tabela 26, com respectivos desvios-padrão entre parênteses:

[43] Esse recorte difere da metodologia do SEBRAE, apresentada no capítulo 1.

Tabela 27 – Resultados da regressão para microempresas

variável	Coeficiente	variável	Coeficiente
qtde funcionários	-0,031 (0,0877)	Elaborou plano negócios**	0,4082 (0,2219)
idade**	-0,0167 (0,0092)	procurava melhoria em produtos	-0,0861 (0,2678)
educação	0,1123 (0,0112)	acompanhava faturamento	-0,3316 (0,2207)
tempo planejamento**	-0,0192 (0,0111)	acompanhava custos produtos	-0,1237 (0,2496)
experiência	-0,2594 (0,1985)	investia na qualificação M.D.O.*	-0,6068 (0,2261)
principal motivação	-0,3354 (0,2172)	conhecia estratégia dos concorrentes	-0,1698 (0,2154)
conhecia concorrência	-0,0542 (0,2123)	buscava qualidade e preço	0,0909 (0,2470)
conhecia custos e investimentos iniciais	0,0178 (0,2358)	buscava atualização tecnologia	0,0363 (0,2640)
inovava em processo de trabalho	-0,0927 (0,2743)		
log likelihood	-603,49		

* significante a 5%

** significante a 10%

Os resultados obtidos mostram três variáveis significativas, sendo elas: idade, investimento na qualificação da mão de obra, tempo de planejamento. Para a variável “tempo de planejamento” encontramos que, para cada mês adicional gasto no planejamento, o risco de falha em algum ponto no tempo é reduzido em 1,9 %.

Para a variável “idade”, verificamos que a cada ano adicional na idade do empreendedor, o risco de falha no tempo é 1,7% menor. Quanto ao investimento em “qualificação da mão de obra”, para o segmento de microempresa, essa variável se mostrou ainda mais importante se comparado à base completa mostrada no item 3.4.1, pois verificamos que as empresas que responderam positivamente a questão, apresentaram risco de falha em algum ponto no tempo 48,2% menor em relação à empresa que não investia.

Embora a variável “plano de negócios” tenha sido significativa, apresentou sinal contrário ao esperado daquele apregoado pela literatura de administração e negócios, que é justamente a eficácia da utilização dessa etapa de pré-concepção da empresa para levantamento das informações de mercado relevantes ao futuro empreendimento.

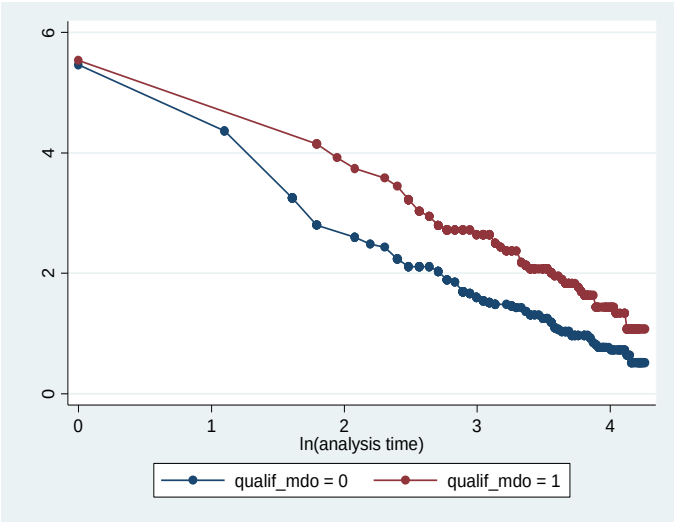
O teste de razão de verossimilhança confirma o modelo com as três variáveis significativas, pois:

$$LR = 2(615{,}32 - 603{,}49) = 23{,}66 \qquad (3.21)$$

O valor de 23,66 para razão de verossimilhança é ainda ligeiramente inferior ao valor crítico (23,685) para $(p = 0{,}05)$ e 14 graus de liberdade. Assim, podemos aceitar a hipótese nula, de que as demais co-variadas tornam-se insignificantes na presença das cinco relevantes para esse modelo.

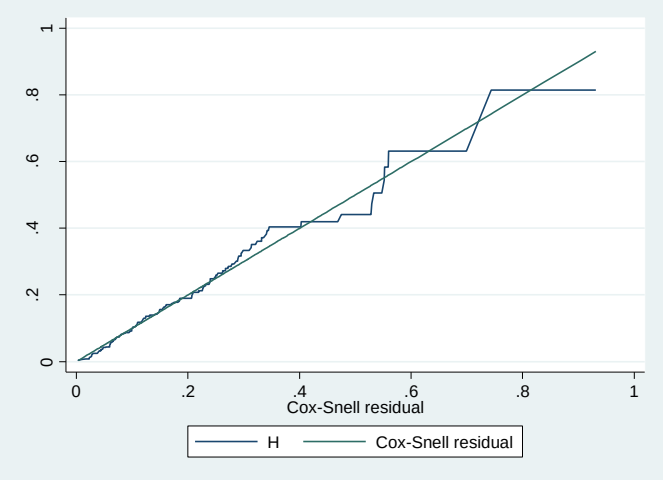
O teste de proporcionalidade para a variável "qualificação da mão de obra" respeita a hipótese de proporcionalidade, como mostra o gráfico 25.

Gráfico 25 - qualificação mão de obra



Os resíduos de Cox-Snell mostram que o modelo de Cox apresenta razoável ajuste aos dados de sobrevivência para o grupo de empresas da amostra, conforme gráfico 26:

Gráfico 26 – Ajuste Cox-Snell



Os gráficos 27 e 28 são os resíduos de martingal para as variáveis “idade” e “tempo de planejamento” para o modelo de microempresas. Pela análise dos gráficos, nenhuma transformação nas variáveis é necessária.

Gráfico 27 – martingal idade

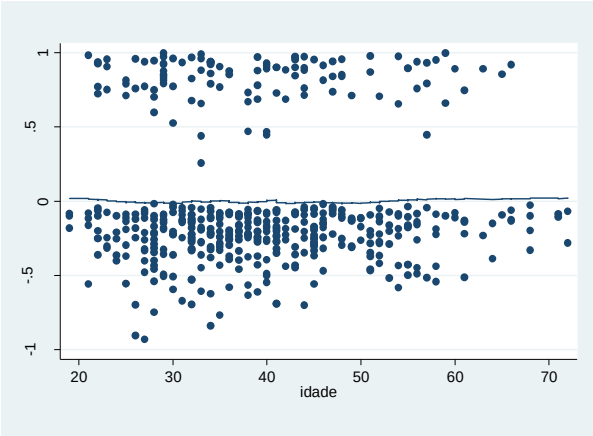
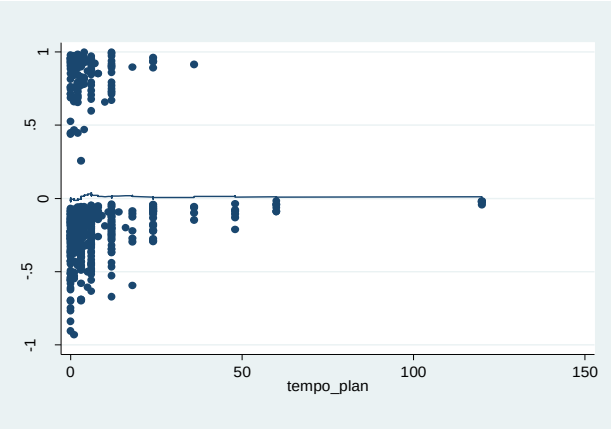


Gráfico 28 - martingal tempo planej.



Os coeficientes de correlação de Pearson (*p*) entre os resíduos padronizados de Schoenfeld contra uma função do tempo *f*(t) para cada co-variável estão na tabela 27:

Tabela 28 – Schoenfeld para resíduos padronizados

	rho(*p*)	**chi2**	**df**	**Prob>chi2**
idade	0,10007	1,18	1	0,2764
tempo de planejamento	0,05316	0,16	1	0,6880
investia qualificação da m.d.o	0,11291	1,39	1	0,2392
teste global		**2,69**	**3**	**0,4423**

Todos os coeficientes individuais das co-variadas relevantes do modelo estão próximos do zero, indicando a não violação da hipótese de riscos proporcionais.

Logo abaixo, apresentamos os gráficos com os testes de proporcionalidade para as variáveis do modelo. Novamente, não há violação da hipótese de riscos proporcionais.

Gráfico. 29 – Teste Schoenfeld - experiência

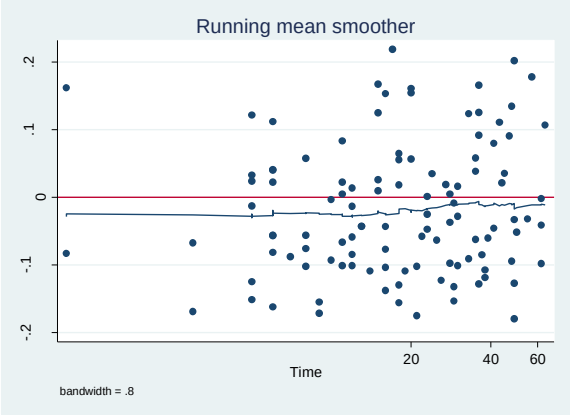
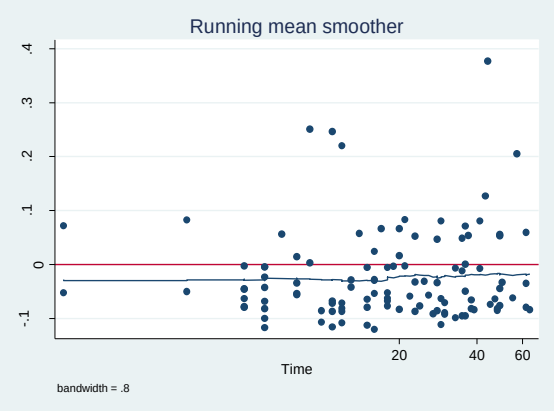


Gráfico. 30 – Teste Schoenfeld - tempo planejam.



A seguir temos os resultados para as pequenas e médias empresas.

3.4.3. Pequena e média empresa

Com o intuito de verificar as particularidades entres os grupos de empresas no tocante à sobrevivência, selecionamos as empresas com mais de cinco funcionários e denominamos o grupo resultante de pequena e média empresa (PME).

À medida que trabalhamos a base de dados, os resultados se tornam mais característicos para os grupos, conforme estudos citados na revisão da literatura discutida desde o primeiro capítulo. No caso das pequenas e médias, outras variáveis se mostraram relevantes para a longevidade da empresa quando comparado à primeira análise com a base completa e a segunda, somente com as microempresas.

No caso das PME, destacamos o fato da elaboração do plano de negócios para a constituição da empresa. A empresa planejada na estrutura dos requisitos de um plano de negócios apresentou redução da probabilidade de falha em 82,9% em relação à empresa criada sem esse planejamento. Outro ponto de destaque é a inovação em processos de trabalho. A empresa que a praticava apresentou redução do risco de falha no tempo em 74,5% em comparação à empresa que não inovava.

A tabela 28 apresenta os resultados da regressão de Cox, com os desvios-padrão entre parênteses[44]:

[44] Vale lembrar que das 11 empresas do setor da indústria, 6 estão na base de pequenas e médias empresas e 5 foram classificadas como microempresas.

Tabela 29 – Resultados da regressão para PME

variável	Coeficiente	variável	Coeficiente
qtde funcionários	0,0013 (0,0039)	Elaborou plano negócios*	-1,7675 (0,6738)
idade	-0,0476 (0,0299)	procurava melhoria em produtos	-1,3509 (1,3103)
educação	-0,5666 (0,3513)	acompanhava faturamento	-0,3642 (0,7510)
tempo planejamento	0,0169 (0,0201)	acompanhava custos produtos	-1,2432 (0,6792)
experiência	-0,8232 (0,6289)	investia na qualificação M.D.O.	0,2412 (0,7284)
principal motivação	-0,026 (0,7889)	conhecia estratégia dos concorrentes	1,3864 (0,8336)
conhecia concorrência	1,3028 (0,8235)	buscava qualidade e preço	-0,2849 (0,9623)
conhecia custos e investimentos iniciais	-1,5071 (0,6823)	buscava atualização tecnologia	0,693 (0,9940)
inovava em processo de trabalho*	-1,3673 (0,8016)		
log likelihood	-59,995		

* significante a 5%

** significante a 10%

A acurácia do modelo pode ser medida pelo teste de razão de verossimilhança, que confirma o modelo para as variáveis significativas, pois:

$$LR = 2(69{,}65 - 59{,}99) = 19{,}32 \tag{3.22}$$

O valor de 19,32 para razão de verossimilhança é ainda ligeiramente inferior ao valor crítico (24,99) para $(p = 0{,}05)$ e 15 graus de liberdade. Assim, podemos aceitar a hipótese nula, de que as demais co-variadas tornam-se insignificantes na presença das duas relevantes para esse modelo.

Como a hipótese de proporcionalidade é uma premissa do modelo, os gráficos 31 e 32 apresentam os gráficos o resultado do teste para as variáveis “plano de negócio” e “inova em processo”, nos quais podemos observar que as variáveis estão de acordo com esse requisito.

Gráfico 31 – inova em processos

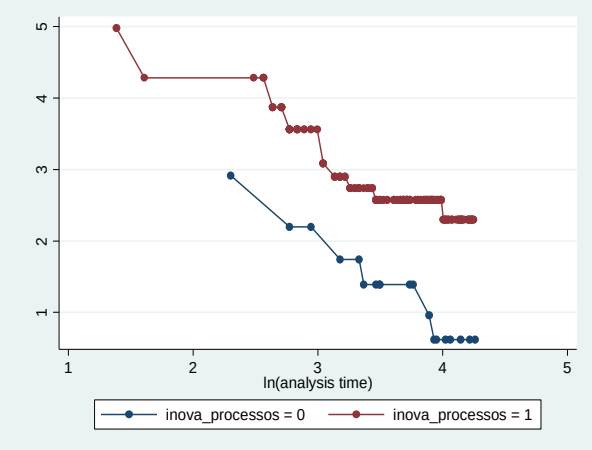
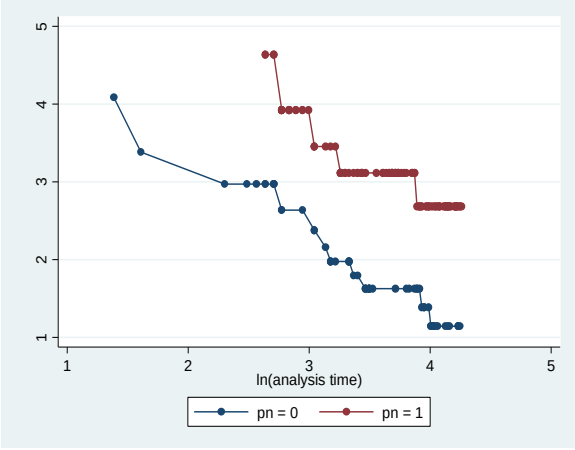
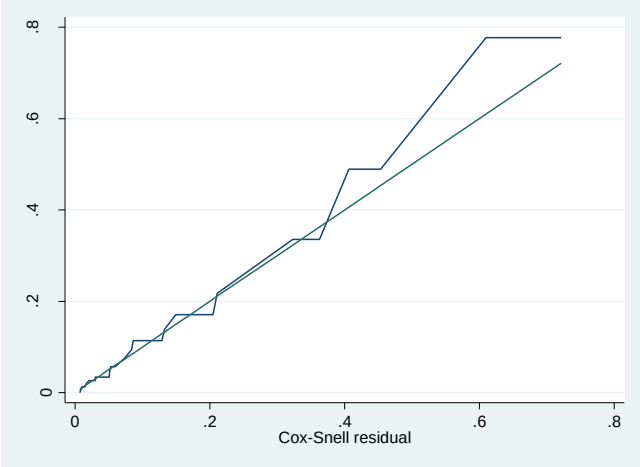


Gráfico 32 – plano de negócio



Além da hipótese de proporcionalidade, o ajuste do modelo pode ser checado por meio dos resíduos de Cox-Snell, como no gráfico 33 Apesar de o modelo ser robusto para as duas variáveis, o ajuste geral do modelo não é capaz, exclusivamente, de explicar a totalidade da longevidade para o grupo de empresas com mais de cinco funcionários, ou seja, além da elaboração de um plano de negócios e de inovar em processos, há outros fatores externos que interferem nessas estruturas, tais como, ambiente macroeconômico, política fiscal, dentre outras.

Gráfico 33 – Ajuste Cox-Snell



Os gráficos 34 e 35 são os resíduos de martingal para as variáveis “inovação em processos” e “elaboração de plano de negócios” para o modelo de microempresas. As variáveis atendem a condição de linearidade e concluímos que nenhuma transformação nessas variáveis é necessária.

Gráfico. 34 – martingal inov. processo

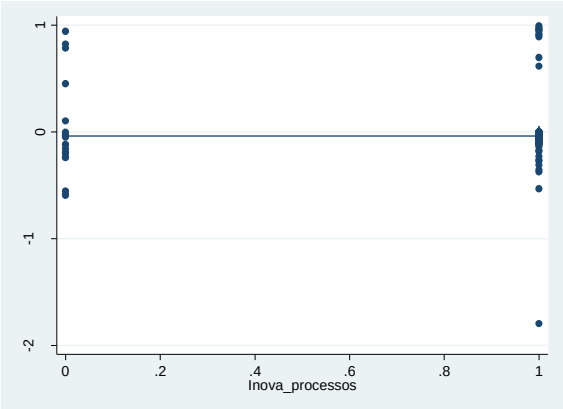
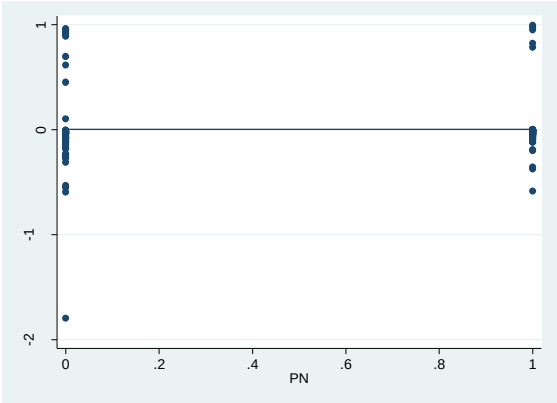


Gráfico. 35 - martingal plano negócio



Os coeficientes de correlação de Pearson (*p*) entre os resíduos padronizados de Schoenfeld contra uma função do tempo *f*(t) para cada co-variável estão na tabela 29:

Tabela 30 – Schoenfeld para resíduos padronizados

	rho(*p*)	chi2	df	Prob>chi2
plano de negócio	0,00928	0	1	0,9694
inovação em processos	-0,22234	0,83	1	0,3629
teste global		**0,84**	**2**	**0,6562**

Todos os coeficientes individuais das co-variadas relevantes para o modelo estão próximos do zero, indicando a não violação da hipótese de riscos proporcionais.

Apresentamos, também, os gráficos com os testes de proporcionalidade para as variáveis do modelo. Verificamos que, novamente, não há violação da hipótese de riscos proporcionais.

Gráfico 36 - Teste Schoenfeld - plano negócio

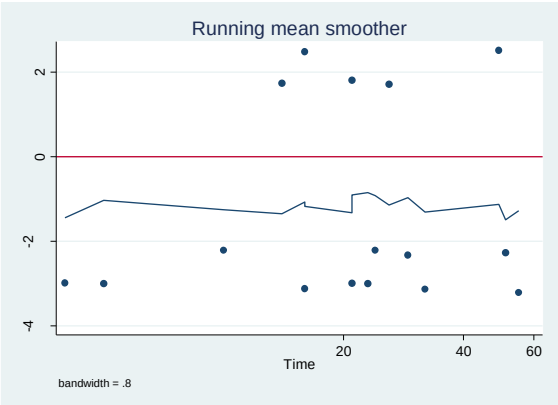
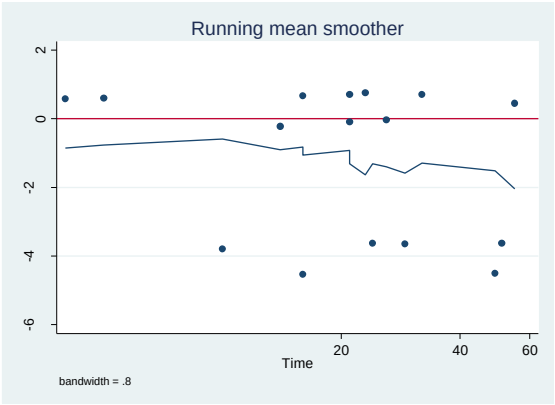


Gráfico. 37 - Teste Schoenfeld –inova processo



Na próxima seção apresentamos os resultados encontrados para as empresas do setor de comércio.

3.4.4. Setor de comércio

Das 662 empresas da base de dados, 447 ou 67,5% pertencem ao setor de comércio. Com tamanho médio de 6,2 funcionários e mediana 3, conforme exposto no item 3.1, os resultados encontrados para esse segmento mostraram-se dentro do esperado para o setor e classificação de tamanho dessas empresas, principalmente pela ausência de significância de variáveis ligadas ao planejamento, mas também de atributos ligados a inovação, idade e experiência.

O destaque maior e único, para o setor, é a prática do acompanhamento do faturamento da empresa por parte do empreendedor. Pelo resultado direto do coeficiente, temos que o valor negativo evidencia que o acompanhamento é uma boa prática em relação à longevidade do negócio, ou seja, praticá-lo, que no questionário tem resposta positiva e igual a 1(um), aumenta a tempo de vida do empreendimento. Ademais, tomando o exponencial encontramos que quem “acompanhava o faturamento”, finanças de um modo geral, apresentou chance de falha ao longo do tempo 51,1% menor em relação a quem não acompanhava.

A tabela 30 mostra os resultados para todas as co-variadas utilizadas na regressão para o modelo de comércio.

Tabela 31 – Resultados da regressão para setor de comércio

variável	Coeficiente	variável	Coeficiente
qtde funcionários	0,0009 (0,0045)	Elaborou plano negócios	0,0522 (0,2682)
idade	-0,0059 (0,0105)	procurava melhoria em produtos	-0,023 (0,3605)
educação	0,0773 (0,1228)	acompanhava faturamento*	-0,7164 (0,2552)
tempo planejamento	-0,0128 (0,0123)	acompanhava custos produtos	0,2185 (0,2999)
experiência	-0,2776 (0,2314)	investia na qualificação M.D.O.	-0,3294 (0,2561)
principal motivação	-0,3001 (0,2730)	conhecia estratégia dos concorrentes	-0,3463 (0,2560)
conhecia concorrência	0,194 (0,2606)	buscava qualidade e preço	-0,0999 (0,2995)
conhecia custos e investimentos iniciais	0,0581 (0,2811)	buscava atualização tecnologia	-0,1309 (0,3254)
inovava em processo de trabalho	-0,1012 (0,3360)		
log likelihood	-448,51		

* significante a 5%

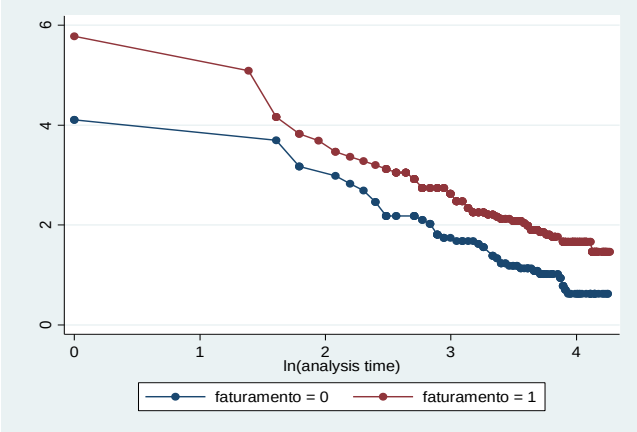
** significante a 10%

Como única variável significativa, aplicamos o teste de razão de verossimilhança (TRV) entre os modelos completo, com todas as 17 co-variadas, e aquele somente com a co-variada e o resultado obtido não rejeitou a hipótese nula, ou seja, as demais co-variadas não são significantes na presença da variável "acompanhamento do faturamento". O valor encontrado, embora muito próximo do valor crítico (26,299) para $(p = 0{,}05)$ e 16 graus de liberdade, é:

$$LR = 2(461{,}62 - 448{,}51) = 26{,}22 \tag{3.23}$$

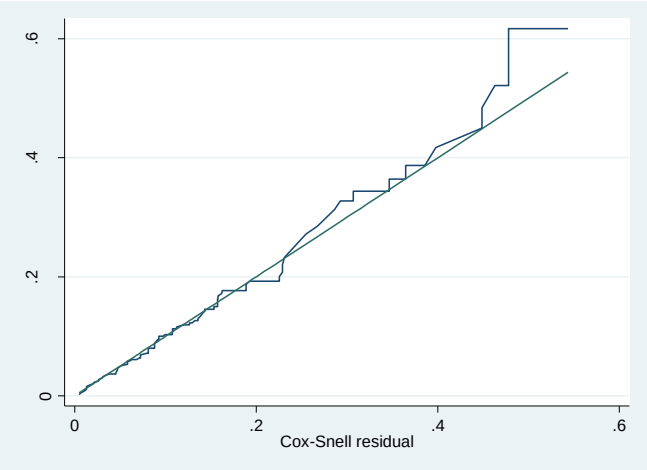
O teste de proporcionalidade mostra que a co-variada "acompanhamento do faturamento", no gráfico 38, respeita a hipótese de proporcionalidade.

Gráfico 38 – Faturamento



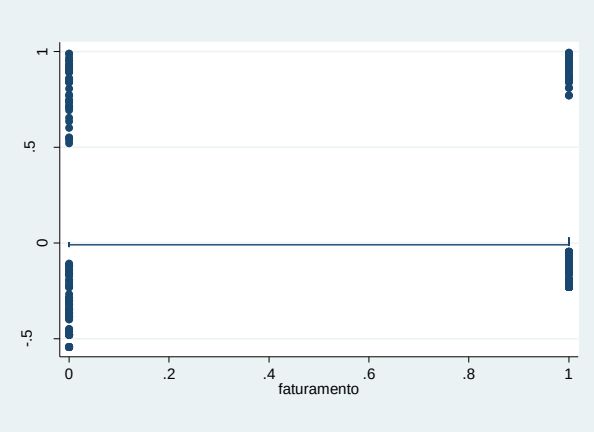
O teste preditivo do modelo por meio do teste dos resíduos de Cox-Snell está no gráfico 39 O modelo apresenta razoável ajuste do ponto de vista estatístico e, de forma geral, é capaz de explicar, em grande parte, a mortalidade das empresas do setor de comércio por meio da variável "acompanhamento de faturamento". Entretanto, fatores externos como os econômicos (juros, política fiscal, tributação, etc.) e/ou políticos influenciam a sobrevivência desse grupo de empresas.

Gráfico 39 – Ajuste Cox-Snell



O gráfico 40 apresenta os resíduos de martingal para a variável "acompanhamento do faturamento" para o modelo de comércio. A variável atende a condição de linearidade e concluímos que nenhuma transformação nessa variável é necessária.

Gráfico 40 – martingal “Faturamento”



Os coeficientes de correlação de Pearson (*p*) entre os resíduos padronizados de Schoenfeld contra uma função do tempo *f*(t) para a co-variada faturamento está na tabela 31:

Tabela 32 – Schoenfeld para resíduos padronizados

	rho(*p*)	**chi2**	**df**	**Prob>chi2**
faturamento	-0,00031	0	1	0,9977
teste global		**0**	**1**	**0,9977**

O coeficiente para a co-variada “acompanhamento do faturamento” está próximo de zero, portanto, não há evidencias da violação da hipótese de riscos proporcionais.

O gráfico 41 mostra o resultado do teste de proporcionalidade de Schoenfeld para a variável “acompanhamento do faturamento”. Constamos que não há violação da hipótese de riscos proporcionais.

Gráfico 41 - Teste Schoenfeld – “Faturamento”

Na seção 3.4.5 a seguir, apresentamos os resultados para as empresas do setor de serviços.

3.4.5. Setor de serviços

De acordo com a PINTEC/IBGE, os setores de serviços e indústria apresentam indicadores mensuráveis em relação à inovação, seja em novos produtos e/ou processos. No entanto, nossa base de dados não nos permite explorar os resultados para um coorte exclusivo às empresas do setor da indústria que, como mencionado, possui pouca representatividade na amostra.

Já para as empresas do setor de serviços, as análises dos resultados se mostraram diferem daqueles obtidos para aquelas do setor de comércio, como veremos a seguir.

No conjunto das empresas da amostra, as classificadas como pertencentes ao setor de serviços representam 31% das empresas da amostra e, como destacamos no item 3.1, possuem a maior média para a variável idade do empreendedor/proprietário que os demais setores (comércio e indústria). Os testes para o modelo de Cox se mostraram significativos para cinco co-variadas, indicadas na tabela 32.

Em relação à quantidade de funcionários, concluímos que essa co-variada mostrou-se significativa de tal forma que a adição de um funcionário no quadro de colaboradores da empresa representa redução de 14,7% na probabilidade de falha em algum ponto do tempo. A co-variada idade também foi significativa nos testes sendo que um ano adicional na idade do empreendedor da empresa representa redução em 3,5% na chance de falha no tempo.

O investimento em mão de obra também se mostrou significante e a empresa que investia na qualificação de seus colaboradores apresentou redução na probabilidade de falha em 62,7% em relação às empresas que não investiam.

A tabela 32 apresenta os resultados para o modelo do setor de serviços.

Tabela 33 – Resultados da regressão para setor de serviços

variável	Coeficiente	variável	Coeficiente
qtde funcionários*	-0,1589 (0,0756)	Elaborou plano negócios	0,2018 (0,3623)
idade*	-0,0398 (0,0165)	procurava melhoria em produtos	-0,3802 (0,4215)
educação	-0,1489 (0,1711)	acompanhava faturamento	0,1542 (0,3737)
tempo planejamento	0,0215 (0,1933)	acompanhava custos produtos	0,1106 (0,4016)
experiência	-0,5628 (0,3288)	investia na qualificação M.D.O.*	-1,2102 (0,4045)
principal motivação	-0,6157 (0,3563)	conhecia estratégia dos concorrentes	0,268 (0,3793)
conhecia concorrência	0,4502 (0,3793)	buscava qualidade e preço	0,2395 (0,3876)
conhecia custos e investimentos iniciais	0,0163 (0,3956)	buscava atualização tecnologia	0,5837 (0,4077)
inovava em processo de trabalho	-0,6757 (0,4186)		
log likelihood	-194,33		

* significante a 5%

** significante a 10%

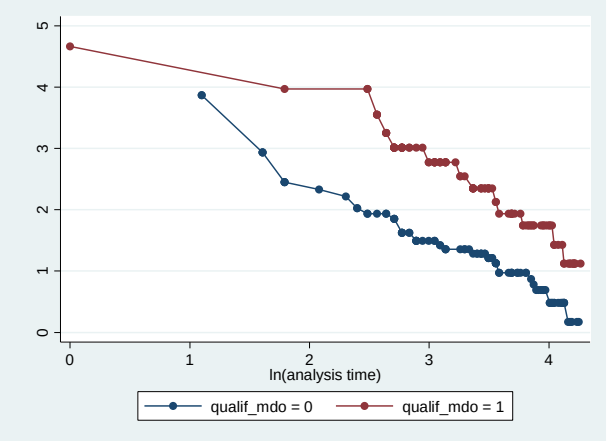
A seguir, a equação 3.28 apresenta o teste de razão de verossimilhança para o modelo:

$$LR = 2(201{,}51 - 193{,}33) = 14{,}36 \tag{3.24}$$

O valor de 19,32 para razão de verossimilhança é inferior ao valor crítico (23,68) para $(p = 0{,}05)$ e 14 graus de liberdade. Assim, podemos aceitar a hipótese nula, de que as demais co-variadas tornam-se insignificantes na presença das duas relevantes para esse modelo.

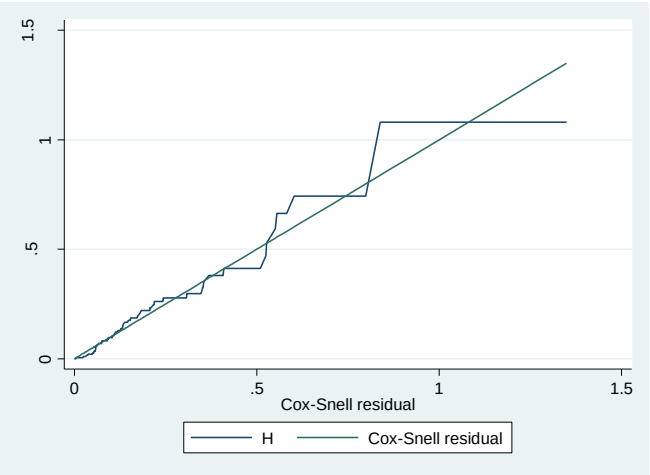
Como a hipótese de proporcionalidade é uma premissa do modelo, o gráfico 3.4.5.1 mostra o resultado do teste de proporcionalidade para a variável dicotômica “qualificação da mão de obra”, em que a hipótese de proporcionalidade não é violada.

Gráfico 42 – Qualificação da mão de obra



Além da hipótese de proporcionalidade, o ajuste do modelo pode ser checado por meio dos resíduos de Cox-Snell, como no gráfico 43. Novamente, o modelo não é totalmente adequado para explicar a projeção da longevidade para esse grupo de empresas, que deve a fatores externos, como os já mencionados anteriormente, a completude da explicação e o ajuste ideal.

Gráfico 43 – Ajuste Cox-Snell



Os gráficos 44 e 45 e 46 são os resíduos de martingal para as variáveis “quantidade de funcionários”, “idade” e “qualificação da mão de obra” para o modelo de serviços. As variáveis atendem a condição de linearidade e concluímos que nenhuma transformação nessas variáveis é necessária.

Gráfico. 44 – martingal qtde. funcionários

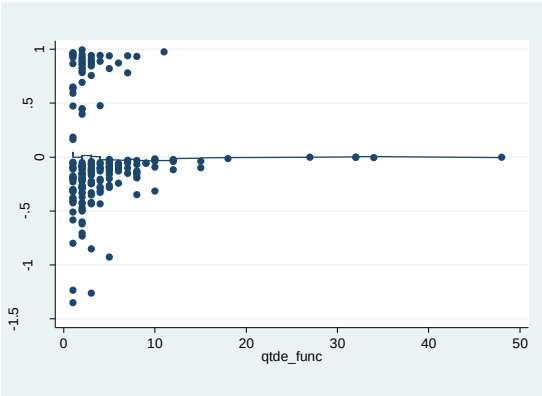


Gráfico 45 - martingal idade

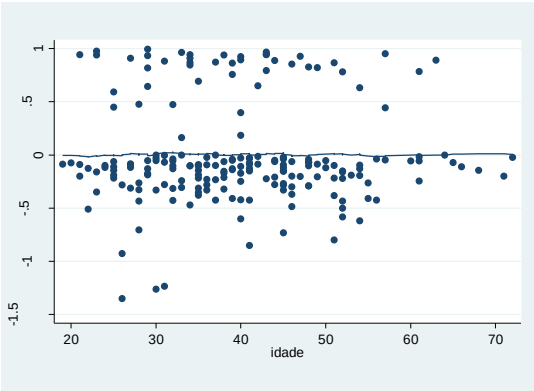
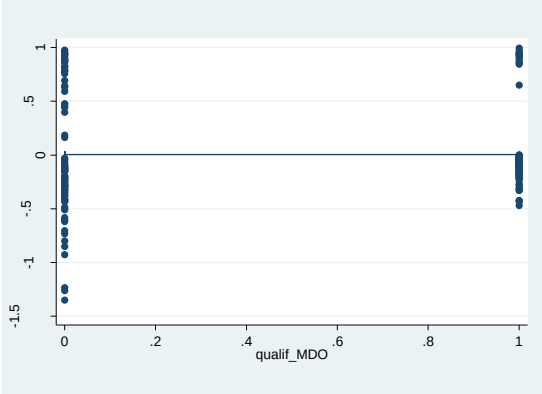


Gráfico 46 – martingal qualif. mdo



Os coeficientes de correlação de Pearson (*p*) entre os resíduos padronizados de Schoenfeld contra uma função do tempo *f*(t) para cada co-variável estão na tabela 33:

Tabela 34 – Schoenfeld para resíduos padronizados

	rho(*p*)	**chi2**	**df**	**Prob>chi2**
quantidade de funcionários	-0,05857	0,16	1	0,6935
idade	0,41390	7,33	1	0,0068
qualificação MDO	-0,22234	1,83	1	0,1756
teste global	**0,1895**	**0,84**	**3**	**0.0376**

Pelo resultado acima, optamos por excluir a variável “idade” do modelo, permanecendo as demais, que apresentaram coeficientes próximos de zero, indicando a não violação da hipótese de riscos proporcionais. Os gráficos que corroboram o resultado da tabela estão logo abaixo:

Gráfico 47 – Teste Schoenfeld – idade

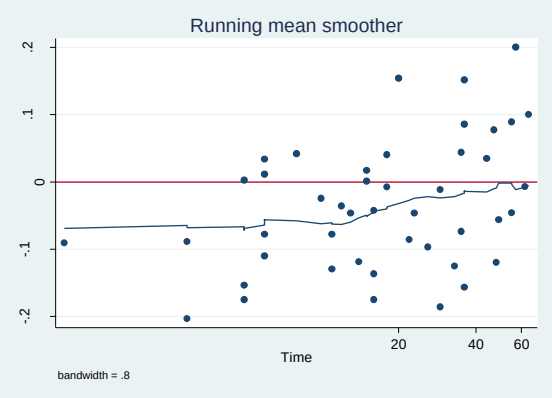


Gráfico 48 - Teste Schoenfeld –qtde funcionários

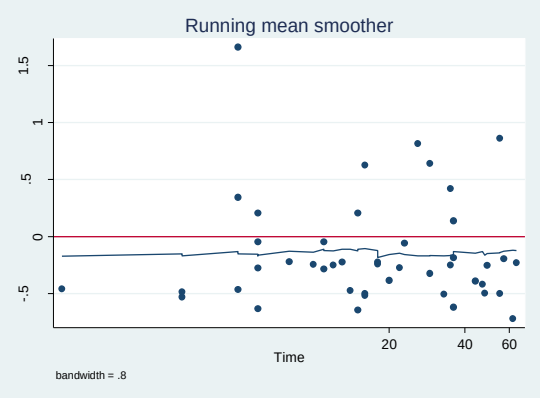
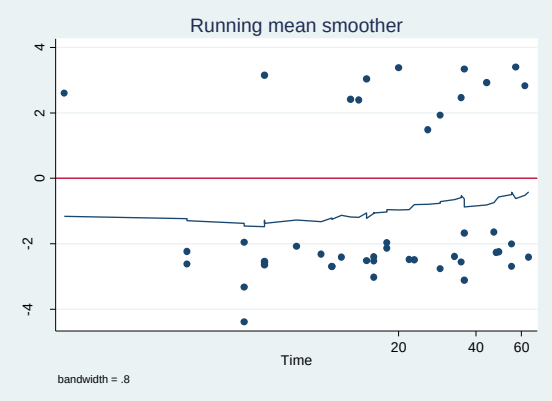


Gráfico 49 – Teste Schoenfeld - qualif. mdo



A seguir, apresentamos a conclusão do presente capítulo.

3.5. Conclusão

Neste capítulo apresentamos os resultados para o modelo de riscos proporcionais, também conhecido na literatura como modelo de Cox. Além dos testes de riscos proporcionais para a mesma base de dados do capitulo 2, segmentamos a amostra por setor, comércio e serviços[45], e também por tamanho em relação à quantidade de funcionários, sendo que denominamos microempresa as empresas até com cinco funcionários e pequena e média empresas aquelas com mais de cinco funcionários.

Os resultados obtidos para todos os modelos de riscos proporcionais testados estão na tabela 4.0. Os sinais negativos dos percentuais indicam o percentual de redução de falha da empresa em algum momento do tempo:

[45] Indústria não foi considerada devido à baixa representatividade das empresas do setor na amostra.

Tabela 35 – Resultados dos modelos

Variáveis - Variáveis por Modelo									
Modelo	Idade	Exper.	Princ. Motivac.	Ac. Finanças	Qualif. M.D.O.	Inovação processos	Tempo planej.	Plano negoc.	Qtde Func.
Completo	-1,6%	-26,4%	-29,2%	-36,6%	-43,8%	-	-	-	-
Micro empresa	-1,7%	-	-	-	-48,2%	-	-1,9%	-	-
Pequena empresa	-	-	-	-	-	-74,5%	-	-82,9%	-
Comércio	-	-	-	-51,1%	-	-	-	-	
Serviços	-3,5%	-	-	-	-62,7%	-	-	-	-14,7%

Os resultados mostraram que um único modelo é menos eficiente para explicar a sobrevivência de um grupo de empresas. Os modelos aplicados a determinado setor ou porte apresentam resultados mais direcionados (sem ruídos) e mais aderentes à literatura. Nesse sentido, por exemplo, era esperado que o setor de comércio apresentasse baixo ou nenhuma inclinação em relação às atividades ligadas a inovação do mesmo modo que uma microempresa fosse constituída pela prévia elaboração de um plano de negócios.

Destacamos ainda, a importância do investimento em qualificação dos funcionários como um dos pilares de sustentação de uma empresa ao longo do tempo. Isso como conseqüência remete às maiores produtividade e qualidade dos processos, sejam eles ligados à produção ou não.

Para a pequena empresa, a elaboração do plano de negócios suplanta a importância da variável idade assim como no setor de comércio, em que o tamanho dos empreendimentos da amostra parece pouco suscetível ao fator idade e a eventual experiência decorrente.

Ademais, os modelos mesmos que segmentados sofrem influência de variáveis exógenas a eles, principalmente, as ligação ao cenário econômico, como taxa de juros, crédito, câmbio, etc. além da política fiscal (tributária).

O capítulo quatro abrange os principais resultados deste trabalho, suas limitações e propõe diferentes abordagens para trabalhos futuros.

Anexo 3

O anexo apresenta os resultados dos testes de proporcionalidade para o modelo completo por meio da análise dos resíduos de Schoenfeld e ajuste pelas curvas de Kaplan-Meier, além das observações dos resíduos martingal.

Testes de Schoenfeld para modelo completo

Gráfico 50 - Schoenfeld - idade

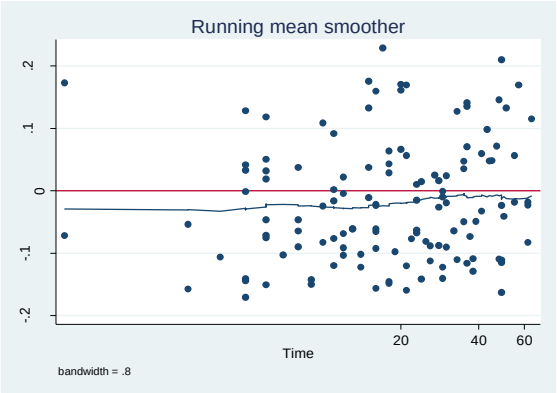


Gráfico 51 - Schoenfeld – principal motivação

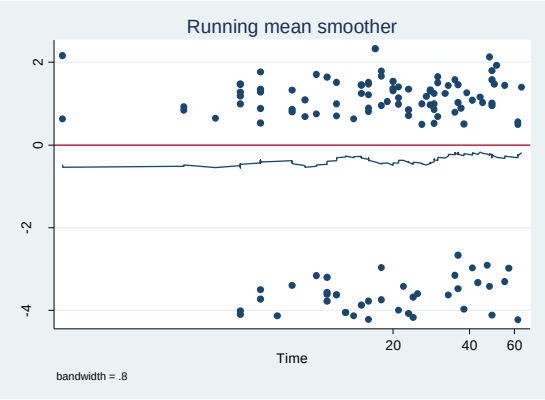


Gráfico 52 - Schoenfeld - experiência

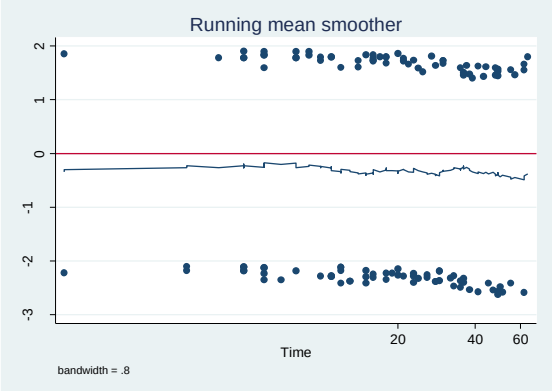


Gráfico 53 - Schoenfeld - faturamento

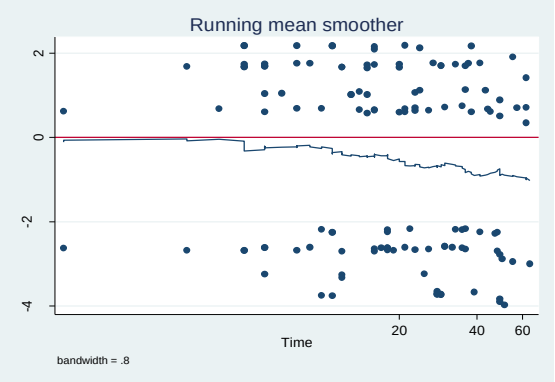
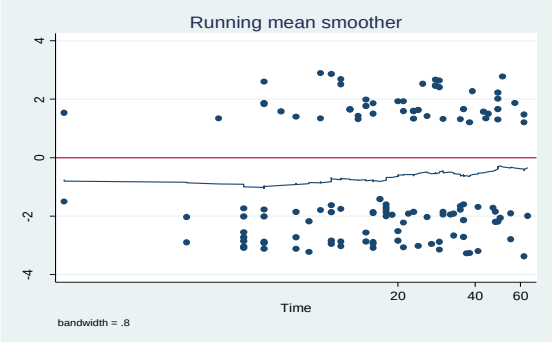


Gráfico 54 - Schoenfeld – qualificação M.D.O



Testes de Proporcionalidade das co-variadas

Gráfico 55 – K-M – p. motivação

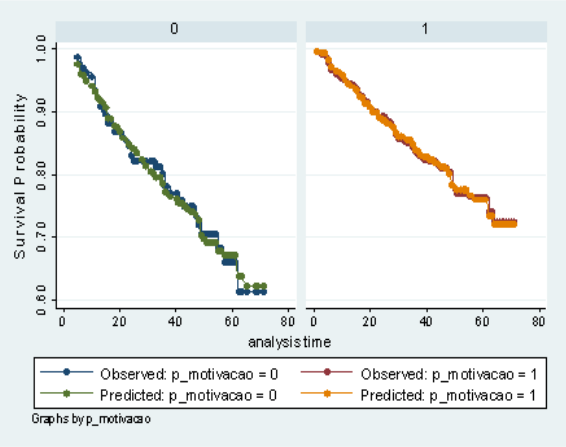


Gráfico 56 - K-M – experiência

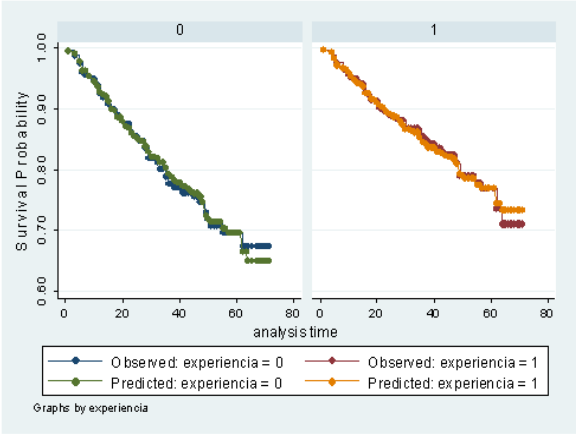
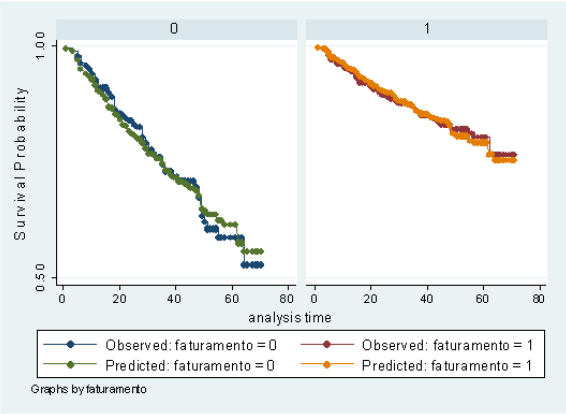


Gráfico 57 - K-M – faturamento



Testes dos resíduos Martingal

Gráfico 58 - Martingal – Qualif. M.D.O

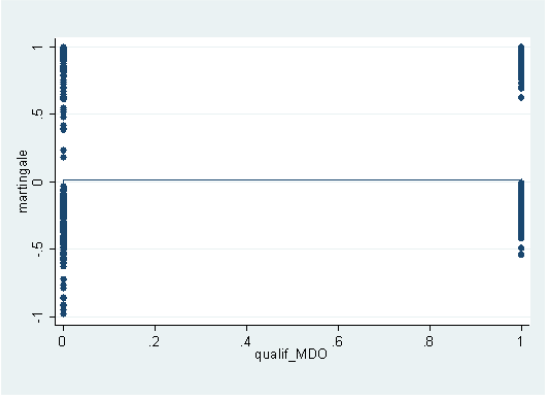


Gráfico 59 - Martingal - principal motivação

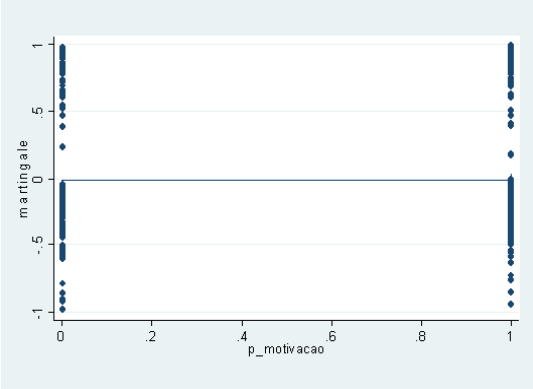


Gráfico 60 - Martingal - experiência

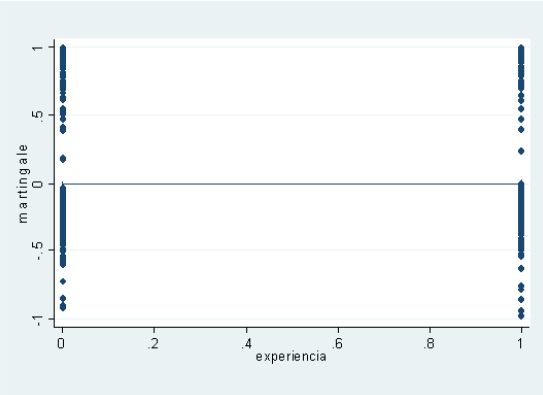
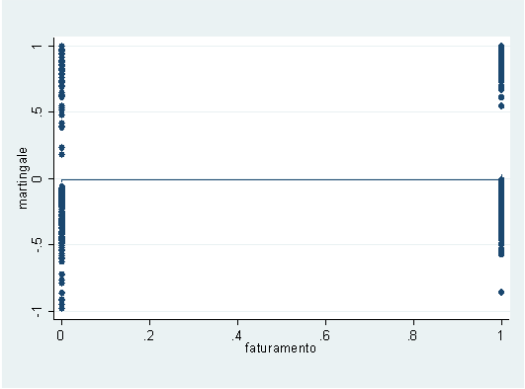


Gráfico 61 - Martingal – faturamento



Testes de Proporcionalidade das co-variadas

Gráfico 62 - Experiência

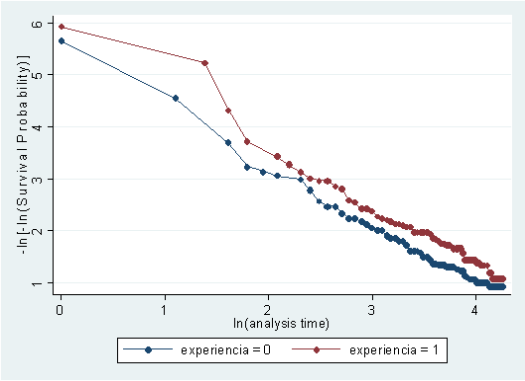


Gráfico 63 - Principal motivação

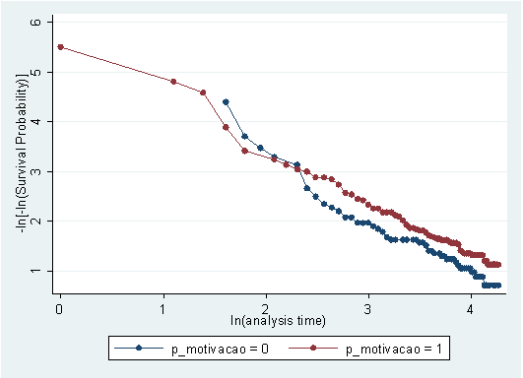
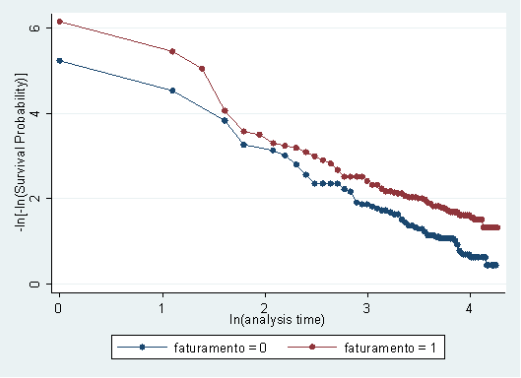


Gráfico 64 - Faturamento



A tabela abaixo mostra a contribuição de cada empresa na composição da função risco e da função sobrevivência.

	id	_t0	_t	_d	h1	s
1.	60	0	.1	1	.01751237	.98248763
2.	369	0	.1	1	.01751237	.98248763
3.	20	0	1	1	.01755237	.96524265
4.	558	0	1	1	.01755237	.96524265
5.	580	0	3	1	.01763321	.94822232
6.	556	0	3	1	.01763321	.94822232
7.	288	0	4	1	.00888968	.93979293
8.	164	0	5	1	.0611102	.882362
9.	425	0	5	1	.0611102	.882362
10.	399	0	5	1	.0611102	.882362
11.	594	0	5	1	.0611102	.882362
12.	549	0	5	1	.0611102	.882362
13.	542	0	5	1	.0611102	.882362
14.	136	0	5	1	.0611102	.882362
15.	446	0	6	1	.07061277	.82005597
16.	500	0	6	1	.07061277	.82005597
17.	436	0	6	1	.07061277	.82005597
18.	151	0	6	1	.07061277	.82005597
19.	523	0	6	1	.07061277	.82005597
20.	621	0	6	1	.07061277	.82005597
21.	494	0	6	1	.07061277	.82005597
22.	366	0	6	1	.07061277	.82005597
23.	202	0	7	1	.00918829	.81252105
24.	168	0	8	1	.03641042	.78293682
25.	616	0	8	1	.03641042	.78293682
26.	342	0	8	1	.03641042	.78293682
27.	329	0	8	1	.03641042	.78293682
28.	248	0	9	1	.01847515	.76847195
29.	258	0	9	1	.01847515	.76847195
30.	287	0	10	1	.0277406	.74715407
31.	638	0	10	1	.0277406	.74715407
32.	270	0	10	1	.0277406	.74715407
33.	31	0	11	1	.0463164	.71254859
34.	413	0	11	1	.0463164	.71254859
35.	267	0	11	1	.0463164	.71254859
36.	454	0	11	1	.0463164	.71254859
37.	630	0	11	1	.0463164	.71254859
38.	520	0	12	1	.0469324	.67910697
39.	152	0	12	1	.0469324	.67910697
40.	367	0	12	1	.0469324	.67910697
41.	154	0	12	1	.0469324	.67910697
42.	362	0	12	1	.0469324	.67910697
43.	637	0	12	0	.	.67910697
44.	132	0	12	0	.	.67910697
45.	662	0	12	0	.	.67910697

46.	499	0	12	0	.	.67910697
47.	266	0	12	0	.	.67910697
48.	133	0	12	0	.	.67910697
49.	427	0	12	0	.	.67910697
50.	38	0	13	1	.01943027	.66591174
51.	103	0	13	1	.01943027	.66591174
52.	183	0	13	0	.	.66591174
53.	105	0	13	0	.	.66591174
54.	19	0	13	0	.	.66591174
55.	652	0	13	0	.	.66591174
56.	361	0	13	0	.	.66591174
57.	311	0	13	0	.	.66591174
58.	290	0	13	0	.	.66591174
59.	646	0	13	0	.	.66591174
60.	360	0	13	0	.	.66591174
61.	544	0	13	0	.	.66591174
62.	12	0	14	1	.01986497	.65268343
63.	498	0	14	1	.01986497	.65268343
64.	265	0	14	0	.	.65268343
65.	61	0	14	0	.	.65268343
66.	554	0	14	0	.	.65268343
67.	126	0	14	0	.	.65268343
68.	36	0	14	0	.	.65268343
69.	423	0	14	0	.	.65268343
70.	426	0	14	0	.	.65268343
71.	543	0	14	0	.	.65268343
72.	131	0	14	0	.	.65268343
73.	655	0	14	0	.	.65268343
74.	545	0	14	0	.	.65268343
75.	477	0	15	1	.04029267	.62638507
76.	142	0	15	1	.04029267	.62638507
77.	392	0	15	1	.04029267	.62638507
78.	557	0	15	1	.04029267	.62638507
79.	324	0	15	0	.	.62638507
80.	37	0	15	0	.	.62638507
81.	269	0	15	0	.	.62638507
82.	540	0	15	0	.	.62638507
83.	421	0	15	0	.	.62638507
84.	651	0	15	0	.	.62638507
85.	312	0	15	0	.	.62638507
86.	161	0	15	0	.	.62638507
87.	424	0	15	0	.	.62638507
88.	359	0	15	0	.	.62638507
89.	591	0	15	0	.	.62638507
90.	162	0	15	0	.	.62638507
91.	629	0	15	0	.	.62638507
92.	279	0	15	0	.	.62638507
93.	35	0	15	0	.	.62638507
94.	496	0	15	0	.	.62638507
95.	642	0	15	0	.	.62638507
96.	645	0	15	0	.	.62638507

97.	122	0	16	1	.07219028	.58116615
98.	495	0	16	1	.07219028	.58116615
99.	163	0	16	1	.07219028	.58116615
100.	488	0	16	1	.07219028	.58116615
101.	555	0	16	1	.07219028	.58116615
102.	200	0	16	1	.07219028	.58116615
103.	298	0	16	1	.07219028	.58116615
104.	280	0	16	0	.	.58116615
105.	518	0	16	0	.	.58116615
106.	592	0	16	0	.	.58116615
107.	553	0	16	0	.	.58116615
108.	626	0	16	0	.	.
109.	661	0	16	0	.	.58116615
110.	628	0	16	0	.	.58116615
111.	434	0	16	0	.	.58116615
112.	184	0	16	0	.	.58116615
113.	395	0	16	0	.	.58116615
114.	541	0	16	0	.	.58116615
115.	236	0	16	0	.	.58116615
116.	325	0	16	0	.	.58116615
117.	654	0	16	0	.	.58116615
118.	350	0	16	0	.	.58116615
119.	323	0	16	0	.	.58116615
120.	93	0	16	0	.	.58116615
121.	104	0	16	0	.	.58116615
122.	326	0	16	0	.	.58116615
123.	130	0	16	0	.	.58116615
124.	465	0	17	1	.01100589	.5747699
125.	492	0	17	0	.	.5747699
126.	422	0	17	0	.	.5747699
127.	18	0	17	0	.	.5747699
128.	182	0	17	0	.	.
129.	536	0	17	0	.	.5747699
130.	241	0	17	0	.	.5747699
131.	539	0	17	0	.	.5747699
132.	538	0	17	0	.	.5747699
133.	160	0	17	0	.	.5747699
134.	94	0	17	0	.	.5747699
135.	510	0	18	1	.04436091	.54927259
136.	327	0	18	1	.04436091	.54927259
137.	363	0	18	1	.	.
138.	135	0	18	1	.04436091	.54927259
139.	508	0	18	1	.04436091	.54927259
140.	173	0	18	0	.	.54927259
141.	537	0	18	0	.	.54927259
142.	353	0	18	0	.	.54927259
143.	358	0	18	0	.	.54927259
144.	278	0	18	0	.	.54927259
145.	237	0	18	0	.	.54927259
146.	493	0	18	0	.	.54927259
147.	349	0	18	0	.	.54927259

148.	418	0	18	0	.	.54927259
149.	352	0	18	0	.	.54927259
150.	11	0	18	0	.	.54927259
151.	150	0	18	0	.	.54927259
152.	34	0	18	0	.	.54927259
153.	206	0	19	1	.01186104	.54275765
154.	240	0	19	0	.	.54275765
155.	17	0	19	0	.	.54275765
156.	660	0	19	0	.	.54275765
157.	59	0	19	0	.	.54275765
158.	653	0	19	0	.	.54275765
159.	473	0	19	0	.	.54275765
160.	394	0	19	0	.	.54275765
161.	351	0	19	0	.	.54275765
162.	371	0	20	1	.03589759	.52327396
163.	601	0	20	1	.03589759	.52327396
164.	435	0	20	1	.03589759	.52327396
165.	125	0	20	0	.	.52327396
166.	590	0	20	0	.	.52327396
167.	357	0	20	0	.	.52327396
168.	552	0	20	0	.	.52327396
169.	239	0	20	0	.	.52327396
170.	420	0	20	0	.	.52327396
171.	235	0	20	0	.	.52327396
172.	268	0	21	1	.04863393	.49782509
173.	339	0	21	1	.04863393	.49782509
174.	98	0	21	1	.04863393	.49782509
175.	320	0	21	1	.04863393	.49782509
176.	535	0	21	0	.	.49782509
177.	641	0	21	0	.	.49782509
178.	419	0	21	0	.	.49782509
179.	491	0	21	0	.	.49782509
180.	181	0	21	0	.	.49782509
181.	551	0	21	0	.	.49782509
182.	92	0	21	0	.	.49782509
183.	650	0	21	0	.	.49782509
184.	222	0	21	0	.	.49782509
185.	640	0	21	0	.	.49782509
186.	301	0	21	0	.	.49782509
187.	497	0	21	0	.	.49782509
188.	511	0	22	1	.01278916	.49145833
189.	534	0	22	0	.	.49145833
190.	289	0	22	0	.	.49145833
191.	490	0	22	0	.	.49145833
192.	416	0	22	0	.	.49145833
193.	58	0	22	0	.	.49145833
194.	300	0	22	0	.	.49145833
195.	261	0	22	0	.	.49145833
196.	636	0	22	0	.	.49145833
197.	532	0	22	0	.	.49145833
198.	322	0	22	0	.	.49145833

199.	639	0	23	1	.05197605	.46591426
200.	66	0	23	1	.05197605	.46591426
201.	438	0	23	1	.05197605	.46591426
202.	319	0	23	1	.05197605	.46591426
203.	550	0	23	0	.	.46591426
204.	32	0	23	0	.	.46591426
205.	533	0	23	0	.	.46591426
206.	33	0	23	0	.	.46591426
207.	417	0	23	0	.	.46591426
208.	393	0	23	0	.	.46591426
209.	275	0	23	0	.	.46591426
210.	517	0	23	0	.	.46591426
211.	649	0	23	0	.	.46591426
212.	159	0	23	0	.	.46591426
213.	95	0	24	1	.02722621	.45322918
214.	355	0	24	1	.02722621	.45322918
215.	233	0	24	0	.	.45322918
216.	234	0	24	0	.	.45322918
217.	16	0	24	0	.	.45322918
218.	63	0	24	0	.	.45322918
219.	433	0	24	0	.	.45322918
220.	10	0	24	0	.	.45322918
221.	111	0	25	1	.01396331	.4469006
222.	429	0	25	0	.	.4469006
223.	531	0	25	0	.	.4469006
224.	26	0	25	0	.	.4469006
225.	101	0	25	0	.	.4469006
226.	415	0	25	0	.	.4469006
227.	310	0	25	0	.	.4469006
228.	91	0	25	0	.	.4469006
229.	474	0	26	1	.02827531	.43426435
230.	90	0	26	1	.02827531	.43426435
231.	410	0	26	0	.	.43426435
232.	610	0	26	0	.	.43426435
233.	412	0	26	0	.	.43426435
234.	102	0	26	0	.	.43426435
235.	356	0	26	0	.	.43426435
236.	348	0	26	0	.	.43426435
237.	146	0	26	0	.	.43426435
238.	472	0	26	0	.	.43426435
239.	414	0	26	0	.	.43426435
240.	611	0	26	0	.	.43426435
241.	221	0	26	0	.	.43426435
242.	242	0	27	1	.01465393	.42790067
243.	614	0	27	0	.	.42790067
244.	529	0	27	0	.	.42790067
245.	30	0	27	0	.	.42790067
246.	431	0	27	0	.	.42790067
247.	411	0	27	0	.	.42790067
248.	100	0	27	0	.	.42790067
249.	432	0	27	0	.	.42790067

250.	89	0	27	0	.	.42790067
251.	80	0	28	1	.04432063	.40893584
252.	259	0	28	1	.04432063	.40893584
253.	328	0	28	1	.04432063	.40893584
254.	457	0	28	0	.	.40893584
255.	57	0	28	0	.	.40893584
256.	336	0	28	0	.	.40893584
257.	9	0	28	0	.	.40893584
258.	338	0	28	0	.	.40893584
259.	548	0	28	0	.	.40893584
260.	116	0	28	0	.	.40893584
261.	296	0	29	1	.06011698	.38435186
262.	569	0	29	1	.06011698	.38435186
263.	148	0	29	1	.06011698	.38435186
264.	560	0	29	1	.06011698	.38435186
265.	589	0	29	0	.	.38435186
266.	623	0	29	0	.	.38435186
267.	56	0	29	0	.	.38435186
268.	29	0	29	0	.	.38435186
269.	516	0	29	0	.	.38435186
270.	337	0	29	0	.	.38435186
271.	99	0	29	0	.	.38435186
272.	530	0	29	0	.	.38435186
273.	232	0	29	0	.	.38435186
274.	147	0	29	0	.	.38435186
275.	79	0	30	1	.04705697	.36626542
276.	365	0	30	1	.04705697	.36626542
277.	452	0	30	1	.04705697	.36626542
278.	391	0	30	0	.	.36626542
279.	172	0	30	0	.	.36626542
280.	171	0	30	0	.	.36626542
281.	180	0	30	0	.	.36626542
282.	299	0	30	0	.	.36626542
283.	613	0	30	0	.	.36626542
284.	647	0	30	0	.	.36626542
285.	316	0	30	0	.	.36626542
286.	220	0	30	0	.	.36626542
287.	204	0	30	0	.	.36626542
288.	284	0	30	0	.	.36626542
289.	347	0	31	0	.	.36626542
290.	515	0	31	0	.	.36626542
291.	97	0	31	0	.	.36626542
292.	27	0	31	0	.	.36626542
293.	489	0	31	0	.	.36626542
294.	633	0	31	0	.	.36626542
295.	456	0	31	0	.	.36626542
296.	179	0	31	0	.	.36626542
297.	400	0	31	0	.	.36626542
298.	528	0	31	0	.	.36626542
299.	609	0	31	0	.	.36626542
300.	321	0	32	1	.01705314	.36001945

301.	409	0	32	0	.	.36001945
302.	608	0	32	0	.	.36001945
303.	547	0	32	0	.	.36001945
304.	618	0	32	0	.	.36001945
305.	527	0	32	0	.	.36001945
306.	48	0	32	0	.	.36001945
307.	584	0	33	1	.03489159	.3474578
308.	139	0	33	1	.03489159	.3474578
309.	526	0	33	0	.	.3474578
310.	8	0	33	0	.	.3474578
311.	158	0	33	0	.	.3474578
312.	346	0	33	0	.	.3474578
313.	28	0	33	0	.	.3474578
314.	484	0	33	0	.	.3474578
315.	546	0	33	0	.	.3474578
316.	588	0	33	0	.	.3474578
317.	455	0	33	0	.	.3474578
318.	96	0	33	0	.	.3474578
319.	170	0	33	0	.	.3474578
320.	622	0	33	0	.	.3474578
321.	344	0	34	0	.	.3474578
322.	276	0	34	0	.	.3474578
323.	231	0	34	0	.	.3474578
324.	524	0	34	0	.	.3474578
325.	612	0	34	0	.	.3474578
326.	525	0	34	0	.	.3474578
327.	430	0	34	0	.	.3474578
328.	340	0	34	0	.	.3474578
329.	55	0	34	0	.	.3474578
330.	292	0	34	0	.	.3474578
331.	277	0	34	0	.	.3474578
332.	354	0	34	0	.	.3474578
333.	615	0	35	1	.0562823	.32790207
334.	512	0	35	1	.0562823	.32790207
335.	343	0	35	1	.0562823	.32790207
336.	607	0	35	0	.	.32790207
337.	149	0	35	0	.	.32790207
338.	389	0	35	0	.	.32790207
339.	308	0	35	0	.	.32790207
340.	586	0	35	0	.	.32790207
341.	345	0	35	0	.	.32790207
342.	78	0	35	0	.	.32790207
343.	54	0	35	0	.	.32790207
344.	264	0	35	0	.	.32790207
345.	587	0	35	0	.	.32790207
346.	487	0	35	0	.	.32790207
347.	364	0	36	1	.07762218	.3024496
348.	502	0	36	1	.07762218	.3024496
349.	479	0	36	1	.07762218	.3024496
350.	579	0	36	1	.07762218	.3024496

351.	408	0	36	0	.	.3024496
352.	178	0	36	0	.	.3024496
353.	405	0	36	0	.	.3024496
354.	333	0	36	0	.	.3024496
355.	238	0	37	1	.02056962	.29622833
356.	428	0	37	0	.	.29622833
357.	251	0	37	0	.	.29622833
358.	390	0	37	0	.	.29622833
359.	203	0	37	0	.	.29622833
360.	24	0	37	0	.	.29622833
361.	229	0	37	0	.	.29622833
362.	297	0	37	0	.	.29622833
363.	285	0	37	0	.	.
364.	106	0	38	1	.04216336	.28373835
365.	243	0	38	1	.04216336	.28373835
366.	304	0	38	0	.	.28373835
367.	219	0	38	0	.	.28373835
368.	45	0	38	0	.	.28373835
369.	115	0	38	0	.	.28373835
370.	13	0	38	0	.	.28373835
371.	76	0	38	0	.	.28373835
372.	383	0	38	0	.	.28373835
373.	378	0	39	1	.0219318	.27751546
374.	14	0	39	0	.	.27751546
375.	218	0	39	0	.	.27751546
376.	87	0	39	0	.	.27751546
377.	88	0	39	0	.	.27751546
378.	606	0	39	0	.	.27751546
379.	509	0	39	0	.	.27751546
380.	72	0	39	0	.	.27751546
381.	120	0	39	0	.	.27751546
382.	77	0	39	0	.	.27751546
383.	404	0	39	0	.	.27751546
384.	656	0	40	0	.	.27751546
385.	604	0	40	0	.	.27751546
386.	118	0	40	0	.	.27751546
387.	271	0	40	0	.	.27751546
388.	643	0	40	0	.	.27751546
389.	307	0	40	0	.	.27751546
390.	224	0	40	0	.	.27751546
391.	585	0	40	0	.	.27751546
392.	380	0	40	0	.	.27751546
393.	514	0	40	0	.	.27751546
394.	75	0	40	0	.	.27751546
395.	73	0	40	0	.	.27751546
396.	114	0	40	0	.	.27751546
397.	483	0	40	0	.	.27751546
398.	605	0	40	0	.	.27751546
399.	22	0	40	0	.	.27751546
400.	381	0	41	1	.04703637	.26446213
401.	185	0	41	1	.04703637	.26446213
402.	250	0	41	0	.	.26446213

403.	382	0	41	0	.	.26446213
404.	581	0	41	0	.	.26446213
405.	155	0	41	0	.	.26446213
406.	74	0	41	0	.	.26446213
407.	260	0	41	0	.	.26446213
408.	113	0	41	0	.	.26446213
409.	84	0	41	0	.	.26446213
410.	471	0	41	0	.	.26446213
411.	217	0	41	0	.	.26446213
412.	169	0	42	0	.	.26446213
413.	603	0	42	0	.	.26446213
414.	53	0	42	0	.	.26446213
415.	249	0	42	0	.	.26446213
416.	25	0	42	0	.	.26446213
417.	451	0	42	0	.	.26446213
418.	625	0	42	0	.	.26446213
419.	291	0	42	0	.	.26446213
420.	186	0	43	1	.02632433	.25750035
421.	306	0	43	0	.	.25750035
422.	406	0	43	0	.	.25750035
423.	44	0	43	0	.	.25750035
424.	295	0	43	0	.	.25750035
425.	403	0	43	0	.	.25750035
426.	522	0	43	0	.	.25750035
427.	659	0	43	0	.	.25750035
428.	453	0	43	0	.	.25750035
429.	582	0	44	1	.02735381	.25045673
430.	294	0	44	0	.	.25045673
431.	274	0	44	0	.	.25045673
432.	230	0	44	0	.	.25045673
433.	583	0	44	0	.	.25045673
434.	119	0	44	0	.	.25045673
435.	201	0	44	0	.	.25045673
436.	385	0	45	1	.0278867	.24347232
437.	332	0	45	0	.	.24347232
438.	644	0	45	0	.	.24347232
439.	263	0	45	0	.	.24347232
440.	85	0	45	0	.	.24347232
441.	602	0	45	0	.	.24347232
442.	83	0	45	0	.	.24347232
443.	305	0	45	0	.	.24347232
444.	521	0	45	0	.	.24347232
445.	450	0	45	0	.	.24347232
446.	157	0	45	0	.	.24347232
447.	86	0	45	0	.	.24347232
448.	331	0	46	0	.	.24347232
449.	167	0	46	0	.	.24347232
450.	70	0	46	0	.	.24347232
451.	177	0	46	0	.	.24347232
452.	624	0	46	0	.	.24347232
453.	216	0	46	0	.	.24347232

454.	635	0	46	0	.	.24347232
455.	572	0	47	1	.03060876	.23601993
456.	379	0	47	0	.	.23601993
457.	3	0	47	0	.	.23601993
458.	658	0	47	0	.	.23601993
459.	141	0	47	0	.	.23601993
460.	577	0	47	0	.	.23601993
461.	482	0	48	1	.06289863	.2211746
462.	109	0	48	1	.06289863	.2211746
463.	283	0	48	0	.	.2211746
464.	41	0	48	0	.	.2211746
465.	112	0	48	0	.	.2211746
466.	576	0	48	0	.	.2211746
467.	15	0	49	1	.15848249	.1861223
468.	376	0	49	1	.15848249	.1861223
469.	480	0	49	1	.15848249	.1861223
470.	330	0	49	1	.15848249	.1861223
471.	255	0	49	1	.15848249	.1861223
472.	156	0	49	0	.	.1861223
473.	166	0	49	0	.	.1861223
474.	388	0	49	0	.	.1861223
475.	196	0	49	0	.	.1861223
476.	470	0	49	0	.	.1861223
477.	145	0	49	0	.	.1861223
478.	71	0	49	0	.	.1861223
479.	449	0	49	0	.	.1861223
480.	315	0	49	0	.	.1861223
481.	341	0	50	1	.03604121	.17941423
482.	407	0	50	0	.	.17941423
483.	262	0	50	0	.	.17941423
484.	318	0	50	0	.	.17941423
485.	7	0	50	0	.	.17941423
486.	199	0	50	0	.	.17941423
487.	632	0	50	0	.	.17941423
488.	49	0	50	0	.	.17941423
489.	42	0	50	0	.	.17941423
490.	448	0	50	0	.	.17941423
491.	469	0	50	0	.	.17941423
492.	134	0	51	1	.03805738	.17258619
493.	578	0	51	0	.	.17258619
494.	21	0	51	0	.	.17258619
495.	571	0	51	0	.	.17258619
496.	198	0	51	0	.	.17258619
497.	197	0	51	0	.	.17258619
498.	23	0	52	0	.	.17258619
499.	69	0	52	0	.	.17258619
500.	124	0	52	0	.	.17258619
501.	129	0	52	0	.	.17258619
502.	574	0	52	0	.	.17258619
503.	575	0	52	0	.	.17258619
504.	600	0	52	0	.	.17258619

505.	507	0	52	0	.	.17258619
506.	175	0	53	0	.	.17258619
507.	402	0	53	0	.	.17258619
508.	573	0	53	0	.	.17258619
509.	377	0	53	0	.	.17258619
510.	447	0	53	0	.	.17258619
511.	195	0	53	0	.	.17258619
512.	468	0	53	0	.	.17258619
513.	225	0	53	0	.	.17258619
514.	506	0	53	0	.	.17258619
515.	467	0	53	0	.	.17258619
516.	215	0	54	0	.	.17258619
517.	223	0	54	0	.	.17258619
518.	317	0	54	0	.	.17258619
519.	272	0	54	0	.	.17258619
520.	68	0	54	0	.	.17258619
521.	505	0	54	0	.	.17258619
522.	140	0	54	0	.	.17258619
523.	273	0	54	0	.	.17258619
524.	478	0	55	1	.09310871	.15651691
525.	513	0	55	1	.09310871	.15651691
526.	52	0	55	0	.	.15651691
527.	387	0	55	0	.	.15651691
528.	192	0	55	0	.	.15651691
529.	634	0	55	0	.	.15651691
530.	466	0	55	0	.	.15651691
531.	445	0	55	0	.	.15651691
532.	627	0	55	0	.	.15651691
533.	398	0	55	0	.	.15651691
534.	257	0	55	0	.	.15651691
535.	504	0	56	0	.	.15651691
536.	620	0	56	0	.	.15651691
537.	110	0	56	0	.	.15651691
538.	213	0	56	0	.	.15651691
539.	67	0	56	0	.	.15651691
540.	194	0	56	0	.	.15651691
541.	214	0	56	0	.	.15651691
542.	176	0	56	0	.	.15651691
543.	212	0	56	0	.	.15651691
544.	247	0	56	0	.	.15651691
545.	4	0	56	0	.	.15651691
546.	40	0	56	0	.	.15651691
547.	475	0	57	1	.0578767	.14745823
548.	335	0	57	0	.	.
549.	401	0	57	0	.	.
550.	570	0	57	0	.	.14745823
551.	211	0	57	0	.	.14745823
552.	6	0	57	0	.	.14745823
553.	565	0	57	0	.	.14745823
554.	657	0	57	0	.	.14745823
555.	444	0	57	0	.	.14745823

556.	46	0	57	0	.	.14745823
557.	246	0	58	0	.	.14745823
558.	227	0	58	0	.	.14745823
559.	190	0	58	0	.	.14745823
560.	375	0	58	0	.	.14745823
561.	282	0	58	0	.	.14745823
562.	373	0	58	0	.	.14745823
563.	597	0	59	0	.	.14745823
564.	193	0	59	0	.	.14745823
565.	226	0	59	0	.	.14745823
566.	208	0	59	0	.	.14745823
567.	128	0	59	0	.	.14745823
568.	293	0	59	0	.	.14745823
569.	50	0	59	0	.	.14745823
570.	568	0	59	0	.	.14745823
571.	107	0	59	0	.	.14745823
572.	599	0	60	0	.	.14745823
573.	144	0	60	0	.	.14745823
574.	619	0	61	0	.	.14745823
575.	123	0	61	0	.	.14745823
576.	370	0	61	0	.	.14745823
577.	256	0	61	0	.	.14745823
578.	117	0	61	0	.	.14745823
579.	476	0	61	0	.	.14745823
580.	443	0	61	0	.	.14745823
581.	503	0	61	0	.	.14745823
582.	374	0	61	0	.	.14745823
583.	566	0	62	1	.23382884	.11297824
584.	51	0	62	1	.23382884	.11297824
585.	253	0	62	1	.23382884	.11297824
586.	595	0	62	0	.	.11297824
587.	210	0	62	0	.	.11297824
588.	481	0	62	0	.	.11297824
589.	165	0	62	0	.	.11297824
590.	598	0	62	0	.	.11297824
591.	137	0	62	0	.	.11297824
592.	441	0	62	0	.	.11297824
593.	127	0	62	0	.	.11297824
594.	567	0	62	0	.	.11297824
595.	2	0	62	0	.	.11297824
596.	108	0	62	0	.	.11297824
597.	209	0	63	0	.	.11297824
598.	228	0	63	0	.	.11297824
599.	174	0	63	0	.	.11297824
600.	82	0	63	0	.	.11297824
601.	138	0	63	0	.	.11297824
602.	442	0	63	0	.	.11297824
603.	302	0	63	0	.	.11297824
604.	463	0	63	0	.	.11297824
605.	121	0	63	0	.	.11297824

606.	464	0	63	0	.	.11297824
607.	47	0	63	0	.	.11297824
608.	281	0	63	0	.	.11297824
609.	485	0	64	1	.12380328	.09899117
610.	303	0	64	0	.	.09899117
611.	245	0	64	0	.	.09899117
612.	254	0	64	0	.	.09899117
613.	461	0	64	0	.	.09899117
614.	462	0	64	0	.	.09899117
615.	372	0	64	0	.	.09899117
616.	191	0	64	0	.	.09899117
617.	314	0	64	0	.	.09899117
618.	64	0	64	0	.	.09899117
619.	519	0	64	0	.	.09899117
620.	384	0	65	0	.	.09899117
621.	207	0	65	0	.	.09899117
622.	564	0	65	0	.	.09899117
623.	596	0	65	0	.	.09899117
624.	486	0	65	0	.	.09899117
625.	62	0	65	0	.	.09899117
626.	501	0	66	0	.	.09899117
627.	39	0	66	0	.	.09899117
628.	1	0	67	0	.	.09899117
629.	368	0	67	0	.	.09899117
630.	460	0	67	0	.	.09899117
631.	631	0	67	0	.	.09899117
632.	397	0	67	0	.	.09899117
633.	563	0	67	0	.	.09899117
634.	648	0	67	0	.	.09899117
635.	244	0	67	0	.	.09899117
636.	143	0	68	0	.	.09899117
637.	5	0	68	0	.	.09899117
638.	309	0	68	0	.	.09899117
639.	81	0	68	0	.	.
640.	286	0	68	0	.	.
641.	458	0	68	0	.	.09899117
642.	593	0	68	0	.	.09899117
643.	189	0	68	0	.	.09899117
644.	617	0	68	0	.	.09899117
645.	440	0	68	0	.	.09899117
646.	439	0	69	0	.	.09899117
647.	562	0	69	0	.	.09899117
648.	386	0	69	0	.	.09899117
649.	437	0	69	0	.	.09899117
650.	43	0	69	0	.	.09899117
651.	188	0	69	0	.	.09899117
652.	65	0	69	0	.	.09899117
653.	187	0	69	0	.	.09899117
654.	205	0	70	0	.	.09899117
655.	334	0	70	0	.	.09899117
656.	396	0	70	0	.	.09899117

```
657. | 313     0   70    0           .   .09899117 |
658. | 561     0   70    0           .   .09899117 |
659. | 252     0   70    0           .   .09899117 |
660. | 559     0   71    0           .   .09899117 |
     |-----------------------------------------------|
661. | 459     0   71    0           .   .09899117 |
662. | 153     0   71    0           .   .09899117 |
     +-----------------------------------------------+
```

Capitulo 4 - Conclusão

No Brasil, os estudos que se utilizam de dados microeconômicos aplicados aos modelos de duração são escassos se excluirmos os trabalhos cuja informação seja oriunda de fontes públicas, como PME/IBGE (pesquisa mensal de emprego do Instituto Brasileiro de Geografia e Estatística), por exemplo. Nesse nível de detalhe, informações sobre empresas são praticamente inexistentes ao domínio público, o que torna este trabalho ainda mais raro e profícuo, ao apresentar resultados e contribuir, eventualmente, com o planejamento de futuras políticas e ações voltadas à promoção do desenvolvimento empresarial seja no Estado de São Paulo quanto no país. O grande diferencial deste trabalho foi o acesso a base de dados ainda inexplorada para a aplicação dos modelos de duração, o que possibilitou quantificar as variáveis, dentro do contexto de sobrevivência, mais importantes para determinar a longevidade das micro, pequenas e médias empresas dentro dos respectivos setores de atuação.

A base de dados original continha 1998 empresas dentre as quais selecionamos, conforme premissas anteriormente mencionadas, 662 delas. Como o levantamento do SEBRAE[46] compreende empresas de todos os pontos do Estado de São Paulo em que a instituição possui plataformas de atendimento, pode ter incorrido, pelos critérios de seleção adotados neste trabalho, eventual concentração de empresas de uma única região em detrimento de outras. Aparentemente, esse fato não seria capaz de influenciar os resultados a ponto de proporcionar viés nos resultados, mesmo admitindo que dentro do Estado de São Paulo haja regiões mais avançadas, do ponto de vista econômico, do que outras.

Dentre os resultados deste estudo, observamos que o mercado atua como um ambiente seletivo em que as empresas que não deixam o mercado (falham) possuem características semelhantes entre si tal como avaliadas pelas variáveis que se considerou neste estudo.

No contexto econômico, mostramos que as variáveis consideradas não sofreram influências significativas do ambiente macroeconômico. Note-se que as duas últimas crises financeiras com efeitos em amplitude mundial ocorreram após o período de coleta da base de dados.

No decorrer dos modelos apresentados neste trabalho, comparamos, quando possível, os resultados encontrados neste estudo com aqueles previamente estabelecidos pela literatura, como, por exemplo, a importância do tamanho da empresa, com o intuito de oferecer

[46] Para detalhes sobre as regiões de coleta ver SEBRAE (2008).

evidências adicionais para a literatura. As informações em nível de firma, na maioria dos casos, abrangem características de cada país ou região, fato que torna difícil replicar exatamente os mesmos resultados de outros trabalhos, o que reforça o significado da complementaridade dos resultados deste estudo em relação aos disponíveis na literatura.

Para as 662 empresas da amostra, selecionamos 33 variáveis que nos possibilitaram identificar quais são os principais fatores e quanto cada um contribui para a sobrevivência das empresas, por grupo de tamanho e setor de atuação. Nos dois primeiros modelos apresentados, com distribuição exponencial e Weibull, encontramos para o primeiro que um ano adicional na idade do empreendedor contribui para redução do risco de falha em 1,4% enquanto que o acompanhamento das finanças resultou em redução de 41,2%. Outras variáveis relevantes para esse modelo foram o investimento em qualificação da mão de obra e tempo de planejamento, com redução do risco de falha em 47,5% e 1,9%, respectivamente.

Para o modelo Weibull, cinco variáveis foram significativas, contribuindo para redução do risco de falha das empresas no tempo, a saber: idade do empreendedor (-1,7% - para cada ano adicional na idade do empreendedor), principal motivação em abrir a empresa (-27,8% - se oportunidade), acompanhamento das finanças (-39,7% - se acompanhava), investimento em qualificação da mão de obra (-47,9% - se investia) e tempo de planejamento (-1,8% - para cada mês adicional de planejamento).

Na abordagem dos modelos de riscos proporcionais, aplicamos o modelo pra a mesma base utilizada nos modelos de Weibull e exponencial e, também, propusemos modelos por setor de atividade e tamanho das empresas. No modelo com a base completa, também encontramos cinco variáveis significativas para redução da probabilidade de falha da empresa no tempo, a saber: idade (-1,6% - para cada ano adicional), experiência (-26,4% - para cada ano adicional, principal motivação (-29,2% - se oportunidade), acompanhamento das finanças (-36,6% - se acompanhava) e investimento em qualificação da mão de obra(-43,8% - se investia).

Para os modelos setoriais, o modelo para comércio destacou que o acompanhamento de finanças contribuiu para redução do risco de falha em 51,1%, enquanto que o modelo de serviços destacou três variáveis, a saber: idade (-3,5% - para cada ano adicional), investimento em qualificação da mão de obra (62,7% - se investia) e quantidade de funcionário (-14,7% - para cada funcionário adicional).

Os resultados obtidos para os modelos que segmentaram as empresas pelo tamanho (quantidade de funcionários), consideraram como micro empresas aquelas com até cinco funcionários e pequena empresa as demais. O modelo encontrado para as micro empresas

apontou as variáveis idade, investimento em qualificação da mão de obra e tempo de planejamento como principais responsáveis para redução do risco de falha no tempo em -1,7%, -48,2% e 1,9%, respectivamente. O modelo para pequenas empresas destacou outras duas variáveis, a saber: inovação em processos e elaboração de plano de negócios como principais fatores, cujas contribuições foram de -74,5% e -82,9%, respectivamente, para redução do risco de falha das empresas.

Os resultados encontrados para os modelos paramétricos: exponencial e Weibull foram próximos, excetuando-se a relevância da variável principal motivação para o último, no caso em que o empreendedor tenha iniciado o negócio por uma oportunidade de mercado. Para os modelos semi-paramétricos ou de riscos proporcionais, os resultados obtidos com a base completa apontam para resultados muito próximos entre si (em que as mesmas variáveis se mostraram relevantes nos três modelos.

No caso dos modelos segregados pelo tamanho das empresas, os resultados evidenciaram que o tamanho das empresas provavelmente tem impacto na determinação das variáveis que são mais importantes para contribuir com a redução de eventual falha no tempo conforme o estagio de desenvolvimento em que a empresa se encontra. Já os modelos setoriais, os resultados permitiram evidenciar que embora as empresas estejam em um mesmo ambiente legal e econômico, os desafios a serem enfrentados pelos empreendedores desses setores, em termos de redução do risco de falha das empresas, seguem por caminhos diferentes.

Finalmente, possíveis linhas de pesquisa sugeridas por este trabalho apontam para uma análise mais ampla, que compreenda empresas de todos os Estados do país, possibilitando encontrar perfis regionais e, também, um perfil “nacional" para as empresas brasileiras, cujas características permitam ao setor público e privado promover outras políticas publicas mais específicas para cada segmento e/ou porte de empresa em cada região do país. Estudos que possibilitem a análise em um período mais longo também podem contribuir para identificar os efeitos dos ciclos econômicos, das mudanças macroeconômicas e institucionais e contribuir para enriquecer a compreensão do tema.

Referências Bibliográficas

AUDRETSCH, D. (1991): "New-firm survival and the technological regime", *Review of Economics and Statistics*, 72 (3), p. 441-450.

AUDRETSCH, D. and MAHMOOD, T. (1995): "New-firm survival: new results using a hazard function", *Review of Economics and Statistics* vol. 77, pp. 97-103.

AUDRETSCH, D., SANTARELLI, E. and VIVARELLI, M. (1999): "Start-up size and industrial dynamics: some evidence from Italian manufacturing", *International Journal of Industrial Organization* vol. 17, pp. 965-83.

BRESLOW, N. and CROWLEY, J. (1974): "A large sample study of the life table and product limit estimates under random censorship". *Annals of Statistic*, Vol. 2, number 3, 437-453.

BRESLOW, N. E. (1974): "Covariance analysis of censored survival data", *Biometrics* 30: 579-594.

CASTROGIOVANNI, G. (1996): "Pre-startup planning and the survival of new small business: theoretical linkages". *Journal of Management*, Vol. 22, nº 6, p. 801-822.

COLLETT, D. (1994): *Modeling survival data in medical research*, London: Chapman and Hall.

COLOSIMO, E. A., GIOLO, S. R. (2006): *Análise de sobrevivência aplicada*. São Paulo: Edgar Blücher.

COX, D. R. (1972): "Regression models and life table", *Journal of Royal Statistics Society*, Series B, 34(2) , pp. 187-202.

COX, D. R. e OAKES, D. (1984): *Analysis on survival data*, New York: Chapman and Hall.

CRAMER, J. S. (1986): *Econometric applications of maximum likelihood methods*. Cambridge University Press.

DUNNE, T., ROBERTS, M. and SAMUELSON, L. (1988): "Patterns of firm entry and exit in US manufacturing industries", *Rand Journal of Economics* vol. 104, 671-98.

EVANS, D., (1987a): "Tests of alternative theories of firm growth", *Journal of Political Economy*, 95(4), pp. 657-674. (August)

EVANS, D., (1987b): "Firms growth, size, and age: estimates for 100 manufacturing industries", *Journal of Industrial Economics*. (June)

GARRET, J. M. (1997): "Graphical assessment of the Cox model proportional hazards assumption". *Stata Technical Bulletin* 35: 9-14. Stata Press.

GEHAN, E. A. (1965): “A generalized wilcoxon test for comparing arbitrarily singly-censored samples”. *Biometrika*, 52(1/2), 203-223.

GEROSKI, P. A. (1991): *Market dynamics and entry*. Oxford: Blackwell.

GEROSKI, P. A. (1995): "What do we know about entry?”, *International Journal of Industrial Organization*, Vol. 13, pp. 421-440.

GLOBAL ENTREPRENEURSHIP MONITOR. Executive Report**,** Babson College Press, 2011.

GRAMBSCH, P. M. and THERNEAU, T. M. (1994): “Proportional hazard test and diagnostic based on weighted residuals”. *Biometrika* 81: 515-26.

GREENE, W. H. (2003): *Econometric Analysis*. Prentice Hall, 1026 p.

HECKMAN, J. and SINGER, B., (1984): “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data”, *Econometrica*, 52(2), pp. 271-320.

HOLMES, P., STONE, I. e BRADFORD, P. (2010), “An Analysis of New Survival using a Hazard Function”, *Applied Economics*, Vol. 42, pp. 185-195.

IBGE. http://www.ibge.gov.br/home/estatistica/economia/industria/pintec/2008/defaultzipcnae2_faixa.shtm: acesso em 10/08/2011.

JOVANOVIC, B. (1982): “Selection and Evolution of Industry”. *Econometrica*, 50(3), 649-670.

KALBFLEISCH, J. D. and PRENTICE, R. L. (1980): *The Statistical Analysis of Failure Time Data*, John Wiley and Sons Inc., New York.

KALBFLEISCH, J. D. and PRENTICE, R. L. (2002): *The Statistical Analysis of Failure Time Data*. J. Wiley and Sons Inc., New York. 439 p.

KAPLAN, E.L. and MEIER, P. (1958): "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association*, **53**, 457-481.

KIEFER, N., (1988): “Econometric Duration Data and Hazard Functions”, Journal of *Economic Literature*, 26(2), pp. 646-679.

LANCASTER, T. (1990). *The Econometric Analysis of Transition Data*, in Econometric Society Monograph **17**, Cambridge University Press, Cambridge.

LEI nº 123, de 14/12/2006. https://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp123.htm#art89: acesso em 25/08/2012.

LIN, D. Y. and WEI, L.J. (1989): “The robust inference for the Cox proportional hazard model” *Journal of the American Statistical Association*, 84(408), p. 1074-1078.

MATA, J. (1992): “Small firms in the Portuguese manufacturing industries” in Audretsch, D. and ACS, Z. (eds.) *Small Firms and Entrepreneurship: An East-West perspective, Cambridge, MA:* Cambridge University Press.

MATA, J. and Portugal, P. (1994): “Life duration of new firms”. *Journal of Industrial Economics*, Vol. 42, pp. 227-245.

MATA, J., PORTUGAL, P. e GUIMARÃES, P. (1995): “The survival of news plants: star-up conditions and pos-entry evolution”. *International Journal of Industrial Organization*, 459-481.

MATA, J. and PORTUGAL, P. (2004): “Patterns of entry, post-entry growth and survival: a comparison between domestic and Foreign owned Firms”. *Small Business Economics*, Springer, vol. 22(3_4), pages 283-298, 04.

PETO, R. and PETO, J. (1972): “Asymptotically efficient rank invariant test procedures”, *Journal of the Royal Statistical Society*, Series A (General) 135(2), 185-207.

PHILLIPS, B. and KIRCHHOFF, B., (1989), “Formation, Growth and Survival: Small Firm Dynamics in the US Economy”, *Small Business Economics*, 1(1), pp. 65-74.

RAIS. http://bi.mte.gov.br/bgcaged/cagedanuariorais/anuario.htm: *Relatório Anual de Informações Sociais (MTE)*, acesso em 05/10/2012.

SEBRAE-SP – Anuário do trabalho na Micro e Pequena Empresa – http://bis.sebrae.com.br /GestorRepositorio/ARQUIVOS_CHRONUS/bds/bds.nsf/25BA39988A7410D78325795D003E8172/$File/NT00047276.pdf – acesso em 03/02/2013.

SEBRAE-SP – Serviço de Apoio às Micro e Pequenas Empresas – *Sobrevivência e Mortalidade das empresas Paulistas de 1 a 5 anos*. SEBRAE-SP. São Paulo, 2005.

SEBRAE-SP – Serviço de Apoio às Micro e Pequenas Empresas - *10 Anos de Monitoramento da Sobrevivência e Mortalidade de Empresas*. SEBRAE-SP. São Paulo, 2008.

SCHUMPETER, J. Capitalismo, Socialismo e Democracia - Rio de Janeiro: Editora Fundo de Cultura, 1961. (Traduzido do inglês *Capitalism, Sociualism and Democracy*. Connectcut 1942).

SCHUMPETER, J. A teoria do desenvolvimento econômico. São Paulo: Nova Cultural (1985)

WOOLDRIDGE, J. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.