



HAL
open science

Régularités spatiales et temporelles dans la réussite et/ou l'échec des entreprises

Dègnon David Dadakpete

► **To cite this version:**

Dègnon David Dadakpete. Régularités spatiales et temporelles dans la réussite et/ou l'échec des entreprises. Economies et finances. Université Bourgogne Franche-Comté, 2023. Français. NNT : 2023UBFCG008 . tel-04394419

HAL Id: tel-04394419

<https://theses.hal.science/tel-04394419>

Submitted on 15 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE
PREPAREE A L'UNIVERSITE DE BOURGOGNE FRANCHE-COMTE

École doctorale n°593

DGEP

Doctorat en Sciences Économiques

Par

Monsieur DADAKPETE Dègnon David

Régularités spatiales et temporelles dans la réussite et/ou l'échec des entreprises

Thèse présentée et soutenue à l'Université de Bourgogne, le 21 novembre 2023

Composition du Jury :

Monsieur, LEDEZMA, Ivan	Professeur à l'Université de Bourgogne, Dijon, France	Président du Jury
Monsieur, DUBE, Jean	Professeur à l'Université de Laval, Québec, Canada	Rapporteur
Monsieur, BRUNELLE Cédric	Professeur à l'INRS, Québec, Canada	Rapporteur
Madame, LE GALLO, Julie	Professeur à l'Institut Agro, Dijon, France	Examinatrice
Madame, GALLIE Émilie-Pauline	Inspectrice générale de l'éducation du sport et de la recherche	Examinatrice
Monsieur, LEGROS Diego	Professeur à l'Université de Bourgogne, Dijon, France.	Directeur de thèse

Titre : Régularités spatiales et temporelles dans la réussite et/ou l'échec des entreprises

Mots clés : Survie des Entreprises, Environnement Spatial, Modèles de Survie

Résumé : Dans le monde d'aujourd'hui où la notion de frontière est constamment repensée, les territoires au sein des nations ne sont pas égaux quant à l'adaptation à de tels changements. Le développement régional devient ainsi un défi majeur pour de nombreuses économies, en particulier celles les plus avancées où l'on note des disparités régionales de plus en plus croissantes. Pour faire face à ce challenge, les autorités publiques ont besoin, d'identifier les différentes sources d'inégalités territoriales, mais surtout de comprendre les mécanismes de création de richesse dans les régions. Dans la littérature sur l'entrepreneuriat, les entreprises sont considérées comme moteur de croissance économique et de développement régional. Dans le même temps, plusieurs études montrent que les entreprises ne peuvent durablement contribuer au développement local que lorsqu'elles survivent et prospèrent sur le marché.

A ce propos, les travaux existants identifient trois groupes de facteurs déterminant la survie des entreprises. Il s'agit des caractéristiques spécifiques à l'entreprise, des caractéristiques spécifiques à l'industrie dans laquelle l'entreprise opère, et des caractéristiques de l'environnement spatial de l'entreprise. Cette thèse aborde en quatre chapitres la question de la survie des entreprises en mettant l'emphase sur les caractéristiques de l'environnement spatial de l'entreprise, qui sont les moins abordées dans la littérature. La mise en œuvre d'une variété de techniques économétriques sur une diversité d'échantillons, a permis de mettre en exergue l'impact significatif qu'exerce l'environnement spatial de l'entreprise sur sa survie.

Title : Spatial and temporal regularities in the success and/or the failure of firms

Keywords : Firm survival, Spatial environment, Survival models

Abstract : In today's world where the notion of borders is constantly being revised, territories within nations are not equal when it comes to adapting to such changes. Regional development is thus becoming a major challenge for many economies, particularly the most advanced ones, where regional disparities appear to be on the increase. To cope with this challenge, public authorities need to identify the sources of territorial inequalities, but more importantly, they need to get a thorough understanding of the mechanisms behind wealth creation in regions. In the literature on entrepreneurship, firms are considered as fundamental drivers of economic growth and regional development. At the same time, several studies show that firms can only make a lasting contribution to local development when they survive and thrive in the market.

In this regard, existing studies identify three groups of factors determining firm survival. These are characteristics specific to the firm, characteristics specific to the industry in which the firm operates, and characteristics of the firm's spatial environment. This thesis addresses the issue of firm survival in four chapters, focusing on the characteristics of the firm's spatial environment, which are the least explored in the literature. Relying on a variety of econometric methods on a diversity of samples, we show that a firm survival is significantly affected by characteristics of its spatial environment.

Université Bourgogne Franche-Comté
32, avenue de l'Observatoire
25000 Besançon

Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'avenir portant la référence ANR-10-EQPX-17 (Centre d'accès sécurisé aux données – CASD)

« L'université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses. Ces opinions doivent être considérées comme propres à leurs auteurs. »

Remerciements

Ce projet de thèse est l'une des expériences les plus passionnantes que j'ai eu l'opportunité d'avoir. Elle fût riche en péripéties et en enseignements. Dès les premières lectures, l'on réalise la pertinence de cette phrase d'Isaac Newton, qui prétend que nos connaissances s'apparentent à une goutte dans un océan de savoir. Cette leçon vient littéralement exacerber notre soif de connaissance. L'un des enseignements que l'on tire également d'une telle aventure, c'est qu'il ne s'agit pas de tout savoir d'avance, mais qu'il s'agit surtout d'avoir la motivation et la curiosité, nécessaires pour aller chercher la connaissance là où elle se trouve.

Cette thèse était également une belle aventure humaine, au cours de laquelle j'ai eu le privilège d'apprendre énormément des collègues doctorants, et des enseignants du Laboratoire d'Economie de Dijon. Je leur exprime, à tous sans exception, ma très profonde gratitude pour les agréables moments que nous avons pu passer.

Le présent document n'aurait pas pu prendre cette forme finale sans la rigoureuse et bienveillante supervision de mon directeur de thèse, le Professeur Diego LEGROS. Sous son encadrement, j'ai pu apprendre la rigueur et l'exigence qui caractérisent l'écriture scientifique, et je lui en sais profondément gré.

Je voudrais également adresser mes sincères remerciements au Professeur Jean DUBE, au Professeur Cédric BRUNELLE, au Professeur Julie LE GALLO, au Professeur Yvan LEDEZMA et à Madame Emilie-Pauline GALLIE, qui ont accepté de faire partir de mon jury de thèse. J'en suis particulièrement honoré car j'ai eu le privilège durant mon parcours, d'admirer la qualité et l'impact de leurs travaux de recherche.

Je dédie ce travail à mon Père, ma Mère, mon Frère et ma Sœur. Murielle!, mes mots ne pourront que très modestement exprimer ma gratitude vis-à-vis de toi. Je te remercie pour ta disponibilité et la minutie dont tu as fait preuve lors des dernières relectures.

*«What we know is a drop, what we don't know
is an ocean.»*

Isaac Newton

«Ce que je sais, c'est que je ne sais rien.»

Socrate

Résumé

Dans le monde d'aujourd'hui où la notion de frontière est constamment repensée, les territoires au sein des nations ne sont pas égaux quant à l'adaptation à de tels changements. Le développement régional devient ainsi un défi majeur pour de nombreuses économies, en particulier celles les plus avancées où l'on note des disparités régionales de plus en plus croissantes. Pour faire face à ce challenge, les autorités publiques ont besoin, d'identifier les différentes sources d'inégalités territoriales, mais surtout de comprendre les mécanismes de création de richesse dans les régions. Dans la littérature sur l'entrepreneuriat, les entreprises sont considérées comme moteur de croissance économique et de développement régional. Dans le même temps, plusieurs études montrent que les entreprises ne peuvent durablement contribuer au développement local que lorsqu'elles survivent et prospèrent sur le marché. A ce propos, les travaux existants identifient trois groupes de facteurs déterminant la survie des entreprises. Il s'agit des caractéristiques spécifiques à l'entreprise, des caractéristiques spécifiques à l'industrie dans laquelle l'entreprise opère, et des caractéristiques de l'environnement spatial de l'entreprise. Cette thèse aborde en quatre chapitres la question de la survie des entreprises en mettant l'accent sur les caractéristiques de l'environnement spatial de l'entreprise, qui sont les moins abordées dans la littérature. La mise en œuvre d'une variété de techniques économétriques sur une diversité d'échantillons, a permis de mettre en exergue l'impact significatif qu'exerce l'environnement spatial de l'entreprise sur sa survie.

MOTS CLÉS : Survie des Entreprises, Environnement Spatial, Modèles de Survie

Abstract

In today's world where the notion of borders is constantly being revised, territories within nations are not equal when it comes to adapting to such changes. Regional development is thus becoming a major challenge for many economies, particularly the most advanced ones, where regional disparities appear to be on the increase. To cope with this challenge, public authorities need to identify the sources of territorial inequalities, but more importantly, they need to get a thorough understanding of the mechanisms behind wealth creation in regions. In the literature on entrepreneurship, firms are considered as fundamental drivers of economic growth and regional development. At the same time, several studies show that firms can only make a lasting contribution to local development when they survive and thrive in the market. In this regard, existing studies identify three groups of factors determining firm survival. These are characteristics specific to the firm, characteristics specific to the industry in which the firm operates, and characteristics of the firm's spatial environment. This thesis addresses the issue of firm survival in four chapters, focusing on the characteristics of the firm's spatial environment, which are the least explored in the literature. Relying on a variety of econometric methods on a diversity of samples, we show that a firm survival is significantly affected by characteristics of its spatial environment.

KEYWORDS : Firm survival, Spatial environment, Survival models

Table des matières

Remerciements	iii
Résumé	v
Abstract	vi
Introduction générale	1
1 Survie des établissements et environnement géographique	7
Introduction	7
1.1 Littérature	9
1.1.1 Littérature théorique	9
1.1.2 Littérature empirique	11
1.1.2.1 Caractéristiques spécifiques à l'entreprise	12
1.1.2.2 Caractéristiques de l'industrie	15
1.1.2.3 Caractéristiques de l'environnement géographique	17
1.2 Analyse économétrique	20
1.2.1 Modèles économétriques	20
1.2.1.1 Modèles de survie	21
1.2.2 Présentation des données et des variables	26
1.2.3 Analyse descriptive	34
1.2.3.1 Statistiques descriptives	35
1.2.3.2 Analyse non paramétrique de survie : Estimateur de Kaplan-Meier	39
1.2.4 Résultats d'estimations	42
1.2.4.1 Modèles à hasards proportionnels de Cox	42
1.2.4.1.1 Environnement spatial de l'entreprise	45
1.2.4.1.1.1 Effets positifs sur la survie	45
1.2.4.1.1.2 Effets négatifs sur la survie	46
1.2.4.1.2 Variables de contrôle	47
1.2.4.2 Modèles à hasards proportionnels de Cox avec stratification	48
1.2.4.2.1 Prise en compte des hétérogénéités spatiales	48
1.2.4.2.2 Résultats	50
1.3 Modèle théorique	57
1.3.1 Les ménages	57

1.3.2	Les entreprises	60
1.3.3	Équilibre	63
1.3.4	Discussion	65
	Conclusion	66
Annexes		68
	Annexe 1.A Catégories d'équipements	68
	Annexe 1.B Tests log-rank	71
	Annexe 1.C Ratios de risques	80
	Annexe 1.D Modèle théorique	82
1.D.1	Ménages	82
1.D.2	Entreprise	84
2	Choix de localisation et survie des jeunes entreprises françaises	86
	Introduction	86
2.1	Choix de localisation des entreprises	87
2.1.1	Littérature théorique	87
2.1.2	Littérature empirique	96
2.2	Environnement spatial et survie des jeunes entreprises	100
2.3	Analyse économétrique	101
2.3.1	Stratégie d'analyse de la survie	102
2.3.1.1	Modèle à hasards proportionnels Exponentiel	103
2.3.1.1.1	Choix d'un modèle Exponentiel	105
2.3.1.2	Modèle à hasards proportionnels de Weibull	106
2.3.1.2.1	Choix d'un modèle de Weibull	106
2.3.2	Stratégie d'estimation du choix de localisation	107
2.3.3	Données et variables	109
2.3.3.1	Données	109
2.3.3.2	Variables explicatives	110
2.3.4	Analyse descriptive	114
2.3.4.1	Statistiques descriptives	114
2.3.4.2	Analyse non paramétrique de survie de Kaplan Meier	119
2.3.5	Résultats d'estimations	121
2.3.5.1	Résultats d'analyse de la survie	121
2.3.5.2	Choix de localisation	126
	Conclusion	130

Annexes	132
Annexe 2.A Tests log-rank	132
Annexe 2.B Ratios de risques	146
Annexe 2.C Coefficients estimés	149
3 Environnement spatial du quartier général et survie de l'entreprise	151
Introduction	151
3.1 Littérature	152
3.2 Les quartiers généraux d'entreprises françaises	156
3.2.1 Données et échantillons	156
3.2.2 Variables	156
3.2.3 Analyse	158
3.3 Localisation du quartier général et la survie de l'entreprise	165
3.3.1 Données et variables	165
3.3.2 Modèles économétriques	171
3.3.3 Résultats	172
3.3.3.1 Signes des coefficients et Significativité	173
3.3.3.2 Discussion	175
3.3.3.3 Stratification par catégories d'entreprises	177
Conclusion	182
Annexes	184
Annexe 3.A Statistiques descriptives	184
Annexe 3.B Ratios de risques	193
4 Competition effect in spatial concentration of firms: A spatial survival model approach	194
Introduction	194
4.1 Competition and firm survival	196
4.2 Spatial dependence in survival analysis	198
4.3 Model and Likelihood	201
4.4 Bayesian Approach	205
4.5 Data and summary statistics	210
4.6 Results	216
Conclusion	219
Annexes	221

Appendix 4.A Implementing the Model using Stata's Matrix programming	221
Appendix 4.B Using 3001 randomly selected plants from Île-de-France	223
Conclusion générale	226
Bibliographie	242

Liste des tableaux

1.1	Codes et libellés des variables	32
1.2	Correspondance des variables dans les tableaux de résultat	34
1.3	Statistiques descriptives : Échantillon	35
1.4	Statistiques descriptives : Établissements fermés vs actifs	37
1.5	Secteurs d'activités	38
1.6	Fréquences des secteurs : Échantillon	39
1.7	Fréquences des secteurs : Établissements fermés vs actifs	39
1.8	Résultats des modèles de Cox	44
1.9	Analyse descriptive des strates	54
1.10	Résultats d'estimation	56
1.B.1	Taux de dépendance économique	71
1.B.2	Taux de chômage	71
1.B.3	Taille de l'établissement	79
1.B.4	Secteurs d'activités	79
1.C.1	Ratios de risques COX sans stratification	80
1.C.2	Ratios de risques COX avec stratification	81
2.3.1	Codes et libellés des variables	113
2.3.2	Description des variables qualitatives	115
2.3.3	Entreprises fermées	116
2.3.4	Description des variables quantitatives	117
2.3.5	Description par catégories	118
2.3.6	Modèle de Cox	122
2.3.8	Résultats d'estimation : coefficients	128
2.3.9	Ratios de risque	129
2.A.1	Age du créateur de l'entreprise	132
2.A.2	Taille de l'entreprise au démarrage	132
2.A.3	Équipements en actions sociales et dans le commerce	133
2.A.4	Équipements dans l'enseignement au premier et au second degré	134
2.A.5	Équipements médicaux et paramédicaux, enseignement supérieur et autres	135
2.A.6	Équipements en Santé, Services aux particuliers	136
2.A.7	Équipements en Sport, loisirs, culture, Transport et Tourisme	137
2.A.8	Taux de dépendance économique et Chômage	138
2.A.9	Nbre d'entreprises pour 1000 et Indice Maurel-Sédillot	139

2.A.10	Expérience en création d'entreprise, et niveau de diplôme	140
2.A.11	Filiale et type de secteur	141
2.A.12	Innovations de procédé et de marketing	142
2.A.13	Innovations en organisation et de produit	143
2.A.14	Citoyenneté et Ambition	144
2.A.15	Sexe du créateur	145
2.B.1	Modèle Exponentiel	146
2.B.2	Modèle de Weibull	147
2.B.3	Modèle de Cox	148
2.C.1	Modèle Exponentiel	149
2.C.2	Modèle de Weibull	150
3.2.1	Variables	157
3.2.2	Statistiques descriptives	158
3.2.3	Caractéristiques des départements	160
3.2.4	Caractéristiques des départements	162
3.3.1	Codes et libellés des variables	168
3.3.2	Statistiques descriptives 2005-2015 pour tout l'échantillon	169
3.3.3	Statistiques descriptives 2005-2015 : Entreprises fermées versus censurées	170
3.3.4	Résultats	174
3.3.5	Description des catégories d'entreprises : 2005-2008	178
3.3.6	Description des catégories d'entreprises : 2009-2012	178
3.3.7	Description des catégories d'entreprises : 2013-2015	178
3.3.8	Test Log-rank	179
3.3.9	Modèle avec stratification	181
3.A.1	Secteurs d'activités	184
3.A.2	Variables quantitatives 2005-2007	185
3.A.3	Variables quantitatives 2008-2010	186
3.A.4	Variables quantitatives 2011-2013	187
3.A.5	Variables quantitatives 2014-2015	188
3.A.6	Variables quantitatives 2005-2007 (entreprises fermées)	189
3.A.7	Variables quantitatives 2008-2010 (entreprises fermées)	190
3.A.8	Variables quantitatives 2011-2013 (entreprises fermées)	191
3.A.9	Variables quantitatives 2014-2015 (entreprises fermées)	192
3.B.1	Ratios de risques modèle sans stratification	193
4.5.1	Variables	213

4.5.2	Statistics	216
4.6.1	Parameters	219
4.B.1	Sampled parameters	225

Table des figures

1.1	Variables quantitatives	40
1.2	Secteur d'activités	41
1.3	Choroplèthe Communes	51
1.4	Choroplèthe Zones d'emploi	52
1.5	I de Moran Communes	52
1.6	I de Moran Zones d'emploi	53
1.7	Clusters Communes	53
1.8	Clusters Zones d'emploi	54
2.1.1	Modèle de Von Thünen	89
2.1.2	Modèle de Christaller	93
2.1.3	Lieux centraux en Allemagne du sud	94
2.1.4	Demande et coût moyen (Modèle de Lösch (1940))	95
2.3.1	Test pour modèle Exponentiel	105
2.3.2	Test pour modèle de Weibull	107
2.3.3	Calendrier SINE	109
2.3.4	Variables quantitatives	119
2.3.5	Dummies	120
3.2.1	Choroplèthe du nombre de QG	159
3.2.2	Les départements les moins attractifs	160
3.2.3	Les départements les plus prisés	162
3.2.4	Matrice de corrélation	164
3.3.1	Fonctions de survie de Kaplan Meier	179
4.5.1	Gross domestic product per capita in USD, France (OECD 2019)	211
4.5.2	Patent PCT applications per million inhabitants, France 2015	211
4.5.3	Population per km ² , France 2020	212
4.5.4	Failed	214
4.5.5	censored	214
4.5.6	Failed and Censored	215
4.6.1	Sample	217
4.6.2	Density	218
4.B.1	Sample	223
4.B.2	Density	224

Introduction générale

Dans un contexte de mondialisation accélérée où les inégalités régionales s'accroissent au sein des économies les plus intégrées (Ezcurra & Rodríguez-Pose, 2013; Ezcurra & Del Villar, 2021), le développement régional s'avère être un défi majeur pour les autorités publiques. En Europe par exemple, plus particulièrement au sein de l'Union Européenne, on observe d'importantes disparités économiques et sociales se caractérisant par la concentration des richesses dans une minorité de régions. D'après les données de 2018 (Regions in Europe, 2020 edition, s. d.), les 10 régions (sur un total de 240) les plus riches représentaient plus de 20% du Produit Intérieur Brut de toute l'Union. De même, les valeurs régionales du Produit Intérieur Brut par habitant oscillaient entre 30% et 263% de la moyenne de l'Union Européenne. Ces inégalités conduisent l'Union Européenne à accorder une place centrale à sa politique régionale mise en œuvre par le biais de trois instruments principaux : le Fonds Européen de Développement Régional (FEDER), le Fonds Social Européen (FSE) et le Fonds de Cohésion (FC). Ces trois fonds représentaient 32,5% (367 milliards d'euros) du budget européen pour la programmation 2014-2020 (Dicharry et al., 2019).

La résolution des problèmes d'inégalités territoriales en termes de développement et de croissance passe avant tout par l'identification précise et la compréhension profonde de leurs sources, et cela a été pendant longtemps l'objectif principal de plusieurs travaux de recherche en science régionale. Deux interrogations majeures reviennent assez souvent dans la littérature sur le développement régional : par quel mécanisme la richesse se crée-t-elle dans les régions? Comment assurer une répartition équitable des richesses dans les régions? (Nijkamp et al., 2009). La première question, sans laquelle la seconde n'a pas lieu, revêt un caractère particulièrement primordial dans la mesure où elle pousse à réfléchir sur les principaux facteurs de création de richesse.

Dans la littérature sur l'entrepreneuriat, le rôle des entreprises dans l'innovation, dans la création d'emploi et les transformations économiques diverses, a été largement démontré (Danson, 2002, chap. 13). Littunen (2001) avance que l'emploi et la croissance économique locale, deux indicateurs clés de la santé économique régionale, peuvent être augmentés via la promotion des nouvelles entreprises. Seninger (1989) argumente que lorsque les ressources humaines locales sont efficacement utilisées, le développement local est positivement impacté de même que la productivité des entreprises, et Malecki (1993) déclare que l'entrepreneuriat est la base même du développement local. Plusieurs travaux empiriques (Audretsch & Fritsch, 2002; Van Stel & Suddle, 2008) aboutissent à une même conclu-

sion selon laquelle les nouvelles entreprises impactent positivement l'emploi régional. Les entreprises jouent donc un rôle d'importance non négligeable dans le développement économique local.

Par ailleurs, il est important de noter qu'une santé économique durable ne saurait être garantie par la création d'entreprises ayant une courte durée de vie. [Haltiwanger et al. \(1999\)](#) trouvent que la productivité des entreprises augmente avec l'âge. [Shane \(2009\)](#) dira même qu'il relève d'une mauvaise politique économique que d'encourager plus de gens à entreprendre, car ce ne sont pas les nouvelles entreprises qui apportent des emplois mais ce sont les entreprises déjà en place. On comprend donc qu'il ne suffit pas que les entreprises naissent pour favoriser le développement de leur territoire d'implantation, mais il leur faut survivre. [Derbyshire \(2013\)](#) trouve à partir d'estimations réalisées sur des données anglaises, que les régions ou les entreprises survivent plus longtemps connaissent une plus forte croissance en termes d'emploi. La survie des entreprises apparaît donc comme une question cruciale qu'il convient de se poser.

Entre 2012 et 2013, le taux de mortalité des entreprises employant des salariés, au sein de l'OCDE¹, oscillait entre 1 et 21% et valait en moyenne 10% ([OCDE, 2017](#)). En France, 28% des entreprises créées en 2010 et 25% de celles créées en 2014 ont fermé trois ans plus tard (Rapport INSEE Première N^o 1751 Avril 2019). Dans la littérature, les travaux sur la survie des entreprises considèrent généralement trois grandes catégories de facteurs : Les caractéristiques spécifiques à l'entreprise, les caractéristiques spécifiques à l'industrie dans laquelle l'entreprise opère et les caractéristiques spécifiques à l'environnement spatial de l'entreprise. Les deux premières catégories de facteurs ont été abondamment abordées alors qu'on note une certaine négligence quant à la caractérisation de l'environnement spatial de l'entreprise. Cette thèse propose en quatre chapitres, de contribuer à combler cette lacune en mettant ainsi l'emphase sur l'environnement géographique de l'entreprise.

L'objectif principal de cette thèse consiste donc à analyser les facteurs de réussite (ou d'échec) des entreprises en fonction de leurs environnements immédiats. L'indicateur de réussite ou d'échec retenu ici est l'état administratif (fermeture ou non) de l'entreprise sur la durée d'observation des données. De manière globale, le cadre conceptuel de notre travail s'inspire de l'économie évolutionniste. Cette dernière s'intéresse aux processus par lesquels l'économie se transforme de l'intérieur et étudie les implications pour les entreprises, les industries, la production, l'emploi et la croissance ([Witt, 2016](#)). Le but ultime de l'économie évolutionniste, c'est d'assimiler la dynamique à l'origine du fonctionnement du mar-

1. Organisation de Coopération et de Développement Économique.

ché et du comportement des entreprises (Nelson & Winter, 1982). Cette thèse s'inscrit plus particulièrement dans la lignée des travaux issus de l'économie géographique évolutionnaire qui, se focalise sur les mécanismes par lesquels l'organisation spatiale de l'économie est transformée dans le temps (Boschma & Martin, 2007). Lorsqu'on mène le raisonnement à une échelle microéconomique, il est question de comprendre la dynamique des entreprises dans un contexte spatial. Elles (les entreprises) choisissent tout d'abord leurs lieux d'implantation, ensuite, elles mettront progressivement en place un mécanisme de production en prenant en compte les conditions locales. Certaines d'entre elles réussiront à prospérer tandis que d'autres quitteront le marché. Le succès ou l'échec des entreprises impactent profondément le paysage économique régional. Les régions, loin d'être de simples unités spatiales passives accueillant les activités économiques, représentent des environnements en perpétuel changement (Essletzbichler & Rigby, 2010). Ce faisant, les caractéristiques d'une région conditionnent pour une part non négligeable la dynamique industrielle.

La survie d'une entreprise ne peut donc être étudiée sur la simple base de ses décisions prétendument rationnelles et de ses caractéristiques, mais il conviendrait de prendre en compte le contexte spatial. Dans les quatre chapitres de cette thèse, nous étudions, toutes choses égales par ailleurs, la survie des entreprises en fonction des caractéristiques de leurs lieux d'implantation.

Le chapitre 1 intitulé *Survie des établissements et environnement géographique*, s'intéresse aux déterminants géographiques de la survie de 1 651 798 établissements français observés de 1995 à 2005 et dont 39,26% ont fermé sur la période. L'originalité du chapitre réside dans le fait que nous y proposons de penser la question aussi bien sous un angle empirique que théorique. L'analyse empirique, basée sur l'estimation de modèles à hasards proportionnels de Cox, a montré que l'environnement spatial affecte significativement la survie des établissements. De la même manière, l'analyse théorique, inspirée du modèle de Schröder & Sørensen (2012), a permis d'établir que l'âge auquel l'entreprise ferme est lié aux caractéristiques de l'environnement spatial lorsque ces dernières sont prises en compte. Nos résultats encouragent particulièrement les politiques publiques visant à augmenter la dotation infrastructurelle des régions. Toutefois, en théorie, l'entreprise est supposée prendre connaissance en amont des réalités de chaque territoire d'implantation possible, et choisir celui qui lui est optimal (profit maximal, coût minimal). L'on pourrait donc penser que les caractéristiques d'une localisation rationnellement choisie impacteraient positivement la survie de l'entreprise contrairement aux caractéristiques d'une localisation choisie par défaut en raison des coûts d'installation plus abordables par exemple. Ainsi, un même

facteur géographique peut favoriser la survie d'une entreprise et causer l'échec d'une autre, d'où l'importance d'inclure le choix de localisation de l'entreprise dans l'analyse de sa survie. Dans le chapitre 2, nous proposons d'aborder la question en suivant un échantillon de jeunes entreprises françaises les cinq premières années de leur vie.

Le second chapitre est intitulé ***Choix de localisation et survie des jeunes entreprises Françaises***, et porte sur 45 268 entreprises créées en 2010 et suivies jusqu'en 2015 grâce au Système d'Information sur les Nouvelles Entreprises (SINE). La question de recherche soulevée dans ce chapitre est induite par certains résultats (Brixy & Grotz, 2007; Huiban, 2011) de la littérature qui montrent que certaines zones géographiques, en l'occurrence les zones urbaines, enregistrent à la fois un fort taux de création, et de fermeture d'entreprises. Ces zones semblent donc aussi attractives, qu'hostiles pour les jeunes entreprises. Par ailleurs, lorsqu'on se réfère aux différentes théories de localisation qui nous expliquent que l'entreprise s'implante là où elle espère réaliser un profit maximal, les résultats de Brixy & Grotz (2007); Huiban (2011) peuvent nous faire penser que les entreprises ont mal choisi leur localisation, ou qu'elles ont effectué des choix par défaut ne pouvant pas se permettre mieux. L'objectif du chapitre 2 est de déterminer la propension de l'entreprise à choisir sa localisation observée, et de prendre ce choix en compte dans l'analyse de sa survie. L'estimation de plusieurs modèles de survie nous montre dans un premier temps que la survie des jeunes entreprises est influencée par les caractéristiques de l'entreprise, de l'entrepreneur, du secteur d'activité, et de l'environnement spatial. Nous avons ensuite mis en œuvre un logit conditionnel afin d'estimer la propension que l'entreprise choisisse sa localisation observée. L'inclusion de ces propensions dans un modèle de Cox provoque des variations au niveau des impacts de certaines variables décrivant la localisation de l'entreprise. Les résultats révèlent également que plus il est probable pour une entreprise de choisir sa localisation observée, moins il est risqué pour elle de fermer.

Dans le troisième chapitre intitulé, ***Environnement spatial du quartier général et survie de l'entreprise***, l'emphase est mise sur le lieu d'implantation du quartier général de l'entreprise. Généralement, le fonctionnement de l'entreprise se structure autour de deux axes aussi importants l'un que l'autre : le management et la production. Le management est assuré par le quartier général et lorsqu'il est défaillant, l'entreprise peut se retrouver en faillite. Dans beaucoup de cas, le quartier général de l'entreprise est géographiquement séparé des unités de production, car ces deux entités n'ont pas nécessairement besoin des mêmes types de ressources pour bien fonctionner. Dans la littérature, plusieurs études (Lovely et al., 2005; Ono, 2006; Aarland et al., 2007; Davis & Henderson, 2008; Strauss-Kahn & Vives, 2009) ont montré que la localisation du quartier général n'est pas choisie de manière

hasardeuse et ce choix peut impacter les performances des différentes unités de productions (Audia & Rider, 2010; Landier et al., 2009; Kalnins & Lafontaine, 2013). Toutefois, les travaux existants n'explorent qu'en partie les conséquences du choix de localisation des quartiers généraux d'entreprise, dans la mesure où ils se limitent à analyser l'impact de la distance entre quartier général et établissements, sur la survie de ces derniers. Le chapitre 3 propose une approche plus générale consistant à caractériser l'environnement spatial du quartier général et évaluer son influence sur la survie de l'entreprise. L'estimation de modèles à hasards proportionnels de Cox sur un échantillon de 761 108 entreprises françaises, montre que les caractéristiques du lieu d'implantation du quartier général de l'entreprise, affectent significativement sa survie.

Le dernier chapitre dont le titre est *Competition effect in spatial concentration of firms : A spatial survival model approach*, vient davantage clarifier l'un des résultats les plus récurrents le long des précédents chapitres, tout en ayant aussi une portée méthodologique. Dans les précédents chapitres, les résultats d'estimation ont montré que les entreprises implantées dans des communes plus densément peuplées ont des chances de survie plus réduites. Nous avons expliqué ce résultat, déjà trouvé dans la littérature, par la concurrence acerbe que les entreprises peuvent rencontrer dans les grandes métropoles. Néanmoins, la variable densité de population mesurée à l'échelle communale, n'a pas de fondements théoriques microéconomiques, et ne permet pas de bien décrire les mécanismes derrière la concurrence entre entreprises. La notion de concurrence elle-même est une interaction entre individus, et d'après la première loi de la Géographie (Tobler, 1970), on peut supposer que les entreprises vivent davantage cette concurrence dans leur voisinage immédiat. De manière générale, l'existence d'interactions (concurrence ou relation commerciale) entre les entreprises entraîne des interdépendances entre leur probabilité de survie (Maté-Sánchez-Val et al., 2018; Vivel-Búa & Lado-Sestayo, 2021) et il est important de modéliser cette réalité. Les modèles de survie proposent les "fragilités" spatialement corrélées (Zhou & Hanson, 2015; Banerjee, 2016) pour prendre en compte les corrélations spatiales au sein des données de survie, mais l'idée derrière la notion de "fragilité" ne convient pas forcément à certains types de données comme les données d'entreprises. Dans la littérature économique, le probit spatial sera utilisé, mais le modèle probit est quelque peu limité lorsqu'il s'agit des données de survie (Kleinbaum & Klein, 2010; Liu, 2012). Dans le chapitre 4, nous proposons un modèle de survie spatiale applicable sur micro-données. Pour ce faire, nous partons d'un modèle à durée de vie accélérée exponentielle, et nous y incluons un décalage spatial. Nous écrivons ensuite la vraisemblance à laquelle nous associons des distributions à priori afin de déterminer une distribution à posteriori. Le modèle est utilisé

sur un échantillon de 300 établissements franciliens opérant dans le secteur manufacturier. Nous avons utilisé des méthodes de Monte-Carlo par chaînes de Markov pour simuler 30 000 observations à partir de la distribution à postériori. Les résultats indiquent que la survie des entreprises est négativement impactée par la survie de leurs voisines.

Chapitre 1

Survie des établissements et environnement géographique

Introduction

La croissance économique et la création d'emploi sont indéniablement deux objectifs clés de la politique économique, et en la matière, le secteur privé représente un pilier important.

En France, les entreprises jouent un rôle prépondérant dans la création d'emplois et de la richesse. En 2017, les entreprises génèrent 68,4% de la valeur ajoutée de l'ensemble de l'économie (INSEE, 2019). Cette même année, 69,18% des 27,9 millions de personnes en emploi exercent dans le secteur tertiaire marchand, l'industrie la construction, et l'agriculture (INSEE, 2019). Ces différentes statistiques donnent un aperçu du poids du secteur privé en France et permettent d'avoir une idée des répercussions que peuvent avoir sa défaillance sur l'économie nationale. Il est ainsi crucial de s'interroger sur les différents facteurs qui déterminent la performance du secteur privé.

Depuis le travail séminal de David Birch (1979) qui montra entre autres l'importance des petites et moyennes entreprises dans la création d'emplois, davantage de travaux se sont intéressés au lien entre l'entrepreneuriat et le développement économique en termes de croissance et de création d'emplois. Bien que le rapport de David Birch (1979) fit l'objet de multiples controverses, plusieurs études ultérieures (Davidsson et al., 1994; Audretsch & Fritsch, 2002; Van Stel & Suddle, 2008; Baptista et al., 2008) aboutiront aux mêmes conclusions. L'impact économique (innovation, croissance, création d'emploi) des entreprises sera depuis lors régulièrement mis en exergue dans la littérature.

Les premiers travaux de recherche, généralement empiriques, portant sur les entreprises peuvent être catégorisés selon deux grandes approches qui se distinguent par l'échelle de l'analyse. On note, d'un côté, une approche macroéconomique qui s'intéresse aux questions relatives aux rôles des entreprises dans la société (innovation, création d'emploi), à la dynamique industrielle et aux agglomérations; et de l'autre côté une approche microéconomique qui traite des sujets comme la gestion, la croissance, et la survie des entreprises (Landström, 2005). Le présent chapitre s'inscrit dans une approche microéconomique en ce sens qu'il s'intéresse à la survie de l'entreprise dans son environnement local.

Les contributions pionnières (empiriques) au sujet de la survie des entreprises remontent aux travaux de [D. S. Evans \(1987\)](#), [Dunne et al. \(1989\)](#) et [Audretsch \(1991\)](#). [D. S. Evans \(1987\)](#) s'intéresse à la croissance et la survie d'entreprises américaines de l'industrie manufacturière et trouve l'un des résultats les plus récurrents dans la littérature. Son analyse révèle un effet positif de l'âge et la taille des entreprises sur leurs chances de survie. Plus tard en 1989, [Dunne et al. \(1989\)](#) soulèvent de même la problématique de la croissance et du succès des entreprises américaines du secteur manufacturier et en viennent à la même conclusion que [D. S. Evans \(1987\)](#) concernant l'effet positif de l'âge et la taille sur la survie. Cependant, [D. S. Evans \(1987\)](#) et [Dunne et al. \(1989\)](#) n'incluent pas les caractéristiques de l'industrie dans leur analyse. C'est ce qui justement, va motiver le travail de [Audretsch \(1991\)](#). Ce dernier va s'intéresser aux facteurs à l'origine de l'hétérogénéité des chances de survie des entreprises à travers différentes industries. Ses résultats suggèrent que la survie des entreprises est positivement impactée par les industries favorisant l'innovation, et celles à faibles économies d'échelle.

Les jalons étant ainsi posés, de nombreux travaux viendront ensuite enrichir la littérature sur la survie des entreprises. Mais bien que la question fit, et continue de faire l'objet de recherche de plusieurs travaux dans la littérature, il n'y a que peu d'études qui s'intéressent à l'influence de l'environnement spatial immédiat de l'entreprise sur sa survie. Elles négligent ainsi un aspect très important de la relation entre l'entreprise et sa localisation. Lorsqu'une entreprise s'implante à un endroit, elle contribue à la transformation économique de celui-ci de par ses différentes activités. Elle participe au dynamisme économique de la région à travers l'innovation. Néanmoins, tout ce processus ne peut aboutir que lorsque l'environnement spatial de l'entreprise est favorable. L'innovation, qui est l'«arme principale» de la concurrence entre entreprises ([Baumol, 2002](#)), peut parfois nécessiter que l'entreprise se localise dans un environnement géographique particulier. De nombreuses études ([Asheim & Isaksen, 2002](#); [Feldman & Audretsch, 1999](#); [Thompson, 2006](#); [Shearmur, 2010](#)) mettent en évidence l'existence d'un facteur géographique dans les activités d'innovation. Il ressort principalement de ces travaux que certaines localisations s'avèrent plus propices pour innover. De manière générale, le lieu d'implantation d'une entreprise peut avoir des répercussions sur les différentes phases de son cycle d'exploitation à savoir : l'approvisionnement en inputs, la production, et la commercialisation. L'approvisionnement en inputs peut s'avérer plus ou moins facile en fonction de la localisation de l'entreprise. Les régions richement dotées de ressources naturelles et d'infrastructures de transport peuvent faciliter l'accès aux matières premières et leur transport. Concernant la phase de production, elle requiert le recours à une main d'oeuvre de qualité. Ainsi les entreprises implan-

tées dans les zones bien dotées en infrastructures d'éducation peuvent mieux réussir cette phase dans la mesure où elles auraient accès à une ressource humaine mieux qualifiée. Quant à la commercialisation, elle nécessite un marché et une demande conséquente, en d'autres termes, il faut des consommateurs ayant un certain pouvoir d'achat. Les entreprises localisées dans des zones dont les habitants ont un niveau de vie élevé, auront probablement plus de débouchés pour leurs produits. Le lieu d'implantation apparaît ainsi comme un élément important pour l'entreprise, et il conviendrait de s'y intéresser lorsqu'on analyse sa survie.

Le but de ce chapitre est de contribuer à la littérature en investiguant sur les déterminants géographiques de la survie des établissements français. Pour ce faire nous menons une analyse empirique sur 1 651 798 établissements français observés de 1995 à 2005. Nous proposons ensuite de penser la question sous un angle théorique en nous inspirant du modèle de [Schröder & Sørensen \(2012\)](#).

La suite du chapitre est organisée autour de trois grandes sections. Nous commencerons par une revue de littérature dans laquelle seront évoqués les facteurs déterminants de la survie des entreprises. Nous présenterons ensuite une analyse économétrique et notre modèle théorique, après quoi nous conclurons.

1.1 Littérature

Il existe une littérature aussi vaste que diversifiée sur la survie des entreprises. Cette littérature est largement dominée par des travaux empiriques en raison de la grande disponibilité des données d'entreprises et la mise au point de nouvelles techniques d'estimation. On retrouve cependant en économie industrielle quelques travaux théoriques évoquant implicitement la question. Dans ces études, l'accent n'est pas forcément mis de façon directe sur la survie des entreprises, mais il est souvent question de dynamique industrielle, c'est à dire de l'entrée et de la sortie des firmes dans l'industrie. Dans cette section, nous présentons les littératures, théorique et empirique, portant sur la survie des entreprises.

1.1.1 Littérature théorique

Dans le domaine de l'économie industrielle, plusieurs études théoriques s'intéressent à la modélisation de la dynamique industrielle. Ces travaux étudient la transformation de l'industrie dans le temps en analysant les facteurs régissant les différentes étapes du cycle

de vie des entreprises. La question de la survie des entreprises est ainsi abordée dans certains modèles expliquant les entrées et sorties d'entreprises sur le marché. A ce propos, [Brock \(1972\)](#) est l'un des pionniers en la matière.

Dans son modèle, [Brock \(1972\)](#) n'inclut pas d'hétérogénéité entre les entreprises, il les suppose toutes identiques au sein de l'industrie. L'équilibre du modèle conduit à une situation où le nombre d'entreprises dans l'industrie est globalement stable dans le temps, et de ce fait il n'y a plus d'entrées ni de sorties sur le long terme. L'absence de caractéristiques idiosyncratiques aux firmes dans le modèle de Brock va inspirer plusieurs autres travaux.

[Jovanovic \(1982\)](#) propose un modèle pour mieux comprendre la croissance et la survie des entreprises. Dans son modèle, l'auteur suppose une industrie avec un bien homogène et une infinité de firmes qui diffèrent par leur efficacité (maîtrise des coûts de production) et qui n'ont aucune influence sur le prix. Le travail de Jovanovic est l'un des premiers qui introduit un aspect stochastique spécifique aux entreprises et susceptible de régir la dynamique de l'industrie. L'auteur définit un choc de productivité spécifique aux entreprises, et matérialisé par un paramètre normalement distribué. Les résultats montrent entre autres, que les firmes de grande taille survivent plus longtemps que celles de petite taille.

Dans le même esprit que Jovanovic, [Hopenhayn \(1992\)](#) analysera les entrées et sorties de firme dans un modèle dynamique stochastique. Il introduira un choc de productivité stochastique idiosyncratique aux firmes, et suivant un processus de markov. La résolution du modèle aboutit à certains résultats cohérents avec la littérature empirique : le risque d'échec est beaucoup plus faible pour les grandes firmes et celles les plus âgées.

D'autres études de la dynamique industrielle vont incorporer dans leur analyse les structures de marché imparfaite. [Fudenberg & Tirole \(1986\)](#) Développent un modèle de duopole avec information incomplète pour étudier la sortie des entreprises du marché. Ils basent leur raisonnement sur les coûts de production. Les entreprises entrent sur le marché avec une connaissance précise de leurs coûts de production et une incertitude quant à ceux de leurs rivales. La sortie du marché a lieu pour une firme lorsqu'elle réalise qu'en raison du pouvoir de sa rivale, il ne lui est plus profitable de continuer à produire. [Klepper \(2002\)](#), impressionné par l'évolution des industries de l'automobile, du pneu, de postes téléviseurs, et de la pénicilline, s'intéresse au mécanisme par lequel les industries comportant un grand nombre de firmes au départ évoluent pour devenir des oligopoles. Dans son modèle, la survie des entreprises est principalement tributaire de leur capacité à innover.

En plus des travaux d'économie industrielle qui étudient les entrées et sorties d'entreprises dans des modèles théoriques, il existe également une branche de la littérature qui

aborde la question de la survie des entreprises. Il s'agit des modèles de commerce international. Dans ce domaine, le travail séminal de [Melitz \(2003\)](#) constitue un point de départ important ouvrant la voie à de nombreuses analyses.

Dans son papier, [Melitz \(2003\)](#) s'intéresse à l'impact du commerce international sur la dynamique industrielle. Pour ce faire il développe un modèle d'économie industrielle qui prend en compte l'hétérogénéité des entreprises. Ses résultats montrent que l'ouverture commerciale est bénéfique aux entreprises les plus productives, et participe à l'augmentation du bien être du pays. Le cadre d'analyse proposé par Melitz est très novateur dans la mesure où il se démarque des anciens modèles qui évoquent la survie. En effet, l'auteur raisonne dans un contexte d'équilibre général où les firmes sont hétérogènes et opèrent dans une industrie où règne la concurrence monopolistique. Cela offre une description relativement plus réaliste comparativement aux travaux antérieurs. Toutefois, le travail de Melitz soulève certaines interrogations. Dans son modèle, [Melitz \(2003\)](#) introduit un paramètre de choc négatif aléatoire qui expose les entreprises aux mêmes risques de décès indépendamment de leur productivité. Cela est contraire aux faits stylisés qui veulent par exemple que les entreprises les plus productives survivent plus longtemps que les moins productives. [Schröder & Sørensen \(2012\)](#) s'inspirent du cadre théorique de [Melitz \(2003\)](#) et proposent une approche permettant de retrouver des conclusions conformes aux faits stylisés. Les auteurs vont introduire un progrès technique exogène dans le modèle de [Melitz \(2003\)](#). Ils s'inspirent de [Dasgupta & Stiglitz \(1981\)](#) et proposent une distribution des productivités qui s'améliore continuellement dans le temps tel dans un modèle à génération de capital. Leurs résultats montrent entre autres, que les firmes les plus productives survivent plus longtemps, la taille des firmes impacte positivement leur survie, les firmes exportatrices survivent plus longtemps que les non exportatrices.

La plupart des conclusions auxquelles ont abouti ces différents travaux théoriques, vont être corroborées par plusieurs études empiriques.

1.1.2 Littérature empirique

Bien que l'hétérogénéité des travaux empiriques suggère une kyrielle d'éléments susceptibles d'influencer la survie des entreprises, trois groupes de facteurs reviennent assez souvent. Il s'agit des caractéristiques propres à l'entreprise, des caractéristiques de l'industrie où elle opère et des caractéristiques de son environnement géographique.

1.1.2.1 Caractéristiques spécifiques à l'entreprise

La Taille, l'Age et la Survie

La taille et l'âge sont les deux caractéristiques spécifiques à l'entreprise les plus récurrentes dans les études sur la survie des entreprises. Ces deux variables apparaissent déjà très tôt dans les premiers travaux qui abordent la question de la survie des entreprises. Ainsi, [D. S. Evans \(1987\)](#) trouve, en travaillant sur les petites entreprises américaines, que l'âge et la taille augmentent la probabilité pour les entreprises de survivre. Cette relation positive sera confirmée par plusieurs études subséquentes à celle de [D. S. Evans \(1987\)](#). [Mata & Portugal \(1994\)](#) étudient la durée de vie des entreprises de l'industrie manufacturière portugaise. Ils découvrent entre autres que la taille affecte positivement la survie des entreprises. [Audretsch & Mahmood \(1995\)](#) estiment un modèle de Cox sur 12 251 établissements américains et trouvent que la taille de l'établissement réduit sa probabilité de fermeture. [Stearns et al. \(1995\)](#) après avoir observé les jeunes entreprises créées entre 1979 et 1985 dans le Minnesota et la Pennsylvanie au États-Unis, concluent que les entreprises âgées survivent mieux que les jeunes entreprises. [Storey & Wynarczyk \(1996\)](#) travaillent sur un échantillon de 298 petites entreprises anglaises et trouvent que l'âge et la taille augmentent les chances de survie. Plus récemment dans les deux dernières décennies¹, la littérature sera enrichie par diverses études ([Fotopoulos & Louri, 2000](#); [Staber, 2001](#); [Cefis & Marsili, 2005](#); [Huiban, 2011](#); [F. M. Neffke et al., 2011](#); [Mas-Verdú et al., 2015](#); [Iwasaki & Kočenda, 2020](#)) portant sur une variété d'échantillons et corroborant l'effet favorable de l'âge et de la taille sur la survie des entreprises.

L'explication, généralement admise, de l'impact positif de la taille, est qu'il existerait dans chaque industrie, une échelle minimale d'efficacité difficilement atteignable par les entreprises de petite taille. L'échelle minimale d'efficacité représente un niveau de production permettant à l'entreprise de minimiser ses coûts moyens sur le long terme et d'être compétitive sur le marché. Les petites entreprises ont généralement moins de ressources pour atteindre un tel niveau de production, et cela est périlleux pour leur survie. Concernant l'âge, les jeunes entreprises feraient face à ce que [Stinchcombe \(1965\)](#) appelle «aléa de la nouveauté»², et cela les exposerait à un plus grand risque d'échec comparativement aux anciennes entreprises. Lorsqu'une jeune entreprise s'établit dans un nouvel environnement, elle a besoin de temps pour s'adapter et se faire une place sur un marché occupé par des firmes qui ont eu le temps de se constituer un capital social, économique et po-

1. 2000-2020.

2. "Liability of newness" en anglais.

litique [Hannan & Freeman \(1977\)](#). La jeune entreprise doit réussir à attirer et recruter des employés qualifiés, créer de la symbiose entre eux et mettre en place les routines qui lui permettront d'exercer efficacement son activité. Il lui faut également se construire un portefeuille conséquent de clients. Durant cette période critique de son développement, il est très ardu pour la jeune entreprise d'être compétitive, et cela réduit ses chances de survie.

Bien que de l'impact positif de l'âge et de la taille soit un résultat prépondérant dans la littérature, il convient de mentionner qu'il y a toutefois des travaux aboutissant à des conclusions inverses. [P. Holmes et al. \(2010\)](#), sur la base de 781 firmes britanniques, trouvent que la taille initiale des micro-entreprises affecte négativement leur survie. Aussi, contrairement à la notion de l'«aléa de la nouveauté», la relation entre l'âge et la survie n'évolue pas toujours dans un même sens. Dans la littérature, on distingue également les notions d'«aléa de l'adolescence» et d'«aléa de la sénescence» ([Hannan, 1998](#)). Selon l'«aléa de l'adolescence», l'âge et les risques de mortalité sont liés par une relation en U inversé. Dans un premier temps, les risques de mortalité augmentent progressivement avec l'âge jusqu'à atteindre un optimum; puis décroissent dans un second temps. Pour ce qui de l'«aléa de la sénescence», il stipule qu'au delà d'un certain âge les entreprises ont de plus en plus de mal à survivre en raison des difficultés d'adaptation à un environnement de plus en plus concurrentiel. L'âge et la taille bien que très présents, ne sont bien évidemment pas les seules caractéristiques étudiées. Les activités d'innovation et d'exportation apparaissent également comme déterminants de la survie des entreprises.

Innovation, Exportation et Survie

Les entreprises, quelque soit leur statut (nouvel entrant, ou déjà établi) sur le marché, ont constamment besoin de recourir à diverses stratégies pour sécuriser ou étendre leur part de marché. Dans cet univers de concurrence, l'innovation apparaît comme l'une des armes les plus redoutables ([Baumol, 2002](#)). Lorsqu'un nouvel entrant innove, elle peut s'accaparer la place d'une entreprise déjà établie. Inversement, lorsqu'une entreprise déjà établie innove, elle peut protéger sa position, gagner de nouvelles parts de marché et prendre de l'avance sur ses concurrents. L'innovation est ainsi un facteur déterminant de la survie des entreprises.

Dans la majorité des cas, l'innovation est positivement associée à la survie des entreprises. [Kimura & Fujii \(2003\)](#) étudient un échantillon d'entreprises japonaises et trouvent que les dépenses en R&D augmentent les chances de survie. [Pérez et al. \(2004\)](#) en viennent à la même conclusion en travaillant sur des données d'entreprises manufacturières espa-

gnoles. Cefis & Marsili (2005) s'intéressent à la relation entre l'innovation et la survie des entreprises hollandaises. Leurs résultats suggèrent l'existence de ce qu'ils appellent une prime d'innovation qui favoriserait la survie des entreprises qui innovent. Helmers & Rogers (2010); Madrid et al. (2015) aboutiront de même à un lien positif entre innovation et survie. Par ailleurs, il faut noter que l'impact positif de l'innovation sur la survie ne fait pas totalement l'unanimité dans la littérature. Pendant que certains travaux démontrent que l'innovation est sans effet sur la survie (Børing, 2015), d'autres concluront qu'elle augmente le risque de mortalité (Hyytinen et al., 2015). Le travail de Børing (2015) porte sur 1 000 entreprises norvégiennes dont 40% opèrent dans l'industrie manufacturière. Ses résultats d'estimation suggèrent que l'innovation n'a guère d'effet significatif sur le risque de fermeture de l'entreprise. Quant à Hyytinen et al. (2015), ils travaillent sur 1 165 jeunes entreprises finlandaises. Leurs résultats montrent que les entreprises innovantes ont un taux de survie de six à sept points de pourcentage inférieur à celui des entreprises non-innovantes. Les conclusions des travaux de Børing (2015) et de Hyytinen et al. (2015), peuvent paraître contre-intuitives, mais peuvent toutefois se comprendre. Les activités d'innovation sont aussi coûteuses qu'incertaines, et les chances qu'elles aboutissent au succès escompté peuvent être très faible. D'après Buddelmeyer et al. (2010) : «seule une fraction des projets R&D seront couronnés de succès en termes de résolution du problème technique préalablement ciblé. Parmi ces projets de R&D réussis, une minorité aboutira à l'innovation d'un nouveau procédé ou au lancement d'un nouveau produit.». Aussi, est-il possible que le lien positif que plusieurs travaux empiriques établissent entre innovation et survie découle d'un biais de sélection. En effet, lorsqu'on étudie la relation entre survie et innovation à partir d'un échantillon d'entreprises suffisamment établies sur le marché, l'on risque de sélectionner des entreprises dont l'activité d'innovation a déjà été un succès. Cela augmente donc les chances d'avoir un effet positif.

Concernant les activités d'exportation, elles sont généralement considérées comme favorables à la survie de l'entreprise. L'exploration de nouvelles opportunités sur un marché étranger, n'est pas à la portée de toutes les entreprises. Celles qui réussissent à braver les éventuelles barrières commerciales et mener des activités d'exportation, sont en amont, des entreprises à forte intensité de travail qualifié et de capital, plus productives, plus grandes et plus disposées à payer des salaires élevés (Bernard et al., 2007). Ce constat suggère l'existence d'une sorte de sélection pour les activités d'exportation, et nous donne une première intuition du lien qu'il peut y avoir entre survie et exportation. Toutefois, même si les entreprises exportatrices sont préalablement les plus performantes, elles bénéficient quand même des activités d'exportation. Bernard & Jensen (1999) travaillent sur des don-

nées américaines et trouvent que les entreprises exportatrices ont 10% de chance de plus de survivre comparativement aux entreprises non exportatrices. Cet effet positif est majoritairement confirmé dans la littérature. [Kimura & Fujii \(2003\)](#) trouvent, à partir de données japonaises, que l'exportation impacte positivement la survie des entreprises. [Pérez et al. \(2004\)](#) et [Esteve-Pérez & Mañez-Castillejo \(2008\)](#) obtiennent les mêmes résultats en travaillant sur des données espagnoles.

Dépenses en Publicité et Profitabilité

Les variables précédemment énumérées font partie des caractéristiques les plus mises en avant dans les travaux empiriques. Toutefois, elles ne constituent pas une liste exhaustive des déterminants de la survie spécifiques aux entreprises. En effet, d'autres variables de nature financière comme les dépenses en publicité et le profit, ont été prises en compte dans diverses études.

Les travaux de [Esteve-Pérez & Mañez-Castillejo \(2008\)](#) montrent que les dépenses en publicité ont un effet positif sur la survie des entreprises, avec notamment un différentiel de risque d'échec de 71% en faveur des entreprises qui investissent dans la publicité. La pérennité d'une entreprise dépend, entre autres, de sa capacité à maintenir une base de clientèle fidèle, et à conquérir de nouvelles parts de marché. L'entreprise a donc grandement intérêt à construire et entretenir une bonne réputation, et en la matière, la publicité s'avère être un outil important. Grâce à la publicité, l'entreprise peut non seulement améliorer son image, faire la promotion de nouveaux produits, mais aussi dissuader de potentiels concurrents en affirmant sa position de leader sur le marché. La capacité d'une entreprise à générer du profit impacte de même positivement sa survie ([Fotopoulos & Louri, 2000](#); [Nogueira et al., 2018](#)). La profitabilité est un indicateur de performance qui est souvent la conséquence de la productivité ou d'une bonne maîtrise des coûts. Ainsi, un profit élevé présage toutes choses égales par ailleurs une longévité sur le marché.

Dans la sous-section suivante, nous présentons les caractéristiques de l'industrie qui influencent la survie des entreprises.

1.1.2.2 Caractéristiques de l'industrie

Les secteurs d'activité d'une économie sont généralement très hétérogènes et caractérisés par des flux entrants et sortants d'entreprises. Cette dynamique industrielle est entre autres le résultat de la concurrence entre les entreprises, qui recherchent chacune un posi-

tionnement avantageux sur le marché. Les nouveaux entrants, arrivés dans l'industrie pour se faire une part de marché, deviennent une menace pour les entreprises déjà établies, et ces dernières doivent désormais mettre tout en œuvre pour conserver leur position. Ce jeu de la concurrence s'intensifie avec le temps et selon les industries, et certaines entreprises ne réussissent pas à y survivre. Les travaux de [Mata & Portugal \(1994\)](#) et de [Brown & Earle \(2000\)](#) montrent justement que les entreprises opérant dans des industries caractérisées par un taux élevé de nouveaux entrants et une forte concurrence, ont des durées de vies réduites. D'une certaine manière, ces résultats, bien qu'étant intuitifs, tendent à présenter la concurrence comme quelque chose de systématiquement nocif alors qu'en terme de bien-être social, la concurrence peut être bénéfique.

Une industrie caractérisée par très peu de concurrence, est une industrie concentrée entre les mains d'un petit groupe de producteurs susceptibles d'abuser de ce pouvoir vis-à-vis des consommateurs et des producteurs ayant moins de pouvoir de marché. Dans ce type de contexte, les entreprises ayant de grandes parts de marchés peuvent facilement recourir à des stratégies beaucoup plus agressives afin d'évincer plus facilement les petits producteurs. [Audretsch & Mahmood \(1995\)](#) démontrent l'existence d'une relation négative entre la survie des entreprises et la valeur de l'indice de Lerner. L'indice de Lerner qui est la marge sur le coût marginal, mesure le pouvoir du monopole et est d'autant plus élevé qu'il a y moins de concurrence. Une industrie concentrée aux mains de quelques entreprises augmente donc le risque d'échec des nouveaux entrants. Ce résultat est corroboré par [Resende et al. \(2016\)](#) qui trouvent que l'indice de Herfindahl réduit la probabilité de survie des entreprises. Le degré de concurrence d'une industrie peut ainsi augmenter ou réduire les chances de survie des entreprises. Par ailleurs, certaines caractéristiques de l'industrie comme la dynamique de la taille et l'échelle minimale d'efficacité affectent significativement la survie des entreprises dans un seul sens et apparaissent systématiquement comme des atouts ou des obstacles.

[Mata & Portugal \(1994\)](#); [Resende et al. \(2016\)](#) trouvent, en travaillant respectivement sur des données portugaises et brésiliennes, qu'une industrie fortement croissante (en termes d'emploi) augmente les chances de survie des nouvelles entreprises. Ainsi, contrairement à une industrie stable laissant peu de place aux nouveaux entrants, une industrie dynamique offre un environnement moins hostile et plus adéquat aux jeunes entreprises.

L'échelle minimale d'efficacité est l'une des principales causes de l'échec des entreprises dans une industrie. Bien qu'elle soit souvent mesurée de différentes manières par les chercheurs, le sens négatif de son impact fait presque toujours l'unanimité. Une industrie

caractérisée par une échelle minimale d'efficacité très élevée est un environnement hostile pour les entreprises. [Audretsch \(1991\)](#) mesure l'échelle minimale d'efficacité comme la taille moyenne des plus grandes entreprises pesant la moitié de l'industrie et confirme l'effet négatif sur la survie. [Fritsch et al. \(2006\)](#) trouvent de même un impact négatif en mesurant l'échelle minimale d'efficacité comme le troisième quartile de la distribution de la taille des établissements. [Resende et al. \(2016\)](#) en viendront à la même conclusion en approxinant l'échelle minimale d'efficacité par la taille médiane des entreprises étudiées.

L'industrie est une sorte de paysage économique pour l'entreprise, et les différents travaux que nous venons de présenter, mettent en évidence l'importance des caractéristiques sectorielles dans la survie de l'entreprise. La section qui suit discute du lien entre l'environnement géographique et la survie de l'entreprise.

1.1.2.3 Caractéristiques de l'environnement géographique

Bien que l'on ne saurait étudier la survie d'une entreprise en faisant abstraction des caractéristiques de son environnement géographique, la littérature incorporant cet aspect est moins abondante que celle prenant en compte les caractéristiques de l'entreprise et de l'industrie. Toutefois, de récents travaux se sont penchés sur la question. Ces travaux sont essentiellement empiriques car les théories d'économie spatiale sur la question décrivent principalement le processus du choix de localisation des entreprises sans mettre d'accent sur la survie.

La prise en considération de la localisation géographique a débuté à une échelle plutôt agrégée en ce sens que les premiers travaux s'intéresseront à l'impact du type de zone (urbaine ou rurale) sur la survie des entreprises. Ainsi, [Fotopoulos & Louri \(2000\)](#) concluront que les entreprises localisées en zones urbaines ont une plus grande chance de survie comparativement à celles qui sont en zones rurales. Cela suggère l'idée que les entreprises localisées en zones urbaines bénéficieraient d'économies d'agglomération liées à l'intensité de l'activité économique. Cet impact positif qui s'apparente à ce que [Perraud et al. \(2008\)](#) appellent «effet offre de ressources», a en revanche un opposé appelé «effet de concurrence». En effet, les régions urbaines en raison des possibilités qu'elles offrent jouissent d'une très forte attractivité susceptible d'induire une concurrence exacerbée à laquelle certaines entreprises peuvent avoir du mal à survivre. [Huiban \(2011\)](#) travaille sur un échantillon de 300 000 établissements français et trouve qu'il est plus facile de créer une entreprise en zones urbaines, mais que les chances de survie y sont faibles comparativement aux zones rurales. Ce résultat met en exergue l'«effet de concurrence». Au delà de la concurrence, les zones at-

tractives peuvent aussi entraîner des coûts élevés (loyers et salaires). Les entreprises doivent ainsi s'assurer de leur capacité à faire face à la concurrence, et à maîtriser leurs coûts de production avant de s'installer dans de telles zones. Les travaux qui viendront plus tard caractériseront plus finement l'environnement géographique.

Staber (2001) trouve que le fait d'avoir beaucoup d'entreprises de la même industrie dans son voisinage (rayon de 10 km) réduit les chances de survie contrairement au fait d'être à proximité d'entreprises de secteurs diversifiés. L'auteur explique que ses résultats suggèrent d'une part, l'idée d'une très forte concurrence autour des ressources, et d'autre part l'existence d'externalités positives issues de la diversification (externalités de Jacobs (1969)). De façon générale, les avantages (externalités) liés à la concentration spatiale d'activités économiques sont regroupés en trois catégories (Dubé et al., 2016). On distingue les externalités d'urbanisation, les externalités de localisation et les externalités de diversification ou de Jacobs (1969). Les externalités d'urbanisation sont les avantages liés au fait d'être localisé dans une grande ville (F. Neffke et al., 2011). En effet, les grandes villes sont caractérisées par une abondance de ressources (marché de grande taille, main d'œuvre qualifiée, infrastructures) susceptibles de promouvoir la prospérité des entreprises. Les externalités de localisation encore appelées externalités MAR du fait des travaux de pionnier de Marshall (1890); Arrow (1962); Romer (1986), découlent de la concentration spatiale d'entreprises issues d'une même industrie. Les entreprises localisées dans ce type d'agglomération bénéficieront donc entre autres d'une baisse des coûts de transport, d'une diffusion des connaissances et d'un partage de ressources. Quant aux externalités de diversification, elles proviennent de la concentration d'activités diversifiées contrairement aux externalités MAR. Jacobs (1969) soutient que la diversification offre aux entreprises l'opportunité de combiner les connaissances et d'aboutir à des idées novatrices. Ces trois externalités vont s'avérer déterminantes dans la survie des entreprises.

F. M. Neffke et al. (2011) travaillent sur échantillon de 25 000 établissements suédois observés de 1970 à 2004. Pour mesurer le degré externalités de Jacobs (1969) auquel une entreprise est exposée, les auteurs utilisent le nombre d'entreprises issue d'une industrie différente. Quant aux externalités de localisation, elles sont, pour chaque entreprise dans une municipalité donnée, approximées par le niveau d'emploi dans la même industrie (ou dans une industrie connexe). Les résultats indiquent que les externalités de Jacobs (1969) impactent positivement la survie des entreprises indépendantes contrairement aux externalités d'urbanisation (niveau de population des municipalités) qui ont une influence négative. Les auteurs n'ont trouvé guère d'effets significatifs en ce qui concerne les externalités de localisation. Cainelli et al. (2014) en viennent à des résultats différents en travaillant

sur des données italiennes. Alors que [F. M. Neffke et al. \(2011\)](#) observent l'entreprise dans sa municipalité d'implantation, [Cainelli et al. \(2014\)](#) mènent leur analyse à une échelle beaucoup plus agrégée. Ils se focalisent sur le taux de survie des entreprises par industries par provinces italiennes sur la période 1998-2007. L'estimation par la méthode des moments généralisés d'un modèle de panel dynamique spatial permet aux auteurs de prendre en compte les dépendances spatiales et les effets décalés du taux de fermeture des entreprises. Leurs résultats montrent que les externalités de localisations réduisent le taux de fermeture des entreprises. En revanche, les externalités de [Jacobs \(1969\)](#) n'ont pas d'effets significatifs. [Arcuri et al. \(2019\)](#) travaillent à une échelle similaire à [Cainelli et al. \(2014\)](#) et estiment également un modèle de panel dynamique spatial. Ils observent les taux de fermeture des entreprises dans 96 départements français entre 2009 et 2013. Leurs résultats d'estimation suggèrent que les externalités de localisation n'ont aucune influence significative sur le taux de fermeture des entreprises. [Basile et al. \(2017\)](#) trouvent des résultats qui vont qualitativement dans le même sens que ceux trouvés par [F. M. Neffke et al. \(2011\)](#) quant aux économies liées à la diversification. Leurs estimations montrent que les externalités de [Jacobs \(1969\)](#) réduisent le risque d'échec des entreprises. Ils trouvent également que les externalités de localisation affectent positivement la survie des entreprises seulement dans le secteur des services alors que les externalités d'urbanisation n'ont pas d'effets robustes. [Ebert et al. \(2019\)](#) se basent sur un panel de 6 776 entreprises allemandes observées entre 2007 et 2011 et montrent que les externalités de localisation ne profitent qu'aux entreprises qui opèrent dans une industrie non-technologique et qui de surcroît innovent moins. Les auteurs ne trouveront aucun effet significatif pour les différentes mesures de la diversité.

Hormis les économies ou déséconomies d'agglomération, d'autres mesures d'ordre général sont également utilisées pour caractériser l'environnement local de l'entreprise. [Perraud et al. \(2008\)](#) incluront par exemple dans leur travail des variables approximant l'état des infrastructures de transport et de tourisme dans la zone où l'entreprise est implantée. Ces infrastructures constituent des ressources importantes dont l'accès peut améliorer la compétitivité des entreprises. Dans le même temps, les territoires possédant de telles ressources peuvent aussi être sujets à une forte concurrence entre les entreprises. Dans l'ensemble, le travail de [Perraud et al. \(2008\)](#) ne décèle qu'un faible pouvoir prédictif pour les caractéristiques territoriales dans la survie des entreprises. Une approche utilisant d'autres types de mesures peut être utile pour mieux cerner l'impact de ces variables.

Le taux de chômage ([Huggins et al., 2017](#); [Brixy & Grotz, 2007](#)), et le taux de croissance économique locale ([Huggins et al., 2017](#)) sont de même étudiés. [Huggins et al. \(2017\)](#);

[Brixy & Grotz \(2007\)](#) trouvent un effet négatif du chômage sur la survie des entreprises. Ce résultat s'explique par le fait qu'un chômage élevé caractérise un environnement socio-économique défavorable (faible niveau de vie) où les entreprises ont très peu de chances d'avoir un niveau de demande nécessaire pour prospérer. Inversement, une croissance économique locale élevée impacte positivement la survie des entreprises ([Huggins et al., 2017](#)).

Nonobstant les réponses intéressantes apportées par tous ces travaux, une question subsiste toujours. Au delà des différentes externalités d'agglomération considérées, quels aspects particuliers; quels types d'infrastructures de la zone d'implantation affectent la survie d'une entreprise, toutes choses égales par ailleurs? Les externalités traduisent généralement les défaillances du marché et peuvent être difficilement régulées par les pouvoirs publics. Les autres caractéristiques territoriales telles que les infrastructures sont l'œuvre de politiques publiques. La compréhension de l'impact de ces caractéristiques sur la survie des entreprises peut conduire à des actions de politiques publiques permettant de créer des environnements favorables aux entreprises. Ainsi, nous proposons dans notre analyse de la survie, de mettre l'accent sur le niveau de dotation des territoires dans un certain nombre d'équipements.

Nous commençons par une analyse économétrique portant sur 1 651 798 établissements français observés de 1995 à 2005. A l'issue de cette analyse, nous proposons d'étudier à nouveau la question dans un cadre théorique inspiré de [Schröder & Sørensen \(2012\)](#).

1.2 Analyse économétrique

Dans cette section, nous présentons successivement les modèles économétriques estimés, les données et variables utilisées, puis les résultats d'estimation.

1.2.1 Modèles économétriques

Dans la littérature, les travaux portant sur la survie des entreprises font appel à diverses méthodologies au nombre desquelles nous avons les moindres carrés ordinaires ([Fritsch et al., 2006](#); [Acs et al., 2007](#); [Perraud et al., 2008](#)), les modèles de choix discrets (probit, logit) ([Helmers & Rogers, 2010](#); [Perraud et al., 2008](#); [Audretsch, 1991](#); [Stearns et al., 1995](#); [Storey & Wynarczyk, 1996](#); [Huiban, 2011](#)) et les modèles de survie ([Staber, 2001](#); [Agarwal & Gort, 2002](#); [Pérez et al., 2004](#); [F. M. Neffke et al., 2011](#); [Mazzotta et al., 2017](#); [Shu & Simmons, 2018](#); [Baumöhl et al., 2019, 2020](#)).

L'utilisation des moindres carrés ordinaires intervient généralement lorsqu'on travaille sur des données agrégées où il est question d'étudier par exemple l'impact des caractéristiques d'une région sur le taux de survie des entreprises de cette région. En ce qui concerne les modèles probit et logit, ils peuvent être utilisés dans le cas d'une analyse de survie sur données individuelles. Toutefois, ils ne permettent pas de prendre en compte les problèmes de censure et la durée de vie des sujets étudiés. Cette limite est cependant levée par les modèles de survie qui s'avèrent plus appropriés lorsque l'information sur la durée de vie des individus est disponible (Kleinbaum & Klein, 2010; Liu, 2012).

Dans le cas de notre étude, nous disposons de données individuelles d'entreprises dont nous avons les informations sur les dates de création et de fermeture (s'il a lieu). Nous avons donc opté pour les modèles de survie.

1.2.1.1 Modèles de survie

L'analyse empirique de ce chapitre s'articule autour de deux principales approches. Une approche descriptive ou non-paramétrique de la survie (méthode de Kaplan-Meier et le test log rank) puis une approche semi-paramétrique (modèles à hasards proportionnels de Cox avec et sans stratification).

Approche non-paramétrique : Méthode de Kaplan-Meier

Les approches non-paramétriques de l'analyse de survie consistent principalement à utiliser les données observées pour calculer la probabilité de survie des individus sans avoir à faire d'hypothèse sur la distribution que suit leur durée de vie. Il existe plusieurs méthodes d'analyse descriptive de la survie, au nombre de celles-ci, nous avons la méthode de Nelson-Aalen et la méthode de Kaplan-Meier pour laquelle nous optons en raison de sa simplicité. Avant de présenter la méthode de Kaplan-Meier, nous introduisons brièvement les fonctions de survie et de hasard.

La fonction de survie $S(t)$ d'un individu, donne la probabilité que cet individu vive au delà d'un instant t . Si T représente la durée de vie de l'individu, alors $S(t) = P(T > t)$ (Kleinbaum & Klein, 2010).

La fonction de hasard $h(t)$ quant à elle, est définie comme le risque d'échec³ encouru par l'individu à chaque instant t (Liu, 2012). Dans le cadre de ce chapitre où l'on s'intéresse

3. L'échec ici correspond à l'événement étudié.

à la survie des entreprises, l'événement étudié (l'échec ou l'état final) est la fermeture, et la population est composée d'entreprises.

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{-d \log(S(t))}{dt}$$

La méthode de Kaplan-Meier permet de calculer les valeurs de la fonction de survie $S(t)$ à chaque instant t , de représenter graphiquement l'évolution de cette probabilité de survie dans le temps (courbe de survie), et d'effectuer des comparaisons entre des catégories d'individus.

Considérons un échantillon de n ($i = 1, 2, \dots, n$) entreprises, ayant chacune une durée de vie observée⁴ t_i . Soit RC une suite croissante des t_i tel que : $RC : t_1 < t_2 < \dots < t_n$. Si on note $\mathcal{R}(t_i)$ l'ensemble des entreprises n'ayant pas encore expérimenté l'échec à l'instant t_i , $t_1 \leq t_2 \Rightarrow \text{Card}(\mathcal{R}(t_1)) \geq \text{Card}(\mathcal{R}(t_2))$. En effet, de t_1 à t_2 , on aurait certaines entreprises qui seraient soit fermées, soit censurées de l'échantillon, ce qui réduit l'ensemble $\mathcal{R}(t_1)$ des d_1 entreprises mortes ou censurées en t_1 , d'où $\text{Card}(\mathcal{R}(t_1)) - d_1 = \text{Card}(\mathcal{R}(t_2))$. Notons n_i le cardinal de $\mathcal{R}(t_i)$, l'estimateur de Kaplan-Meier de $S(t)$ est donné par :

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}$$

Ainsi, il est possible de représenter graphiquement la fonction de survie à partir des valeurs obtenues par la méthode de Kaplan-Meier.

La méthode de Kaplan-Meier, à travers les graphiques, permet également de réaliser des études comparatives⁵ lorsque l'on dispose de plusieurs catégories d'individus. Supposons que nous nous intéressions dans notre analyse descriptive à comparer la fonction de survie des entreprises d'une certaine catégorie A à celle des entreprises d'une certaine catégorie B. Il suffira de calculer les fonctions de survie par la méthode de Kaplan-Meier, de les représenter sur le même graphe et de voir la position de l'une par rapport à l'autre. Cependant il est important de souligner que l'éventuelle différence observée entre les courbes n'est pas systématiquement significative, il faut donc un test statistique. De tous les tests [log Rank, Wilcoxon, Tarone-Ware, Peto, Flemington-Harrington, (Mantel & Haenszel, 1959; Gehan, 1965; Peto & Peto, 1972; Tarone & Ware, 1977; Prentice, 1978)] qui permettent d'évaluer la significativité des dissemblances observées entre des courbes de survie, le test «log rank»

4. La durée de vie observée représente la durée de vie réelle des entreprises ayant fermé durant l'étude. Quant aux autres, il s'agit de la durée de vie censurée.

5. Sans toutefois établir une causalité.

est l'un des plus simples à mettre en œuvre. Les travaux de [Kleinbaum & Klein \(2010\)](#) et de [Liu \(2012\)](#), présentent tous ces tests de manière détaillées.

Approche semi-paramétrique : Modèle à hasards proportionnels de Cox

L'approche non paramétrique présentée ci-dessus s'inscrit beaucoup plus dans le cadre d'une analyse descriptive où l'on ne peut considérer qu'une seule variable à la fois. Pour mieux identifier les déterminants de la survie des individus étudiés, il conviendrait d'utiliser une approche permettant de considérer simultanément l'influence de plusieurs variables. Nous utilisons ainsi le modèle de [Cox \(1972\)](#).

Le modèle à hasards proportionnels de [Cox \(1972\)](#) propose une façon de contourner une limite inhérente aux modèles de survie paramétriques. Dans l'approche paramétrique des modèles de durée, l'on est amené à faire des hypothèses sur la distribution que suit la durée de vie T . Bien qu'il existe des méthodes graphiques permettant de conjecturer la distribution de T , les décisions issues de ses vérifications restent toujours un peu subjectives. Ainsi, il y a toujours des risques d'erreurs susceptibles de conduire à des biais dans les paramètres estimés. Ces limites de l'approche paramétrique conduisent [Cox \(1972\)](#) à proposer un modèle à hasards proportionnels semi-paramétrique appelé **modèle à hasards proportionnels de Cox**. Le modèle de Cox ne nécessite guère d'hypothèses sur la distribution de la durée de vie T , c'est en cela qu'il est un modèle semi paramétrique. En effet, [Cox \(1972\)](#) propose une façon d'écrire la vraisemblance de l'échantillon, qui permet de simplifier les hasards de base h_0 , et donc d'estimer les paramètres sans avoir à définir une forme pour le hasard de base.

Comme dans la plupart des modèles de régression, l'idée est de quantifier l'impact (soit β) d'une ou plusieurs variables explicatives (représentées par une matrice X) sur la durée de vie (soit T) d'un individu. Ici, le taux de hasard conditionnellement aux explicatives s'écrit :

$$h(t | X) = h_0(t)\exp(X\beta)$$

Le taux de hasard étant une quantité positive, le choix de la fonction exponentielle est justifié. $h_0(t)$ est le hasard de base commun à tous les individus. L'estimation du modèle passe par la définition d'une vraisemblance partielle qui ne prend pas directement en compte les observations censurées. L'idée derrière la vraisemblance partielle de Cox consiste à calculer pour un individu A ayant vécu l'événement étudié à l'instant t_A , la probabilité que cet individu expérimente l'événement à l'instant t_A sachant qu'un des individus sous risque en

t_A a expérimenté l'événement. Il s'agit donc d'une probabilité conditionnelle. La vraisemblance de l'échantillon s'écrit principalement par rapport aux individus non censurés. Les données censurées ne sont pas explicitement prises en compte mais elles le sont implicitement. En effet, un individu censuré à l'instant t_g ne contribue certes pas à la vraisemblance en t_g mais appartient à l'ensemble des individus sous risque en t_f ($t_f < t_g$). Ainsi, ce dernier aura été pris en considération dans la vraisemblance de l'individu ayant connu l'événement en t_f .

Considérons à présent un échantillon de n individus. Pour l'individu i ayant expérimenté l'événement en t_i , la vraisemblance L_i est donnée par :

$$L_i = \frac{h_0(t) \exp(X_i \beta)}{\sum_{j \in \mathcal{R}(t_i)} h_0(t) \exp(X_j \beta)} = \frac{\exp(X_i \beta)}{\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta)}$$

Où, comme précédemment défini, $\mathcal{R}(t_i)$ est l'ensemble des individus n'ayant pas expérimenté l'événement en t_i .

Supposons à présent que tous les individus ont des durées de vie bien distinctes, c'est à dire que nous n'avons pas d'observations liées⁶. Pour écrire la vraisemblance $L(\beta)$ des n individus de l'échantillon, les différents instants t_i sont rangés par ordre croissant ($t_1 < t_2 < t_3 < \dots < t_n$), et on obtient :

$$L(\beta) = \prod_{i=1}^n \left[\frac{\exp(X_i \beta)}{\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta)} \right]^{\delta_i}$$

où $\delta_i = 0$ si l'observation est censurée et 1 sinon

$$\text{Pour } \delta_i = 0; \text{ on a } \prod_{i=1}^{i=n} \left[\frac{\exp(X_i \beta)}{\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta)} \right]^{\delta_i} = 1$$

$$\text{On peut donc écrire } L(\beta) = \prod_{i=1}^{n_1} \frac{\exp(X_i \beta)}{\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta)}$$

Où n_1 est le nombre d'individus non censurés.

$$\log L(\beta) = \sum_{i=1}^{n_1} \left\{ X_i \beta - \log \left[\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta) \right] \right\}$$

le score $S(\beta)$ est donné par :

6. Individus ayant vécu l'événement au même instant.

$$S(\beta) = \frac{\partial \log L(\beta)}{\partial \beta} = \sum_{i=1}^{n_1} \left\{ X'_i - \left[\frac{\sum_{j \in \mathcal{R}(t_i)} X'_j \exp(X_j \beta)}{\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta)} \right] \right\} = 0$$

L'équation $S(\beta) = 0$ n'admettant pas analytiquement de solution, l'algorithme de Newton-Raphson est souvent utilisé avec une valeur initiale $\beta^0 = 0$ (Liu, 2012). Lorsque la vraisemblance remplit certaines conditions de régularités (Ruud, 2000) et que les estimateurs $\hat{\beta}$ sont convergents et asymptotiquement normalement distribués, les tests usuels de significativités peuvent être aisément construits.

La matrice de variances-covariances (soit $\Omega(\hat{\beta})$) est obtenue en inversant la matrice d'information de Fisher (soit $I(\hat{\beta})$) qui est une matrice de dimension $K \times K$ où K est le nombre de paramètres.

$$\Omega(\hat{\beta}) = [I(\hat{\beta})]^{-1}$$

L'élément de la k ième ligne de la k' ième colonne de la matrice d'information de Fisher se calcule par : $-\frac{\partial^2 \log L(\beta)}{\partial \beta_k \partial \beta_{k'}}$. On obtient :

$$\left\{ \frac{\left[\sum_{j \in \mathcal{R}(t_i)} X'_{jk} X'_{jk'} \exp(X_j \beta) \right] \left[\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta) \right] - \left[\sum_{j \in \mathcal{R}(t_i)} X'_{jk'} \exp(X_j \beta) \right] \left[\sum_{j \in \mathcal{R}(t_i)} X'_{jk} \exp(X_j \beta) \right]}{\left[\sum_{j \in \mathcal{R}(t_i)} \exp(X_j \beta) \right]^2} \right\}$$

Interprétation des résultats

Dans les modèles à hasards proportionnels le signe des coefficients peut être directement interprété comme étant le sens (positif ou négatif) de l'impact des explicatifs sur le hasard parce que les coefficients ont le même signe que les effets marginaux. En ce qui concerne les valeurs des paramètres estimés, pour les interpréter on utilise le ratio des hasards noté HR . Le ratio de hasards entre deux individus A et B pour une variable explicative k (toutes choses égales par ailleurs) prenant respectivement les valeurs x_k^A et x_k^B donne :

$$\begin{aligned} HR(x_k^A, x_k^B) &= \frac{h_0(t) \exp(\beta_0 + \dots + \beta_k x_k^A + \dots + \beta_K x_K)}{h_0(t) \exp(\beta_0 + \dots + \beta_k x_k^B + \dots + \beta_K x_K)} \\ HR(x_k^A, x_k^B) &= \exp[\beta_k (x_k^A - x_k^B)] \end{aligned}$$

Dans le cas d'une variable binaire, $HR(x_k^A, x_k^B) = \exp[\beta_k(1-0)] = \exp(\beta)$ et s'interprète comme le hasard des individu A relativement à celui des individus B . Dans le cas de variables continues, le HR indique la proportion de hasard associé à une variation d'une unité de x_k . Ainsi, $HR > 1$ indiquerait un hasard élevé pour les individus A et $HR < 1$ indiquerait un hasard faible pour les individus A .

1.2.2 Présentation des données et des variables

Données

L'analyse économétrique de ce chapitre porte sur un échantillon de 1 651 798 établissements français suivis de 1995 à 2005 et issus du champs I.C.S (industrie, construction, commerce ou services hors agriculture et activités financières). Le champs I.C.S est le plus large des statistiques d'entreprises et représente une proportion prépondérante de la valeur ajoutée de l'économie française⁷.

Un établissement relève du champs I.C.S s'il :

- appartient à une entreprise au sens de la démographie des entreprises
- exerce une activité économique réelle
- exerce une activité économique à caractère marchand
- exerce une activité dans les secteurs de l'industrie, de la construction, du commerce et des services, hors agriculture et activités financières.

Ces établissements sont répertoriés dans les bases de données REE (Répertoire des Entreprises et des Établissements) qui sont produites par l'INSEE. Ces bases dénombrent les entreprises et les établissements exploitants du champs marchand non agricole actifs au premier janvier de chaque année, et nous fournissent des informations relatives à l'identification de l'établissement, sa localisation et quelques unes de ses données économiques. Ainsi, nous avons des informations telles que le numéro d'identification de l'établissement, sa catégorie juridique, sa date de création, sa localisation (Région, département, commune, zone d'emploi), son activité principale, son effectif, les caractéristiques de l'entrepreneur (sexe, âge, date de naissance) pour les entreprises individuelles, et le statut de l'établissement (siège ou non). Il est important de noter que ces informations ne sont pas toujours disponibles d'une année à une autre et peuvent être manquantes au niveau de plusieurs établissements. Afin donc d'avoir plus d'informations sur les établissements suivis à partir des REE, nous avons également utilisé les bases SIRENE (Système Informatique pour le

7. Les entreprises en France, édition 2017- Insee Références.

Répertoire des Entreprises et Établissements).

Les bases SIRENE répertorient l'intégralité des entreprises, et établissements français actifs ou non, et cela correspond à un total d'environ 20 millions d'entreprises et 28 millions d'établissements. En raison de la grande quantité d'informations qu'apportent les bases SIRENE, elles nous auraient difficilement permis de cibler avec précision les entreprises du champs I.C.S, d'où l'avantage des REE. L'apport principal des bases SIRENE a été de nous permettre d'obtenir les dates de création et de fermeture des établissements, parfois manquantes dans les REE.

À ces deux bases, nous avons joint la base Déclaration Annuelle des Données Sociales (DADS) de 1995 pour avoir la taille précise des établissements au début de l'étude. La DADS est produite par l'INSEE et contient des informations relatives à l'établissement et aux salariés. Nous avons donc pour chaque salarié des données sur la nature de son emploi, sa qualification, les heures effectuées, la condition (temps partiel ou complet) et les rémunérations.

Les variables explicatives caractérisant le territoire d'implantation de l'établissement sont au centre de ce chapitre. Pour commencer, nous avons téléchargé une version de la base SIRENE contenant les adresses géocodées du stock des établissements, ce qui nous a permis d'avoir une localisation précise (longitude et latitude) des établissements de notre échantillon. Cette information nous a permis de calculer la distance qui sépare chaque établissement de la capitale de sa région. Cela peut être considéré comme une mesure du degré d'exposition aux externalités d'agglomération. Nous avons également utilisé l'Atlas des zones d'emploi, la Base Permanente des Équipements (BPE) de l'INSEE, les bases ADMIN EXPRESS et GEOFLA de l'Institut National de l'Information Géographique et Forestière.

Une zone d'emploi est définie comme «un espace géographique à l'intérieur duquel la plupart des actifs résident et travaillent, et dans lequel les établissements peuvent trouver l'essentiel de la main d'œuvre nécessaire pour occuper les emplois offerts.» (INSEE⁸). En 2010, les délimitations font état de 322 zones d'emploi sur toute l'étendue du territoire. L'Atlas des zones d'emploi nous a permis d'avoir des informations relatives aux caractéristiques démographiques de la population des zones d'emploi (taux d'évolution de la population, taux de dépendance économique, indicateur de jeunesse), à leur niveau de revenu et à l'emploi (taux d'emploi et taux de chômage).

La base permanente d'équipements s'appuie sur une multitude de sources pour nous renseigner sur le niveau d'équipements et de services rendus à la population sur un ter-

8. <https://www.insee.fr/fr/information/2114596>

ritoire. Les sources de la BPE sont : le Fichier National des Établissements Sanitaires et Sociaux du ministère chargé de la santé (FINESS), le Répertoire Académique et Ministériel Sur les Établissements du Système Éducatif (RAMSESE), le Répertoire Partagé des Professionnels de Santé (RPPS), le Système Informatisé du Répertoire des Entreprises et des Établissements (SIRENE), le Recensement des Équipements Sportifs (RES), Gendarmerie nationale (Ministère de la Défense), enquêtes HÔTEL et CAMPING (INSEE), Préfecture de Police de Paris, Direction Générale de l'Enseignement et de la Recherche (Ministère de l'Agriculture et de l'Alimentation), Système de Référence Justice (Ministère de la Justice), La Poste, Centre national des arts du cirque, de la rue et du théâtre (ARTCENA), Pôle Emploi, base du Ministère de l'Économie pour les carburants, Direction Générale de l'Aviation Civile (DGAC), fichier SNCF (Société Nationale des Chemins de fer Français). Dix grandes catégories d'équipements sont distinguées : Sport, loisirs et culture, Commerce, Services aux particuliers, Actions sociales; Services de santé; Fonctions médicales et paramédicales, Tourisme et transports, Enseignement du premier degré, Enseignement du second degré, Enseignement supérieur, formation et services de l'éducation. L'annexe 1.A apporte plus de détails sur chacune de ces catégories.

Grâce aux informations apportées par cette base, nous avons calculé le nombre d'équipements par commune rapporté à la densité de population dans la commune.

Les bases ADMIN EXPRESS et GEOFLA de l'Institut National de l'Information Géographique et Forestière nous offrent une description détaillée du découpage administratif du territoire métropolitain et ultra-marin. Ces jeux de données nous ont, dans un premier temps permis d'avoir des informations sur les caractéristiques géographiques à l'échelle communale (superficie); et dans un deuxième temps, elles nous ont permis d'effectuer des analyses cartographiques.

Les 1 651 798 établissements observés sur la période 1995-2005 résultent d'un traitement des données. Les bases REE sont constituées de plusieurs jeux de données organisées en deux séries : une série de 1987 à 2000, et une autre série à partir de 2000. Nous avons ainsi fusionnées les différentes tables à partir des identifiants de chaque établissements et de l'année. La base obtenue après la fusion, a été jointe à la base Stock d'établissements SIRENE afin d'avoir les informations nécessaires pour le calcul de la durée de vie des établissements. Nous avons ainsi obtenu une base avec plusieurs dizaines de millions d'observations car nous avons retenu toutes les tables REE disponibles. Cela a engendré des difficultés liées à la taille de l'échantillon (puissance de calcul), et une quantité importante d'informations manquantes. Nous avons donc progressivement réduit l'échantillon. Pour

ce faire, nous avons commencé par limiter la plage temporelle à 1995-2005, et avons ensuite exclu tous les établissements opérants à l'étranger car nous n'avons d'informations que sur les caractéristiques des territoires français. Après l'ajout des données issues des bases BPE, Atlas des zones d'emploi, ADMIN EXPRESS et GEOFLA, nous avons supprimé tous les établissements pour lesquels les variables sélectionnées n'étaient pas renseignées. Ceci nous a conduit à notre échantillon final de 1 651 798 établissements.

Variables

Les différentes sources de données nous ont permis de retenir ou de construire un certain nombre d'indicateurs utilisés comme variables explicatives dans les modèles estimés (voir tableau 1.1 pour la liste). À l'échelle des établissements, les bases mobilisées ne fournissent pas beaucoup d'informations. La plupart de nos variables sont donc à l'échelle de la commune, de la zone d'emploi et du secteur d'activité.

Les variables issues des bases et retenues telles quelles, sont les suivantes : le taux de dépendance économique de la zone d'emploi (**taux_dep_eco_09**) le taux de chômage de la zone d'emploi (**tx_chom_t1_03**), la taille de l'établissement en terme d'emploi (**size_95**). Toutes les autres variables incluses dans notre modèles sont construites.

Les variables **taux_dep_eco_09** et **tx_chom_t1_03** viennent de l'atlas des zones d'emploi 2010 publié conjointement par l'INSEE, la Délégation interministérielle à l'aménagement du territoire et à l'attractivité régionale (DATAR), et La Direction de l'Animation de la Recherche, des Études et des Statistiques (Dares). Ces variables nous permettent de mesurer la disponibilité de la main d'œuvre dans la zone d'emploi où l'établissement est implanté. La variable **Taux_dep_eco_09** est calculée à partir du recensement de la population de 2009 de l'INSEE et mesure la proportion d'individus de moins de 20 ans et de plus de 60 ans par rapport aux individus en âge de travailler (20 à 59 ans). **Tx_chom_t1_03** est le taux de chômage localisé au premier trimestre 2003 estimé sur la base des enquêtes emploi 2009 et le recensement de la population 2002 et 2007 de l'INSEE. La variable **size_95** est obtenue grâce aux DADS et représente la taille de l'établissement au début de l'étude.

Le taux de dépendance économique (**Taux_dep_eco_09**) est pris en compte dans notre analyse car il peut être considéré comme un indicateur de demande potentielle et de main d'œuvre disponible. Une zone d'emploi avec un taux de dépendance économique élevé est une zone dont la population est majoritairement composée de personnes non éligibles au marché de travail en raison de leurs jeunes âges (moins de 20 ans), ou de leurs âges avancés

(plus de 60 ans). Les entreprises implantées dans de telles zones peuvent être confrontées à des risques de fermeture plus élevés à cause d'un manque de main d'œuvre et d'une faible demande.

Quant au taux de chômage, il a déjà été pris en compte dans plusieurs travaux sur la survie des entreprises. [Brixy & Grotz \(2007\)](#) et [Huggins et al. \(2017\)](#) travaillent respectivement sur des données allemandes et anglaises et trouvent que le chômage impacte négativement la survie des entreprises. Un taux de chômage élevé reflète un déséquilibre du marché de travail, et plus généralement un environnement économique peu favorable. Ainsi, une zone d'emploi caractérisée par un niveau élevé de chômage est plus susceptible de limiter les chances de survie des entreprises.

Concernant la taille de l'établissement, c'est l'une des caractéristiques les plus considérées dans les travaux sur la survie (voir [Fotopoulos & Louri, 2000](#); [Staber, 2001](#); [Cefis & Marsili, 2005](#); [Huiban, 2011](#); [F. M. Neffke et al., 2011](#); [Mas-Verdú et al., 2015](#); [Iwasaki & Kočenda, 2020](#)). Dans la majorité des cas, la taille est positivement liée à la survie de l'entreprise car les entreprises de grande taille ont plus les moyens d'atteindre des niveaux de production permettant de bénéficier des économies d'échelle.

Les trois variables que nous venons de décrire sont les seules prises telles quelles dans les bases de données. Toutes les autres variables utilisées sont construites.

La variable **dpop16** mesure la densité de population par commune en 2016. Elle est calculée en divisant la population par la superficie en kilomètre carré :

$$\text{dpop16} = \frac{\text{Population communale}}{\text{Superficie de la commune en km}^2}$$

La densité de la population est un indicateur du niveau d'agglomération des communes ([Brixy & Grotz, 2007](#)). Elle permet de distinguer les zones urbaines des zones rurales ([Stearns et al., 1995](#); [Huiban, 2011](#)). L'impact de la densité de population sur la survie des entreprises a déjà été évalué dans la littérature et les résultats montrent deux types d'effets. D'un côté, on trouve un effet négatif de la densité sur la survie ([Brixy & Grotz, 2007](#)) et de l'autre côté on trouve un effet positif ([Wennberg & Lindqvist, 2010](#)). Les entreprises localisées dans des régions densément peuplées (grandes agglomérations) peuvent, bénéficier d'un accès privilégié à de nombreuses ressources (mains d'œuvres, infrastructures, externalités), et dans le même temps, être exposées à une concurrence féroce.

Par ailleurs, l'établissement n'a pas besoin d'être dans une zone urbaine pour être exposée aux externalités d'agglomération, il lui suffit juste d'en être à proximité. En effet, la va-

riable densité de population étant calculée à l'échelle communale, ne prend pas en compte le fait que certaines communes très peu denses peuvent se situer très proches de grandes agglomérations. Pour prendre cet aspect en considération, nous avons calculé la distance qui sépare chaque établissement de la capitale de sa région (**distcaptl_km**). Le calcul est effectué par la méthode de Pythagore en utilisant les coordonnées géographiques des établissements et des capitales régionales. Les coordonnées étant au départ en longitude et latitude, elles ont d'abord été exprimées en degrés décimaux.

$$\text{distcaptl_km} = 1,852 \times 60 \times \sqrt{\left(\left[(L_r - L_i) \times \cos\left(\frac{l_r + l_i}{2}\right) \right]^2 + [l_r - l_i]^2 \right)}$$

où L et l sont respectivement les longitudes et latitudes des établissements i et capitales régionales r .

Nous avons également calculé le nombre d'établissements par km^2

$$\text{d_km2_nbredetabs} = \frac{\text{Nombre d'établissements dans la commune}}{\text{Superficie de la commune}}$$

puis l'indice de concentration sectorielle de Herfindahl par zone d'emploi nommé **h_index** et calculé avec le nombre d'établissements. Pour une zone d'emploi Z on a :

$$\text{h_index}_Z = \sum_S (\alpha_Z^S)^2$$

où α_Z^S est la proportion d'établissements du secteur S dans la zone d'emploi Z . Nous avons utilisé la nomenclature agrégée en 10 postes donc **h_index** prend ses valeurs entre $\frac{1}{10}$ (répartition équitable) et 1 (parfaite spécialisation).

Les travaux de [Mata & Portugal \(1994\)](#); [Audretsch \(1995\)](#); [Baumöhl et al. \(2020\)](#) mettent en évidence le fait que les chances de survie des entreprises varient d'un secteur (ou industrie) à un autre. Cela a motivé la construction de dummies sectoriels calculés à partir de la Nomenclature des Activités Françaises (NAF) agrégées en 10 postes.

TABLE 1.1 – Codes et libellés des variables

Codes	Libellés
taux_dep_eco_09	Taux de dépendance économique de la zone d'emploi
tx_chom_t1_03	Taux de chômage de la zone d'emploi
dpop16	Densité de population dans la commune
atsoc_p_km2	Nombre d'équipements en actions sociales, pour 100 personnes par km ² , par commune
cmrce_p_km2	Nombre d'équipements dans le commerce pour 100 personnes par km ² , par commune
ens1erdeg_p_km2	Nombre d'équipements dans l'enseignement au 1er degré pour 100 personnes par km ² , par commune
ens2ndeg_p_km2	Nombre d'équipements dans l'enseignement au 2nd degré pour 100 personnes par km ² , par commune
ensupformautr_p_km2	Nombre d'équipements dans l'enseignement supérieur formation et services d'éducation, pour 100 personnes par km ² , par commune
medparmed_p_km2	Nombre d'équipements, dans les services médicaux et paramédicaux, pour 100 personnes par km ² , par commune
sante_p_km2	Nombre d'équipements dans les services de santé, pour 100 personnes par km ² , par commune
servparticul_p_km2	Nombre d'équipements dans les services aux particuliers, pour 100 personnes par km ² , par commune
sprtlsrctult_p_km2	Nombre d'équipements dans le sport les loisirs et la culture, pour 100 personnes par km ² , par commune
trsptrsm_p_km2	Nombre d'équipements dans le transport et le tourisme pour 100 personnes par km ² , par commune
sect_be	Dummy prenant la valeur 1 pour les établissements du secteur de l'industrie manufacturière, extractive et autre
sect_fz	Dummy prenant la valeur 1 pour les établissements du secteur de la construction
sect_gi	Dummy prenant la valeur 1 pour les établissements du secteur du commerce de gros et de détail, transport, hébergement et restauration
sect_mn	Dummy prenant la valeur 1 pour les établissements du secteur des activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien
sect_oq	Dummy prenant la valeur 1 pour les établissements du secteur de l'administration publique, enseignement, santé humaine et actions sociales
distcaptl_km	Distance à la capitale régionale
d_km2_nbredetabs	Nombre d'établissements, par km ² , par commune
h_index	Proxy de l'indice de Herfindahl
et_siege	Dummy prenant la valeur 1 pour les établissements qui sont le siège de l'entreprise
size_95	Taille de l'établissement en 1995

La NAF est créée en 1993 et successivement révisée en 2003 (NAF rév 1) et en 2008 (NAF rév 2). Grâce aux tables de passage des NAF élaborées par l'INSEE, nous avons réussi à harmoniser les secteurs d'activités des établissements de l'échantillon suivant la version NAF rév 2.

L'une des particularités de notre approche est que nous proposons une fine caractérisation de l'environnement spatial de l'établissement grâce aux données des bases BPE. Nous avons construit au niveau communal, un certain nombre de variables mesurant le nombre d'équipements pour 100 personnes par km² (voir tableau 1.1). Ces variables, considérées comme des indicateurs de la dotation de chaque commune en diverses infrastructures, peuvent jouer un rôle déterminant dans le succès d'un projet d'entreprise. Lorsqu'un établissement s'implante dans un nouvel environnement spatial, les caractéristiques de ce dernier peuvent influencer l'établissement à deux niveaux : les affaires (les activités) et l'humain (le personnel). Dans un premier temps, il faut que l'environnement offre à l'établissement de bonnes opportunités d'affaires et des possibilités d'exercer aisément son activité. Ensuite, il est aussi important que le personnel de l'établissement soit épanoui, d'où l'importance de certains équipements relatifs aux conditions de vie.

Ainsi, pour un équipement e dans une commune c on a :

$$100 \times \frac{\left(\frac{N_{ec}}{P_c} \right)}{S_c}$$

où N_{ec} représente le nombre d'équipements e dans la commune c ; P_c la population de la commune c et S_c la superficie de la commune c en km².

Pour faciliter la lecture de nos résultats, le tableau 1.2 établit une correspondance entre les codes des variables et les dénominations utilisées dans les tableaux de résultats.

TABLE 1.2 – Correspondance des variables dans les tableaux de résultat

Codes	Noms dans les tableaux de résultat
taux_dep_eco_09	Dépendance économique
tx_chom_t1_03	Chômage
dpop16	Densité de population
atsoc_p_km2	Équipements en actions sociales
cmmrce_p_km2	Équipements dans le commerce
ens1erdeg_p_km2	Enseignement au 1er degré
ens2nddeg_p_km2	Enseignement au 2nd degré
ensupformautr_p_km2	Enseignement supérieur et autres
medparmed_p_km2	Services médicaux et paramédicaux
sante_p_km2	Services de santé
servparticul_p_km2	Services aux particuliers
sprtlsrscult_p_km2	Sport loisirs et culture
trsptrsm_p_km2	Transport et tourisme
sect_be	Dummy Industrie
sect_fz	Dummy Construction
sect_gi	Dummy Commerce, transport, hébergement et restauration
sect_mn	Dummy activités spécialisées
sect_oq	Dummy l'administration publique, enseignement, santé humaine et actions sociales
distcaptl_km	Distance à la capitale régionale
d_km2_nbredetabs	Nombre d'établissements par km ²
h_index	Indice de Herfindahl
et_siege	Établissement siège
size_95	Taille

1.2.3 Analyse descriptive

Dans cette section, nous menons une analyse descriptive de nos données. Notre échantillon est composé de 1 651 798 établissements observés entre 1995 et 2005. Au cours de cette période, 648 636 soit 39,27% des établissements ont fermé.

1.2.3.1 Statistiques descriptives

TABLE 1.3 – Statistiques descriptives : Échantillon

	Moyenne	Écart-type	Min	Max
Dépendance économique	88.87	11.73	70.51	130.18
Chômage	8.41	1.85	4.00	16.20
Densité de population	4311.97	7792.41	0.44	40 168.58
Équipements en actions sociales	0.01	0.02	0.00	3.51
Équipements dans le commerce	0.14	0.50	0.00	7.91
Enseignement au 1er degré	0.01	0.01	0.00	1.79
Enseignement au 2nd degré	0.00	0.01	0.00	0.30
Enseignement supérieur et autres	0.00	0.01	0.00	0.31
Services médicaux et paramédicaux	0.09	0.19	0.00	10.71
Services de santé	0.01	0.02	0.00	2.34
Services aux particuliers	0.37	1.08	0.00	251.57
Sport loisirs et culture	0.02	0.10	0.00	62.89
Transport et tourisme	0.05	0.16	0.00	34.88
Distance à la capitale régionale	54.48	53.98	0.00	254.41
Nombre d'établissements par km ²	334.42	977.07	0.01	7189.90
Indice de Herfindahl	0.22	0.02	0.18	0.32
Age	8.30	7.42	1.00	102.16
Taille	6.81	55.71	0.00	16342.00
Nombre d'observations : 1 651 798				

Le tableau 1.3 présente les statistiques descriptives des variables quantitatives pour l'intégralité de l'échantillon. On note en moyenne un faible taux de dépendance économique (88.87%) dans les zones d'emploi où sont implantés les établissements de l'échantillon. La proportion de la population en âge de travailler est donc en moyenne supérieure celle des jeunes (moins de 20 ans) et des seniors (plus de 60 ans).

Le taux de chômage dans les zones d'emploi vaut en moyenne 8.41% ce qui est relativement élevé. La moyenne de l'indice de Herfindahl dans les zones d'emploi est de 0,22. Sachant qu'il oscille entre 0,1 (répartition équitable des secteurs) et 1 (parfaite spécialisation), on peut dire que les différents secteurs sont répartis dans les zones d'emploi de manière relativement équitable.

Au niveau communal, on note une forte hétérogénéité marquée par une moyenne très élevée de la densité de population (4 312) et un écart-type également très élevé (7 792). Cette hétérogénéité est également constatée au niveau du nombre d'établissements par

km² dont la moyenne est de 334 et l'écart-type 977.

En ce qui concerne les équipements, nous en avons en moyenne moins d'un pour 100 personnes par kilomètre carré. Cette valeur reste difficilement interprétable d'autant plus qu'on a pas d'informations sur la capacité d'accueil des équipements. Au début de l'étude (1995), les établissements sont en moyenne âgés de 8 ans et ont en moyenne 7 employés. Ils sont en moyenne situés à 54 km de la capitale de la région où ils sont implantés. Afin de détailler davantage les résultats du tableau 1.3, nous reprenons dans le tableau 1.4 les mêmes statistiques en considérant séparément les établissements ayant fermé et ceux n'ayant pas fermé.

Les résultats du tableau 1.4 nous montrent que les établissements ayant survécu sur la période de l'étude sont localisés dans des zones d'emploi avec un taux de dépendance moyen de 89,24% supérieur à la moyenne de l'échantillon tandis que les établissements ayant fermé se trouvent dans des zones d'emploi avec un taux de dépendance moyen de 88,28% inférieur à la moyenne de l'échantillon. La même tendance est observée pour les variables comme la distance à la capitale régionale, l'indice de Herfindah, l'âge et la taille de l'établissement. Quant au chômage et à la densité de la population, ils sont en moyenne légèrement plus élevés dans les zones d'emploi où sont implantés les établissements ayant fermé. Concernant les équipements, on en compte en moyenne moins d'un pour 100 personnes par km² que ce soit dans les communes des établissements ayant fermé ou celles des établissements ayant survécu. Toutefois, on note que les communes des établissements ayant survécu sont en moyenne moins dotées que les autres communes, sauf pour les équipements de sport, loisir et culture.

TABLE 1.4 – Statistiques descriptives : Établissements fermés vs actifs

	Non Fermés				Fermés			
	Moyenne	Écart-type	Min	Max	Moyenne	Écart-type	Min	Max
Dépendance économique	89.25	11.71	70.51	130.18	88.28	11.74	70.51	130.18
Chômage	8.38	1.86	4.00	16.20	8.46	1.84	4.00	16.20
Densité de population	4111.41	7671.91	0.44	40168.58	4622.15	7965.27	0.44	40168.58
Équipements en actions sociales	0.01	0.02	0.00	3.51	0.01	0.02	0.00	3.51
Équipements dans le commerce	0.14	0.50	0.00	7.91	0.14	0.51	0.00	7.91
Enseignement au 1er degré	0.01	0.01	0.00	1.79	0.01	0.01	0.00	1.79
Enseignement au 2nd degré	0.00	0.01	0.00	0.30	0.00	0.01	0.00	0.30
Enseignement supérieur et autres	0.00	0.01	0.00	0.31	0.00	0.01	0.00	0.14
Services médicaux et paramédicaux	0.09	0.19	0.00	10.71	0.09	0.19	0.00	10.71
Services de santé	0.01	0.02	0.00	2.34	0.01	0.02	0.00	2.34
Services aux particuliers	0.36	1.04	0.00	251.57	0.38	1.14	0.00	251.57
Sport loisirs et culture	0.02	0.08	0.00	62.89	0.02	0.12	0.00	62.89
Transport et tourisme	0.04	0.17	0.00	34.88	0.05	0.15	0.00	34.88
Distance à la capitale régionale	55.98	54.07	0.00	254.41	52.15	53.75	0.00	250.71
Nombre d'établissements par km ²	318.99	954.10	0.01	7189.90	358.29	1011.10	0.01	7189.90
Indice de Herfindahl	0.22	0.02	0.18	0.32	0.22	0.02	0.18	0.32
Age	8.42	7.28	1.00	102.16	8.11	7.64	1.00	95.86
Taille	8.08	63.16	0.00	16342.00	4.85	41.55	0.00	9145.00
Nombre d'observations : 1003162 censurées et 648636 fermées								

Les tableaux 1.5 et 1.6 décrivent les secteurs d'activités et présentent leur table de fréquences. Le tableau 1.7 montre les fréquences pour les établissements ayant fermé et ceux ayant survécu.

Les secteurs de l'industrie (BE), de la construction (FZ), du commerce transport hébergement restauration (GI), des activités spécialisées (MN), et d'administration publique (OQ), sont les plus représentés au sein de notre échantillon. Ces secteurs comptent respectivement 11,05%, 11,82%, 39,45%, 10,20% et 11,41% des établissements étudiés. Des tendances similaires sont constatées lorsqu'on considère séparément les établissements ayant fermé et ceux ayant survécu.

TABLE 1.5 – Secteurs d'activités

Code	Description
AZ	Agriculture, sylviculture et pêche
BE	Industrie manufacturière, industries extractives et autres
FZ	Construction
GI	Commerce de gros et de détail, transports, hébergement et restauration
JZ	Information et communication
KZ	Activités financières et d'assurance
LZ	Activités immobilières
MN	Activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien
OQ	Administration publique, enseignement, santé humaine et actions sociales
RU	Autres activités de services

TABLE 1.6 – Fréquences des secteurs : Échantillon

	Proportions	
	Effectifs	Pourcentages
AZ	21587	1.31
BE	182517	11.05
FZ	195255	11.82
GI	651622	39.45
JZ	32463	1.97
KZ	32883	1.99
LZ	83944	5.08
MN	168518	10.20
OQ	188484	11.41
RU	94525	5.72
Total	1651798	100.00
<i>N</i>	1651798	

TABLE 1.7 – Fréquences des secteurs : Établissements fermés vs actifs

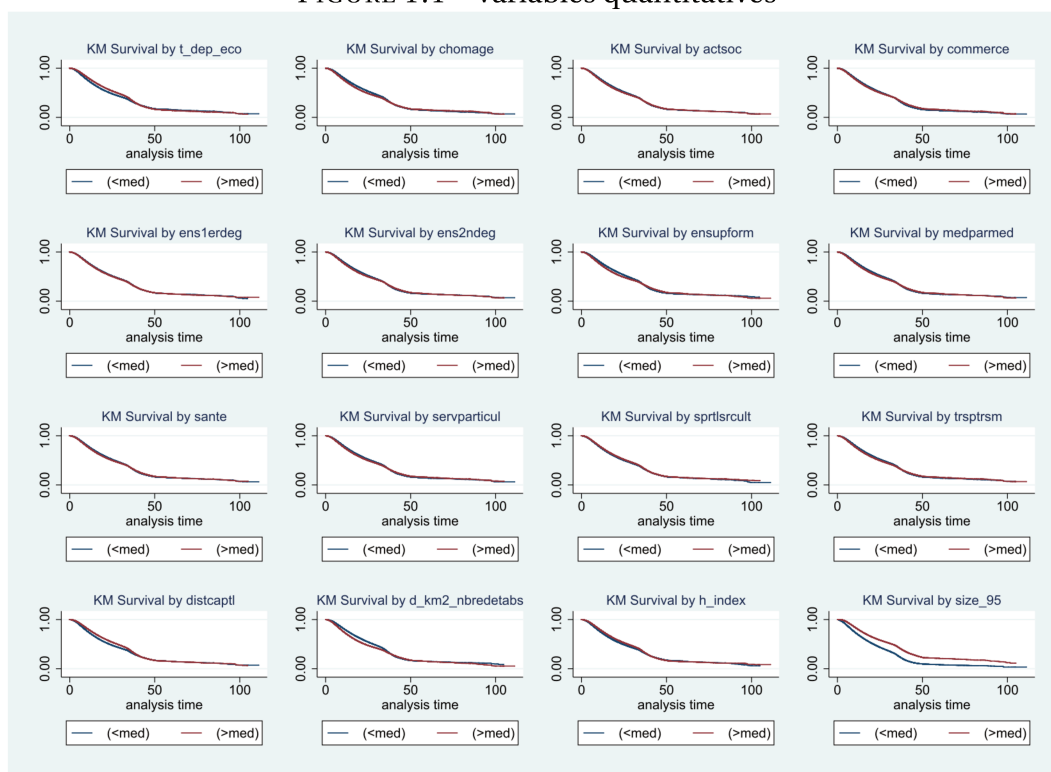
	Non-Fermés		Fermés	
	Effectifs	Pourcentages	Effectifs	Pourcentages
AZ	9371	0.93	12216	1.88
BE	109170	10.88	73347	11.31
FZ	115833	11.55	79422	12.24
GI	377388	37.62	274234	42.28
JZ	19645	1.96	12818	1.98
KZ	20217	2.02	12666	1.95
LZ	65481	6.53	18463	2.85
MN	105921	10.56	62597	9.65
OQ	120948	12.06	67536	10.41
RU	59188	5.90	35337	5.45
Total	1003162	100.00	648636	100.00
<i>N</i>	1003162		648636	

1.2.3.2 Analyse non paramétrique de survie : Estimateur de Kaplan-Meier

Pour aller plus loin dans la description de notre échantillon, nous avons procédé à une analyse non paramétrique de la survie des établissements. Nous avons utilisé la méthode de Kaplan-Meier pour calculer les fonctions de survie représentées sur les figures 1.2 et 1.1. Au niveau des variables quantitatives continues, les individus sont séparés en deux groupes : ceux à droite de la médiane et ceux à gauche de la médiane. Les fonctions de survie de

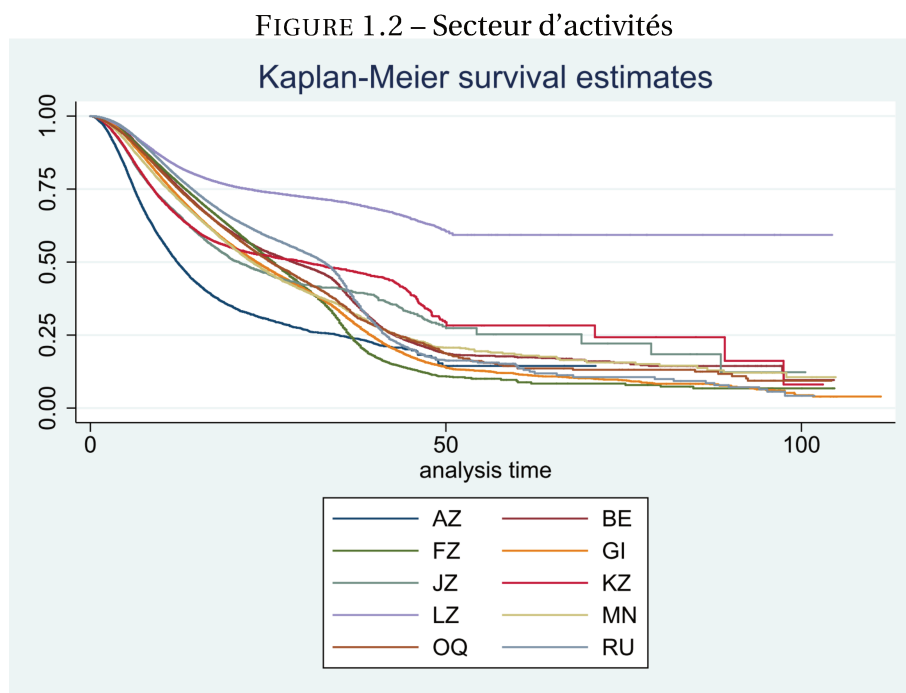
ces deux catégories sont représentées sur un même graphique afin de voir s'il existe une différence. La significativité de la différence constatée est testée par le test log-rank (voir sous-section 1.2.1.1) dont les résultats peuvent être retrouvés dans l'annexe 1.B.

FIGURE 1.1 – Variables quantitatives



Les tests Log-rank indiquent que les différences de survie observées sont toutes significatives. Toutefois, les graphiques de la figure 1.2 ne montrent pas dans la plupart des cas une différence véritablement perceptible au niveau des fonctions de survie. Ce constat peut être expliqué par le fait que plusieurs établissements partagent les mêmes valeurs au niveau des variables calculées à l'échelle communale.

Les quelques différences observées sont au niveau de la distance à la capitale (médiane = 38,91km), le nombre d'établissements par km² (médiane = 30,25 établissements par km²) et la taille de l'établissement en 1995 (médiane = 0 employé).



La fonction de survie des établissements situés à plus de 38,91km de la capitale de leur région paraît très légèrement au dessus de celle des établissements situés à moins de 38,91km de la capitale. Cela indique que les établissements plus éloignées de la capitale vivent plus longtemps. Concernant le nombre d'établissements par km², les courbes montrent que les établissements implantés dans les communes où on dénombre moins de 30 établissements par km² vivent plus longtemps que ceux localisés dans les communes ayant plus de 30 établissements par km². Pour ce qui est de la taille, les établissements situés à droite de la médiane de la distribution du nombre d'employés, ont une fonction de survie très distinctement au dessus de celle des autres établissements. Au niveau des secteurs d'activités, les établissements opérant dans le secteur de l'immobilier ont la durée de vie la plus élevée.

Les différents résultats obtenus à l'issue de cette analyse non-paramétrique permettent certes d'avoir une idée de départ mais restent cependant insuffisants pour tirer de véritables conclusions. Les courbes de survie de Kaplan-Meier, lorsqu'elles sont représentées à partir d'une variable, ne contrôlent pas d'autres variables. Les différences de survies observées pour chaque variable considérée sont donc potentiellement imputable à d'autres facteurs.

Dans la section qui suit, nous menons une analyse semi-paramétrique de la survie qui consiste en l'estimation de plusieurs modèles à hasards proportionnels de Cox.

1.2.4 Résultats d'estimations

Cette section comprend deux parties. Dans un premier temps nous estimons 4 modèles à hasards proportionnels de Cox dont les résultats sont présentés dans le tableau 1.8. Dans un second temps nous estimons un modèle de Cox avec stratification afin de prendre en compte les hétérogénéités intrinsèquement liées au territoire d'implantation des établissements. Les résultats figurent dans le tableau 1.10. Les ratios de risques de tous les modèles sont disponibles dans l'annexe 1.C.

1.2.4.1 Modèles à hasards proportionnels de Cox

Le tableau 1.8 ci-dessous contient les résultats d'estimation de quatre modèles à hasards proportionnels de Cox. Dans le premier modèle, nous incluons exclusivement les variables liées à la zone d'emploi et à la commune de l'établissement. Dans le second modèle nous utilisons uniquement les variables liées à l'établissement. Dans le troisième modèle nous mettons ensemble les variables des deux premiers modèles et dans le quatrième modèle nous rajoutons les variables indicatrices des secteurs d'activités aux variables du modèle 3. Dans les modèles 3 et 4 nous avons effectué les tests de rapport de vraisemblance en prenant le modèle 2 comme modèle contraint. L'idée, c'est d'évaluer la pertinence d'une prise en considération des caractéristiques de l'environnement spatial et des secteurs d'activité. Dans chacun des cas, la statistique du test et la probabilité critique (P-value) montrent que ces variables ont un impact significatif sur la survie des établissements.

Le signe du coefficient estimé indique le sens de l'impact de la variable sur le risque instantané de mortalité des établissements. Ainsi, un coefficient positif indique un risque accru de mortalité et un coefficient négatif indique un effet positif sur la survie. On peut également interpréter les ratios de risques en pourcentage en faisant $100 \times (\text{Ratio de risque} - 1)$. Le résultat obtenu représente la variation en pourcentage du risque de mortalité, associée à la variation de l'explicative d'une unité.

Dans le premier modèle, à l'exception du nombre d'équipements dans les services médicaux et paramédicaux pour 100 personnes par km^2 , toutes les variables ont un effet significatif. Le taux de dépendance économique; les équipements dans le commerce, dans l'enseignement au premier degré, dans l'enseignement supérieur la formation et les services d'éducation, dans les services de santé, dans le sport les loisirs et la culture, dans le transport et le tourisme; le nombre d'établissements par km^2 ; et l'indice de Herfindahl

impactent négativement le risque de mortalité des établissements donc augmentent leur chance de survie. Le taux de chômage; la densité de population; les équipements dans le domaines actions sociales, dans l'enseignement au second degré, dans les services aux particuliers; réduisent les chances de survie des établissements.

Les trois variables prises en compte dans le second modèle sont tous significatifs à 1‰. Les résultats montrent que les établissements sièges ont des risques de mortalité réduits comparativement aux autres établissements. La taille et la distance à la capitale favorisent la survie des établissements.

Dans les modèles 2 et 3, on constate que l'ajout de variables de contrôle supplémentaires n'a pas modifié les résultats des modèles 1 et 2 d'un point de vue qualitatif. On note des variations mineures au niveau des valeurs des coefficients sans aucun changement de signe. Toutefois, les nombres d'équipements dans l'enseignement au premier et au second degré, ont perdu leur significativité.

Dans notre discussion, nous nous baserons sur les résultats du modèle 4 car il est le plus complet. L'objectif principal de cette étude étant d'analyser la survie des établissements en fonction de leur environnement spatial, nous nous focaliserons sur les variables décrivant la zone géographique où l'établissement s'est implanté.

TABLE 1.8 – Résultats des modèles de Cox

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Dépendance économique	-0.00690*** (0.000140)		-0.00498*** (0.000171)	-0.00495*** (0.000171)
Chômage	0.0335*** (0.000711)		0.0308*** (0.000711)	0.0313*** (0.000711)
Densité de population	0.00000289*** (0.000000279)		0.00000287*** (0.000000280)	0.00000274*** (0.000000280)
Équipements en actions sociales	0.366*** (0.0817)		0.268** (0.0827)	0.270** (0.0833)
Équipements dans le commerce	-0.160*** (0.00883)		-0.164*** (0.00880)	-0.171*** (0.00881)
Enseignement au 1er degré	-0.430** (0.135)		0.130 (0.124)	-0.00399 (0.127)
Enseignement au 2nd degré	0.681* (0.287)		0.275 (0.286)	0.420 (0.287)
Enseignement supérieur et autres	-2.097*** (0.178)		-1.968*** (0.178)	-2.062*** (0.178)
Services médicaux et paramédicaux	0.0251 (0.0199)		-0.00328 (0.0197)	0.0236 (0.0199)
Services de santé	-0.453*** (0.129)		-0.418** (0.128)	-0.471*** (0.129)
Services aux particuliers	0.145*** (0.00669)		0.147*** (0.00670)	0.147*** (0.00671)
Sport loisirs et culture	-0.582*** (0.0292)		-0.575*** (0.0278)	-0.578*** (0.0280)
Transport et tourisme	-0.199*** (0.0126)		-0.199*** (0.0127)	-0.202*** (0.0127)
Nombre d'établissements par km ²	-0.0000396*** (0.00000375)		-0.0000412*** (0.00000376)	-0.0000397*** (0.00000377)
Indice de Herfindahl	-2.396*** (0.0790)		-2.141*** (0.0789)	-2.315*** (0.0790)
Établissement siège		-0.531*** (0.00255)	-0.527*** (0.00255)	-0.530*** (0.00256)
Taille		-0.00584*** (0.0000853)	-0.00578*** (0.0000850)	-0.00593*** (0.0000869)
Distance à la capitale régionale		-0.00151*** (0.0000239)	-0.000517*** (0.0000319)	-0.000522*** (0.0000319)
Dummy Industrie				0.129*** (0.00499)
Dummy Construction				0.189*** (0.00489)
Dummy Commerce , transport, hébergement et restauration				0.249*** (0.00383)
Dummy activités spécialisées				0.193*** (0.00523)
Dummy l'administration publique, enseignement, santé humaine et actions sociales				0.0727*** (0.00512)
N	1651788	1651788	1651788	1651788
lrtest_df			15	20
lrtest_chi2			5197.8	10500.6
lrtest_p			0	0

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1.2.4.1.1 Environnement spatial de l'entreprise

1.2.4.1.1.1 Effets positifs sur la survie

Les variables suivantes ont une influence positive sur la survie des établissements :

- Taux de dépendance économique de la zone d'emploi
- Nombre d'équipements dans le commerce pour 100 personnes par km² dans la commune
- Nombre d'équipements dans l'enseignement supérieur la formation et les services d'éducation pour 100 personnes par km² dans la commune
- Nombre d'équipements dans les services de santé pour 100 personnes par km² dans la commune
- Nombre d'équipements dans le sport les loisirs et la culture pour 100 personnes par km² dans la commune
- Nombre d'équipements dans le transport et le tourisme pour 100 personnes par km² dans la commune
- Nombre d'établissements par km² dans la commune
- L'indice de Herfindahl

L'augmentation du taux de dépendance économique d'un point en pourcentage réduit le risque de mortalité des établissements de 0,5% toutes choses égales par ailleurs. Autrement dit, les établissements ont plus de chances de survivre dans les zones d'emploi majoritairement peuplées de jeunes et de seniors. D'après l'atlas des zones d'emploi 2010, les taux élevés de dépendance économique sont généralement constatés dans des zones de taille moyenne ou petite et dans des zones rurales. Les établissements implantés dans ces zones feront face à très peu de concurrence, ce qui pourrait favoriser leur survie. Ce résultat peut-être également vu sous le même angle que celui de [Huiban \(2011\)](#) qui trouve que les entreprises ont une plus grande probabilité de survie dans les zones rurales.

Un équipement additionnel dans le commerce pour 100 personnes par km² dans une commune entraîne une baisse du risque instantané de fermeture d'établissements de 15,7%. La baisse de risque est respectivement de 87,3%; 37,5%; 43,9%; 18,3% pour les équipements dans l'enseignement supérieur la formation et les services d'éducation; dans les services de santé; dans le sport les loisirs et la culture; dans le transport et le tourisme. Ce résultat est assez intuitif dans la mesure où ces équipements représentent des infrastructures vitales pour l'économie de la région. Il est également intéressant de constater que l'impact le plus fort est associé aux équipements dans l'enseignement supérieur, formation et services d'éducation. Il s'agit à ce niveau de centres de production de connaissances et

pourvoyeurs de main d'œuvre qualifiée. Ces centres peuvent donc être de grandes sources d'externalités positives pour les entreprises.

Le nombre d'établissements par km² dans la commune, et l'indice de concentration sectorielle de Herfindahl dans la zone d'emploi améliorent également les chances de survie des établissements. L'effet du nombre d'établissements par km² est certes significatif à 1‰ mais reste extrêmement faible. Le coefficient estimé vaut -0,0000397 ce qui conduit à un ratio de risque approximativement égal à l'unité. Quant à l'indice de Herfindahl, il peut être considéré comme un proxy quantifiant les économies de localisation (Marshall, 1890; Arrow, 1962; Romer, 1986). Le coefficient négatif associé à cette variable signifie que la concentration d'un secteur dans une zone d'emploi augmente la survie des établissements. Cet impact favorable vient du fait qu'une agglomération de cette nature engendre la réduction des coûts de transport, la diffusion des connaissances et le partage des ressources. Cainelli et al. (2014) aboutissent à un résultat similaire en utilisant comme mesure le nombre d'entreprises du même secteur par km².

1.2.4.1.1.2 Effets négatifs sur la survie

Les variables ci-dessous ont un effet négatif sur la survie des établissements :

- Le taux de chômage dans la zone d'emploi
- La densité de population de la commune
- Le nombre d'équipements en actions sociales pour 100 personnes par km² dans la commune
- Le nombre d'équipements dans les services aux particuliers pour 100 personnes par km² dans la commune

Lorsque le taux de chômage de la zone d'emploi augmente d'un point en pourcentage, le risque instantané de mortalité des établissements augmente de 3.2%. Le taux de chômage renseigne indirectement sur le pouvoir d'achat de la population. Un fort taux de chômage implique, toutes choses égales par ailleurs, une faible demande adressée aux entreprises locales et donc un faible chiffre d'affaire. Cette situation peut effectivement dégrader les chances de survie.

La densité de la population, un indicateur d'urbanisation, réduit également les chances de survie des établissements. Le ratio de risque avoisine l'unité en raison d'un coefficient proche de 0 bien que très significatif (1‰). Le niveau élevé d'agglomération caractéristique des zones urbaines peut s'avérer hostile pour beaucoup d'activités économiques. Dans un premier temps, s'implanter dans une grande agglomération peut engendrer d'énormes coûts que tous les établissements ne pourront pas supporter. Ensuite le niveau de concur-

rence dans les zones urbaines peut augmenter les risques de mortalité des établissements. [Brixy & Grotz \(2007\)](#) ont également trouvé que la densité de la population affecte négativement la survie des entreprises.

Un équipement supplémentaire en actions sociales pour 100 personnes par km² augmente de 31% le risque instantané de fermeture des établissements. L'accroissement du risque est de 15,8% lorsqu'il s'agit d'équipements dans les services aux particuliers. La présence d'un grand nombre d'équipements de cette nature dans une région n'est pas forcément le reflet d'une prospérité économique. Cela pourrait expliquer ce résultat.

1.2.4.1.2 Variables de contrôle

Les coefficients associés aux autres explicatives incluses dans le modèle sont tous statistiquement significatifs. Les résultats d'estimations montrent que les établissements sièges ont 41,1% de chance de moins de fermer comparativement aux autres établissements. Comme plusieurs études dans la littérature ([Fotopoulos & Louri, 2000](#); [Staber, 2001](#); [Cefis & Marsili, 2005](#); [Huiban, 2011](#); [F. M. Neffke et al., 2011](#); [Mas-Verdú et al., 2015](#)), nous trouvons que la taille favorise la survie des établissements. Plus précisément, nos résultats d'estimation suggèrent qu'un employé supplémentaire réduit le risque de mortalité instantané de 0,6%. Être éloigné de la capitale de sa région s'avère également bénéfique pour les établissements. Se situer à 1 km plus loin de la capitale régionale baisse le risque de fermeture de 0,1%.

Concernant les variables indicatrices sectorielles, elles affichent toutes des effets négatifs sur la survie des établissements. Les établissements du secteur de l'industrie manufacturière, extractive et autres ont 13,8% de risque de plus que les autres de fermer. Les établissements du secteur de la construction ont 20,8% de risque de plus que les autres de fermer. Les établissements du secteur du commerce de gros et de détail, transport, hébergement et restauration ont 28,3% de risque de plus que les autres de fermer. Les établissements du secteur des activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien ont 21,3% de risque de plus que les autres de fermer. Les établissements du secteur de l'administration publique, enseignement, santé humaine et actions sociales ont 7,5% de risque de plus que les autres de fermer. Ces secteurs pour lesquels nous avons construits des dummies sont les plus représentés dans notre échantillon (voir [tableau 1.6](#)). Le risque accru de mortalité suggéré par nos résultats pourrait donc être expliqué par le fractionnement des parts de marché dû au grand nombre d'établissements qui se concurrencent. Remarquons par exemple que le secteur du commerce de gros et de détail, transport, hébergement et restauration où le risque de fermeture est le plus élevé (28,3%),

a la fréquence relative la plus élevée 39.45% (voir tableau 1.6).

1.2.4.2 Modèles à hasards proportionnels de Cox avec stratification

1.2.4.2.1 Prise en compte des hétérogénéités spatiales

L'une des principales hypothèses des modèles à hasards proportionnels, $h(t | X) = h_0(t)\exp(X\beta)$, c'est que tous les individus ont un même hasard de base h_0 , et donc le différentiel de hasard réside exclusivement dans les variables explicatives incluses dans le modèle. Cette hypothèse est assez forte car il peut y avoir des hétérogénéités non reflétées par les variables explicatives du modèle.

Dans les statistiques descriptives précédemment présentées, nous avons pu constater que les communes et les zones d'emploi considérés sont très hétérogènes. Cette hétérogénéité peut être la conséquence de facteurs spatiaux non observés et peut entraîner des hasards de base h_0 différents pour les entreprises en fonction de leur localisation. La négligence de cet aspect dans la spécification des modèles estimés peut conduire à des conclusions inappropriées.

Dans le but de prendre en compte les éventuelles hétérogénéités intrinsèquement liées au territoire d'implantation de l'établissement, nous avons donc estimé un modèle de Cox avec une stratification spatiale. L'idée de la stratification spatiale consiste à regrouper les entreprises par plusieurs catégories c ($c = 1, 2, \dots, C$) correspondant à des zones géographiques bien définies. Ensuite on supposera dans le modèle de Cox que le hasard de base n'est pas commun à toutes les entreprises, mais qu'il l'est aux entreprises de la même catégorie. On aura donc C hasards de base dans notre modèle (h_{0c} avec $c = 1, 2, \dots, C$).

Pour les entreprises de la classe c on a :

$$h_c(t | X_c) = h_{0c}(t) \exp(X_{ic}\beta)$$

Leur vraisemblance est donnée par :

$$L_c(\beta) = \prod_{i=1}^{d_c} \frac{\exp(X_{ic}\beta)}{\sum_{\mathcal{R}(t_{ic})} \exp X_c\beta}$$

Où d_c représente le nombre d'entreprises de la catégorie c qui ont fermé, et $\mathcal{R}(t_{ic})$ correspond à l'ensemble des entreprises de la catégorie c qui sont toujours actives à l'instant

t_{ic} . La vraisemblance pour tout l'échantillon est donnée par $L(B) = \prod_{c=1}^C L_c(\beta)$.

Le choix des strates a été guidé par une analyse descriptive spatiale à l'échelle des communes et des zones d'emploi. Le but est de détecter d'éventuels schémas spatiaux globaux ou locaux, sources d'hétérogénéités. Pour ce faire, nous avons utilisé les bases Admin Express de l'Institut National de l'information Géographique et Forestière pour construire des cartes choroplèthes et calculer des statistiques d'autocorrélation spatiale. Ces statistiques nous permettent de quantifier pour une variable donnée, la nature de la relation spatiale liant des observations voisines (Dubé & Legros, 2014). L'indice I de Moran (Moran, 1948, 1950) est l'une des mesures les plus courantes de l'autocorrélation spatiale. Le calcul s'effectue à deux niveaux différents à savoir : niveau global et niveau local.

Niveau global

Nous considérons ici le territoire étudié dans sa globalité et nous cherchons à savoir s'il existe un regroupement (ou une dispersion) de valeurs similaires, ou si les valeurs sont aléatoirement distribuées. Le I de Moran se calcule suivant la formule :

$$I = \frac{V^{*'} W^* V^*}{V^{*'} V^*}$$

V^* représente le vecteur de la variable d'intérêt, en l'occurrence le taux de fermeture des établissements par commune et par zone d'emploi. La taille de V^* correspond au nombre de communes et de zones d'emploi et ses éléments sont centrés par la valeur moyenne.

$$V^* = \begin{pmatrix} v_1 - \bar{v} \\ \cdot \\ \cdot \\ \cdot \\ v_N - \bar{v} \end{pmatrix}$$

W^* est une matrice de pondération spatiale standardisée construite par une dichoto-

misation des éléments de distances.

$$W^* = \begin{pmatrix} 0 & w_{12}^* & w_{13}^* & w_{14}^* & \cdot & \cdot & w_{1N}^* \\ w_{21}^* & 0 & w_{23}^* & \cdot & \cdot & \cdot & w_{2N}^* \\ w_{31}^* & w_{32}^* & 0 & \cdot & \cdot & \cdot & w_{3N}^* \\ \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{N1}^* & w_{N2}^* & w_{N3}^* & \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

$$w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^N w_{ij}}$$

$$w_{ij} = \begin{cases} 1 & \text{si } d_{ij} \leq \bar{d} \text{ et } i \neq j \\ 0 & \text{sinon} \end{cases}$$

Les indices i et j représentent l'unité spatiale (communes et zones d'emploi), d_{ij} représente la distance et \bar{d} la distance seuil au delà de laquelle deux unités spatiales ne sont plus considérées comme voisines. La construction s'est effectuée sur le logiciel GeoDa et \bar{d} a été choisi de façon à ce que toutes les unités spatiales aient au moins une voisine. Ce type de matrice a été choisi en raison de sa simplicité.

Niveau local

L'analyse au niveau local ne nous permet pas de saisir les éventuels schémas spatiaux locaux que l'autocorrélation spatiale globale ne permet pas détecter. Il est principalement question à ce niveau d'identifier les regroupements locaux de valeurs faibles ou de valeurs élevées. Pour ce faire, nous avons dans notre travail eu recours aux indicateurs locaux d'autocorrélation spatiale ou LISA (Anselin, 1995) puis construit des cartes de clusters afin de visualiser les schémas spatiaux locaux significatifs. Nous avons donc pu identifier les communes et zones d'emploi ayant des taux de fermeture d'établissements semblables ou non à ceux de leurs voisines.

1.2.4.2.2 Résultats

À l'échelle communale, la carte choroplèthe 1.3 et le graphique 1.5 mettent en évidence l'existence d'une faible autocorrélation spatiale globale du taux de fermeture des établissements. Cette analyse est approfondie par la réalisation d'une carte de clusters (figure 1.7) qui révèle l'existence de schémas spatiaux locaux significatifs. On note ainsi une absence de schéma spatial local significatif du taux de fermeture au niveau de 19 471 communes. Concernant les communes restant, nous en avons : 3 704 qui ont des taux de fermeture élevés et qui sont entourées de communes voisines ayant des taux de fermeture élevés; 3 264 ont des taux de fermeture faibles et sont entourées de voisines ayant des taux de fermeture faibles; 2 006 ont des taux de fermeture faibles et sont entourées de voisines ayant des taux de fermeture élevés; et 1 723 ont des taux de fermeture élevés et sont entourées de voisines ayant des taux de fermeture faibles.

La même analyse est menée au niveau des zones d'emploi (confère cartes 1.4 , 1.8 et figure 1.6). Nous nous baserons sur les résultat obtenus à ce niveau pour déterminer nos strates car les zones d'emploi constituent un découpage beaucoup plus pertinents que les communes.

FIGURE 1.3 – Choroplèthe Communes

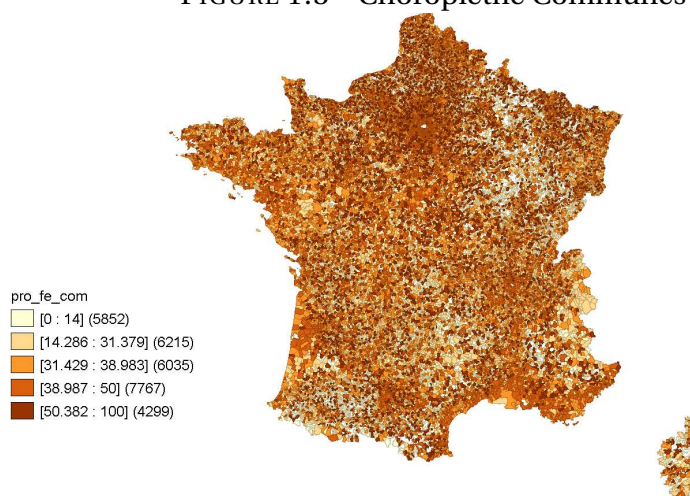


FIGURE 1.4 – Choroplèthe Zones d'emploi

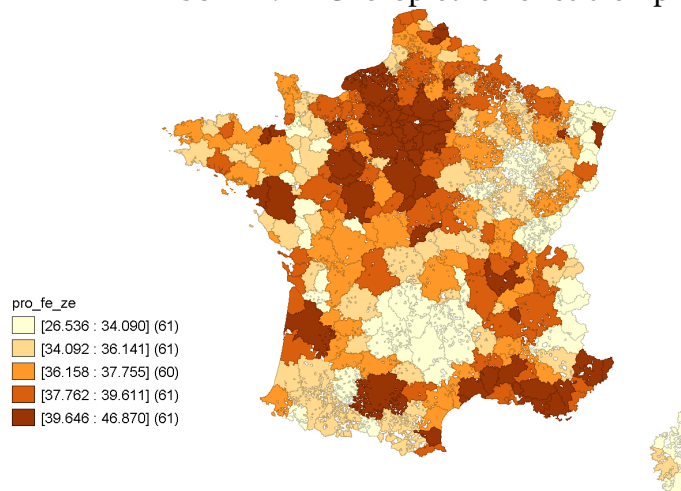


FIGURE 1.5 – I de Moran Communes

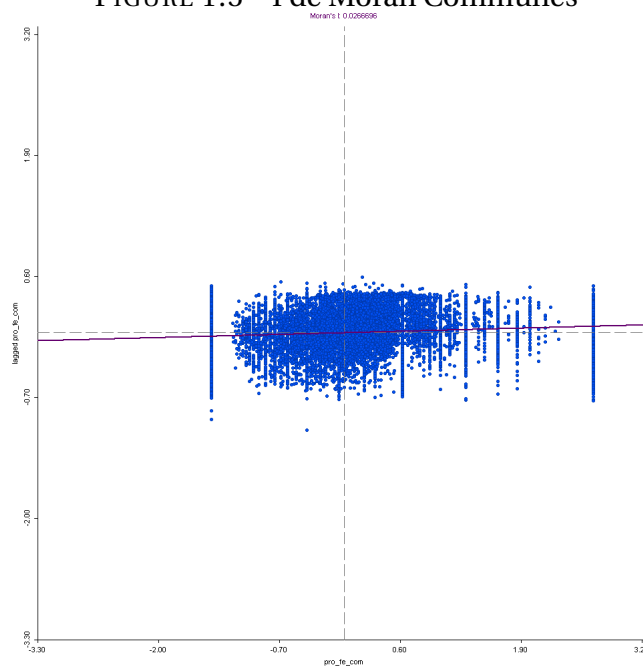


FIGURE 1.6 – I de Moran Zones d'emploi

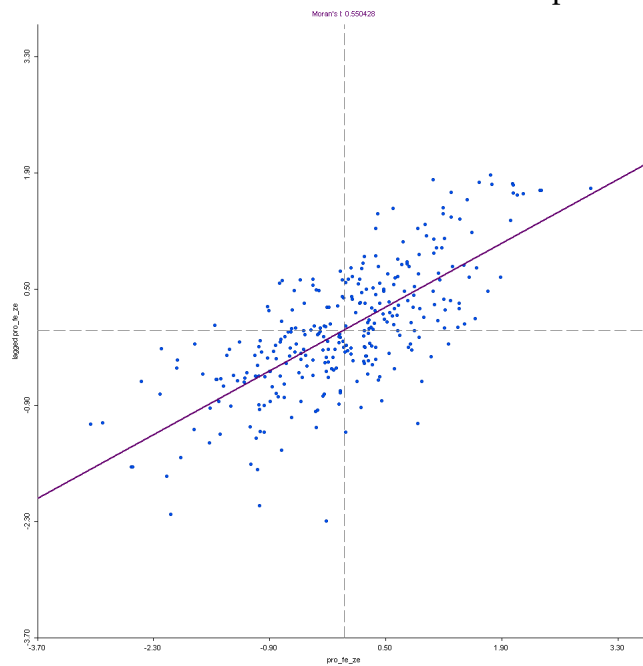


FIGURE 1.7 – Clusters Communes

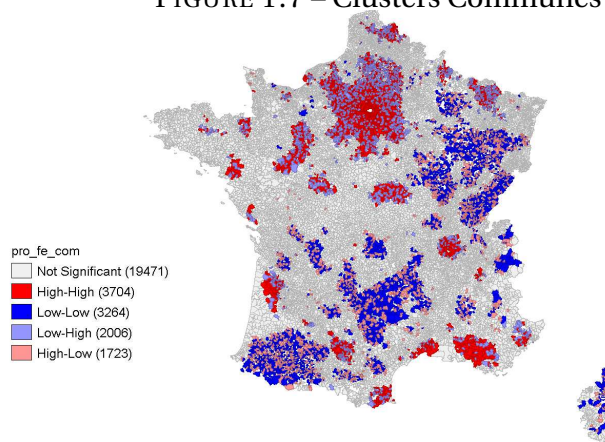


FIGURE 1.8 – Clusters Zones d’emploi

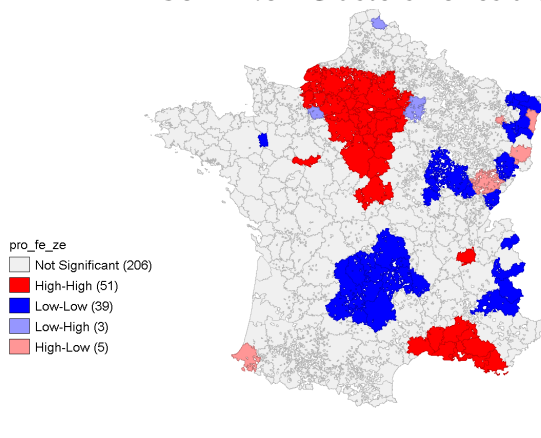


TABLE 1.9 – Analyse descriptive des strates

	Échantillons		Non-fermés		Fermés	
	effectifs	pourcentages	effectifs	pourcentages	effectifs	pourcentages
HH	603 615	36.54291	348506	34.74075	255 109	39.33007
HL	41 194	2.49389	25 341	2.52611	15 853	2.44405
LH	5 972	0.36155	3 834	0.38219	2 138	0.32961
LL	99 579	6.02852	65 740	6.55328	33 839	5.21695
NS	90 1438	54.57314	559 741	55.79767	341 697	52.67931
Total	1651798	100.00000	1003162	100.00000	648 636	100.00000
N	1 651 798		1 003 162		648 636	

Le tableau 1.9 présente une analyse descriptive des 5 niveaux de strates que nous distinguons. La première strate *HH* regroupe les établissements implantés dans les zones d’emploi ayant un fort taux de fermeture et dont les voisins ont également un fort taux de fermeture d’établissements. La seconde strate *HL* regroupe les établissements implantés dans les zones d’emploi ayant un fort taux de fermeture et dont les voisins ont un faible de fermeture. La troisième strate *LH* regroupe les établissements implantés dans les zones d’emploi ayant un faible taux de fermeture et dont les voisins ont un fort de fermeture. La quatrième strate *LL* regroupe les établissements implantés dans les zones d’emploi ayant un faible taux de fermeture d’établissements au même titre que leurs voisins. La dernière strates *NS* regroupe les établissements implantés dans les zones d’emploi où le taux de fermeture d’établissements est aléatoirement distribué. Les strates une fois définies, nous avons procédé à l’estimation.

Les résultats du modèle de Cox avec stratification sont présentés dans le tableau 1.10 qui contient également les résultats du modèle 4 (tableau 1.8) précédemment interprété. Les coefficients estimés sont très proches, ce qui fait que les ratios de risques sont pour la

plus part identiques à quelques décimales près. Les conclusions précédentes basées sur le tableau 1.8 s'avèrent donc robustes à la stratification du modèle.

TABLE 1.10 – Résultats d'estimation

	Modèle sans strates	Modèle avec strates
Dépendance économique	-0.00495*** (0.000171)	-0.00342*** (0.000176)
Chômage	0.0313*** (0.000711)	0.0228*** (0.000742)
Densité de population	0.00000274*** (0.000000280)	0.000000971*** (0.000000283)
Équipements en actions sociales	0.270** (0.0833)	0.304*** (0.0839)
Équipements dans le commerce	-0.171*** (0.00881)	-0.167*** (0.00887)
Enseignement au 1er degré	-0.00399 (0.127)	-0.166 (0.129)
Enseignement au 2nd degré	0.420 (0.287)	-0.140 (0.295)
Enseignement supérieur et autres	-2.062*** (0.178)	-1.774*** (0.178)
Services médicaux et paramédicaux	0.0236 (0.0199)	0.0370 (0.0200)
Services de santé	-0.471*** (0.129)	-0.457*** (0.130)
Services aux particuliers	0.147*** (0.00671)	0.139*** (0.00670)
Sport loisirs et culture	-0.578*** (0.0280)	-0.541*** (0.0278)
Transport et tourisme	-0.202*** (0.0127)	-0.205*** (0.0136)
Nombre d'établissements par km ²	-0.0000397*** (0.00000377)	-0.0000376*** (0.00000376)
Indice de Herfindahl	-2.315*** (0.0790)	-1.708*** (0.0811)
Établissement siège	-0.530*** (0.00256)	-0.530*** (0.00256)
Taille	-0.00593*** (0.0000869)	-0.00593*** (0.0000867)
Distance à la capitale régionale	-0.000522*** (0.0000319)	-0.000347*** (0.0000326)
Dummy Industrie	0.129*** (0.00499)	0.132*** (0.00499)
Dummy Construction	0.189*** (0.00489)	0.190*** (0.00489)
Dummy Commerce , transport, hébergement et restauration	0.249*** (0.00383)	0.254*** (0.00383)
Dummy activités spécialisées	0.193*** (0.00523)	0.190*** (0.00523)
Dummy l'administration publique, enseignement, santé humaine et actions sociales	0.0727*** (0.00512)	0.0805*** (0.00512)
<i>N</i>	1651788	1651788

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dans cette section qui s'achève, nous avons, à travers nos estimations, montré que l'environnement géographique de l'établissement n'est pas neutre dans sa survie. Ces résultats en plus de ceux que l'on trouve dans la littérature empirique, permettent d'un côté de minimiser les doutes quant à l'existence d'un lien entre la survie d'une entreprise et son environnement spatial, et de l'autre côté ces travaux contribuent à l'établissement de faits stylisés sur la question.

En considérant ces résultats empiriques comme des faits stylisés, nous proposons dans la section suivante, de repenser la question sous un angle théorique car les travaux théoriques existants omettent souvent la dimension spatiale.

1.3 Modèle théorique

Si les études théoriques sur la survie des entreprises ont pu prendre en considération la spécificité des entreprises et de l'industrie, la plupart d'entre elles omettent le rôle de l'environnement géographique de l'entreprise. En considérant nos résultats précédents et les études empiriques ayant mis en exergue le lien entre la survie d'une entreprise et son environnement spatial, nous proposons d'introduire un facteur environnemental dans une version simplifiée du modèle de [Schröder & Sørensen \(2012\)](#).

Dans ce qui suit, nous supposons que les entreprises opèrent seulement sur le marché domestique et nous rajoutons au profit de l'entreprise une composante qui représente l'impact de l'environnement spatial. Cette composante représente pour l'entreprise l'avantage (ou l'inconvénient) net associé au fait d'être localisée à son lieu d'implantation.

1.3.1 Les ménages

Considérons un ménage représentatif disposant d'un niveau de richesse instantanée W_t ; offrant une quantité exogène L de travail en contrepartie d'un salaire $w \equiv 1$, et recevant une partie aliquote Π_t du profit des entreprises. En supposant un taux d'escompte nul comme [Melitz \(2003\)](#), le ménage représentatif fait face au problème d'optimisation dynamique suivant⁹ :

9. L'annexe [1.D.1](#) contient des détails supplémentaires relatifs aux calculs.

$$\begin{cases} \text{Max } U_t = \int_t^\infty \ln(C_s) ds \\ \text{S.C } \dot{W}_t = r_t W_t + \Pi_t + L - D_t \end{cases}$$

où r_t et $D_t = P_t C_t$ sont respectivement le taux d'intérêt et la dépense du ménage. C_t et P_t sont les indices de consommation et de prix.

$$\begin{cases} \text{Max } U_t = \int_t^\infty \ln(C_s) ds \\ \text{S.C } \dot{W}_t = r_t W_t + \Pi_t + L - D_t \end{cases} \Rightarrow \begin{cases} \text{Max } U_t = \int_t^\infty \ln(D_s) - \ln(P_s) ds \\ \text{S.C } \dot{W}_t = r_t W_t + \Pi_t + L - D_t \end{cases}$$

Nous pouvons dès lors, par une substitution, réduire le problème afin de le résoudre par la méthode du calcul de variations (Chiang, 1992). Il s'agit désormais de maximiser la valeur :

$$U_t = \int_t^\infty \ln(r_t W_s + \Pi_s + L - \dot{W}_s) - \ln(P_s) ds$$

L'équation d'Euler conduit à :

$$\begin{aligned} \Rightarrow & \frac{\partial [\ln(r_t W_t + \Pi_t + L - \dot{W}_t) - \ln(P_t)]}{\partial W_t} - \frac{d \frac{\partial [\ln(r_t W_t + \Pi_t + L - \dot{W}_t) - \ln(P_t)]}{\partial \dot{W}_t}}{dt} = 0 \\ \Rightarrow & r_t = \frac{\dot{D}_t}{D_t} = r \quad \text{qui est constant sur le sentier de croissance équilibrée.} \end{aligned}$$

La richesse W_t du ménage et la part Π_t de profit croissent aussi de façon constante sur le sentier de croissance équilibrée. Si τ est le taux de croissance de W_t , la contrainte budgétaire nous permet d'écrire

$$\begin{aligned} \frac{\dot{W}_t}{W_t} &= r + \frac{\Pi_t + L - D_t}{W_t} = \tau \\ \tau \text{ constant} &\Rightarrow r + \frac{\Pi_t + L - D_t}{W_t} \text{ constant} \end{aligned}$$

Donc D_t est constant dans le temps.

A chaque instant t , le ménage représentatif dépense de façon optimale $D_t = D = P_t C_t$ dans la consommation d'un continuum de variétés $\omega \in \Omega_t$. L'indice de consommation C_t est de fait représenté par un agrégat Dixit-Stiglitz (Dixit & Stiglitz, 1977).

$$C_t = \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}$$

où $c_t(\omega)$ représente la consommation de la variété ω et $\sigma > 1$ représente l'élasticité de substitution entre les variétés. La rationalité du ménage représentatif dans ses choix de consommation peut être formalisée par :

$$\begin{cases} \text{Max}_{c_t(\omega)} C_t = \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}} \\ \text{S.C} \quad D = P_t C_t = \int_{\omega \in \Omega_t} p_t(\omega) c_t(\omega) d\omega \quad \text{où } p_t(\omega) \text{ est le prix de la variété } \omega \end{cases}$$

En utilisant la méthode du lagrangien, on obtient :

$$\begin{cases} \frac{\sigma}{\sigma-1} \frac{\sigma-1}{\sigma} c_t^{-\frac{1}{\sigma}}(\omega) \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{1}{\sigma-1}} - \lambda p_t(\omega) = 0 \quad \forall \omega \in \Omega_t \\ (\lambda \text{ est le multiplicateur de lagrange}) \\ D = \int_{\omega \in \Omega_t} p_t(\omega) c_t(\omega) d\omega \end{cases}$$

Au niveau du ménage, nous avons finalement

$$\begin{cases} c_t(\omega) = \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \\ C_t = \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}} \\ P_t = \left[\int_{\omega \in \Omega_t} [p_t(\omega)]^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \end{cases} \quad (1.a)$$

1.3.2 Les entreprises

Nous simplifions le cadre de [Schröder & Sørensen \(2012\)](#) en supposant que les entreprises n'opèrent que sur le territoire national d'autant que nous excluons de notre analyse empirique les établissements opérant à l'étranger. Nous rajoutons ensuite une composante ϵ au profit de l'entreprise et nous nous intéressons à la possibilité qu'elle soit liée à la survie de l'entreprise.

L'entrée sur le marché requiert un coût f_e , et chaque entreprise offre une variété ω unique dans un contexte de concurrence monopolistique où le travail est le seul facteur de production. Après l'introduction de sa variété, l'entreprise s'installe avec un niveau de productivité $\phi(\omega)$ issue d'une distribution de Pareto.

En se basant sur [Dasgupta & Stiglitz \(1981\)](#), [Schröder & Sørensen \(2012\)](#) proposent une distribution des productivités qui s'améliore continuellement dans le temps telle dans un modèle à génération de capital. Ainsi, lorsqu'une entreprise produisant la variété ω entre sur le marché à l'instant t , sa productivité $\phi_t(\omega)$ dérive de $G_t(\phi_t(\omega)) = 1 - \left(\frac{\phi_t(\omega)}{\bar{\phi}_t}\right)^{-k}$ où $\phi_t(\omega) \geq \bar{\phi}_t$ et $\bar{\phi}_t = e^{\beta t}$. Précisons que $\bar{\phi}_t$ et k sont respectivement les paramètres d'échelle et de forme de la distribution. En posant $\phi_0 = \frac{\phi_t(\omega)}{\bar{\phi}_t}$ on obtient $\phi_t(\omega) = \phi_0 e^{\beta t}$ pour une entreprise entrée sur le marché à l'instant t et $\phi_{t,\theta}(\omega) = \phi_0 e^{\beta(t-\theta)}$ pour une entreprise âgée de θ années à l'instant t . De plus :

$$\begin{aligned} \text{si } \phi_t(\omega) &\sim G(k, \bar{\phi}_t) \\ \text{alors } \phi_0 &\sim G(k, 1) \end{aligned}$$

La quantité de travail nécessaire pour produire $c_t(\omega)$ unités de la variété ω est $l_t(\omega) = \frac{1}{\phi(\omega)} c_t(\omega) + f$ avec f représentant le coût fixe de production.

$$\begin{aligned} \text{Le coût marginal est : } c_m &= \frac{\partial \frac{1}{\phi(\omega)} c_t(\omega) + f}{\partial c_t(\omega)} = \frac{1}{\phi(\omega)} \\ \text{L'indice de Lerner donne : } \frac{p_t(\omega) - c_m}{p_t(\omega)} &= \left| \frac{\partial c_t(\omega)}{\partial p_t(\omega)} \frac{p_t(\omega)}{c_t(\omega)} \right| \\ \Rightarrow \frac{c_m}{p_t(\omega)} &= 1 - \frac{1}{\sigma} = \frac{\sigma - 1}{\sigma} \end{aligned}$$

$$\text{d'où} \quad p_t(\omega) = \frac{\sigma}{\phi(\omega)(\sigma - 1)} \quad (1.1)$$

Toutes choses égales par ailleurs, le profit d'une entreprise produisant la variété ω pourrait s'écrire :

$$\pi_t(\phi(\omega)) = p_t(\omega)c_t(\omega) - \frac{1}{\phi(\omega)}c_t(\omega) - f$$

Toutefois, la particularité de notre approche repose sur l'hypothèse d'existence d'un impact de l'environnement spatial noté ϵ appartenant à $] - f; f[$. On a désormais :

$$\begin{aligned} \pi_t(\phi(\omega)) &= p_t(\omega)c_t(\omega) - \frac{1}{\phi(\omega)}c_t(\omega) - f + \epsilon \\ &= c_t(\omega) \left(\frac{\sigma}{\phi(\omega)(\sigma - 1)} - \frac{1}{\phi(\omega)} \right) - f + \epsilon \\ &= \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \left(\frac{\sigma}{\phi(\omega)(\sigma - 1)} - \frac{1}{\phi(\omega)} \right) - f + \epsilon \\ \pi_t(\phi(\omega)) &= Q_t(\phi(\omega))^{\sigma-1} - f + \epsilon \\ \text{où } Q_t &= \frac{1}{\sigma - 1} \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma} DP_t^{\sigma-1} \end{aligned} \quad (1.2)$$

Si g_Q est le taux de croissance de Q_t , le profit d'une entreprise âgée de θ années à l'instant t s'écrit :

$$\begin{aligned} \pi_{t,\theta}(\phi_{t,\theta}(\omega)) &= Q_0 e^{g_Q t} \phi_{t,\theta}(\omega)^{\sigma-1} - f + \epsilon \\ \pi_{t,\theta}(\phi_0) &= Q_0 e^{g_Q t} \left(\phi_0 e^{\beta(t-\theta)} \right)^{\sigma-1} - f + \epsilon \end{aligned}$$

En supposant que l'entreprise ferme dès que son profit est nul, on peut calculer l'âge θ^s de fermeture en posant $\pi_{t,\theta^s}(\phi_0) = 0$

$$\begin{aligned} \pi_{t,\theta^s}(\phi_0) = 0 &\Rightarrow \pi_{t,\theta^s}(\phi_0) = Q_0 e^{g_Q t} \left(\phi_0 e^{\beta(t-\theta^s)} \right)^{\sigma-1} - f + \epsilon = 0 \\ &\Rightarrow \ln(Q_0 \phi_0^{\sigma-1}) + g_Q t + \beta(t - \theta^s)(\sigma - 1) = \ln(f - \epsilon) \end{aligned}$$

$$\Rightarrow \theta^s = t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right) + \ln \left(\frac{Q_0 \phi_0^{\sigma-1}}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}}$$

Le niveau ϕ_0^s de productivité minimum pour opérer sur le marché est tel que :

$$\begin{aligned} \theta^s(\phi_0^s) &= 0 \\ \Rightarrow t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right) + \ln \left(\frac{Q_0 (\phi_0^s)^{\sigma-1}}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} &= 0 \\ \phi_0^s &= \left(\frac{f-\epsilon}{Q_0 e^{g_Q t} e^{\beta(\sigma-1)t}} \right)^{\frac{1}{\sigma-1}} \end{aligned}$$

Le profit net actualisé d'une entreprise entrée sur le marché à l'instant t se calcule :

$$\begin{aligned} \pi(\phi_0) &= \int_t^{t+\theta^s} \left(Q_0 e^{g_Q \rho} e^{\beta(\sigma-1)t} \phi_0^{\sigma-1} - f + \epsilon \right) d\rho \\ &= \left[\frac{1}{g_Q} Q_0 e^{g_Q \rho} e^{\beta(\sigma-1)t} \phi_0^{\sigma-1} - (f + \epsilon) \rho \right]_t^{t+\theta^s} \\ \text{Sachant que } \phi_0^s &= \left(\frac{f-\epsilon}{Q_0 e^{g_Q t} e^{\beta(\sigma-1)t}} \right)^{\frac{1}{\sigma-1}} \\ \text{On obtient } \pi(\phi_0) &= Q_0 e^{[\beta(\sigma-1)+g_Q]t} \phi_0^{\sigma-1} \frac{1}{g_Q} \left[e^{g_Q t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right)} \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{g_Q}{\beta(\sigma-1)}} \phi_0^{\frac{g_Q}{\beta}} - 1 \right] \\ &\quad - (f-\epsilon) \left[t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right) \right] - (f-\epsilon) \ln \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} - \frac{1}{\beta} \ln(\phi_0) \end{aligned}$$

Le profit espéré $E[\pi(\phi_0)]$ donne ¹⁰ :

$$\begin{aligned} E[\pi(\phi_0)] &= \int_{\phi_0^s}^{\infty} \pi(\phi_0) \frac{dG(\phi_0)}{d\phi_0} d\phi_0 \\ &= \int_{\phi_0^s}^{\infty} \pi(\phi_0) k \phi_0^{-k-1} d\phi_0 \end{aligned}$$

10. détails dans annexe 1.D.2.

$$E[\pi(\phi_0)] = (f - \epsilon)(\phi_0^s)^{-k} \left[\frac{1}{g_Q} \left(\frac{k}{k - (\sigma - 1 + \frac{g_Q}{\beta})} - \frac{k}{k - (\sigma - 1)} \right) - \frac{1}{\beta k} \right]$$

La libre entrée sur le marché est conditionnée par :

$$(f - \epsilon)(\phi_0^s)^{-k} \left[\frac{1}{g_Q} \left(\frac{k}{k - (\sigma - 1 + \frac{g_Q}{\beta})} - \frac{k}{k - (\sigma - 1)} \right) - \frac{1}{\beta k} \right] = f_e$$

$$\phi_0^s = \left\{ \left(\frac{f - \epsilon}{f_e} \right) \left[\frac{1}{g_Q} \left(\frac{k}{k - (\sigma - 1 + \frac{g_Q}{\beta})} - \frac{k}{k - (\sigma - 1)} \right) - \frac{1}{\beta k} \right] \right\}^{\frac{1}{k}} \text{ où } f_e \text{ est le coût d'entrée.}$$

1.3.3 Équilibre

Nous pouvons à présent déterminer certaines grandeurs comme la demande totale de travail instantanée L_t et l'indice de prix P_t . Soit donc N le nombre total d'entreprises actives sur le marché à l'équilibre, $N_{t,\theta}^c$ le nombre d'entreprises âgées de θ ans¹¹ à l'instant t et $L_{t,\theta}$ leur demande de travail. En supposant que $N_{t,\theta}^c$ croît au taux g_{N^c} on a :

$$L_t = N_0^c e^{g_{N^c} t} f_e + \int_0^\infty L_{t,\theta} d\theta$$

$$\begin{aligned} \text{Notons que }^{12} : L_{t,\theta} &= N_{t,\theta}^c \int_{\phi_t^s}^\infty (f + Q_t(\sigma - 1)\phi^{\sigma-1}) k \bar{\phi}_{t,\theta}^k \phi^{-k-1} d\phi \\ &= N_0^c e^{g_{N^c} t} e^{-\theta(\beta k + g_{N^c})} (\phi^s)^{-k} \left[f \left(\frac{k\sigma - (\sigma - 1)}{k - (\sigma - 1)} \right) - \left(\frac{k(\sigma - 1)}{k - (\sigma - 1)} \right) \epsilon \right] \end{aligned}$$

11. Cohorte entrée sur le marché à l'instant $t - \theta$.

$$\begin{aligned}
L_t &= N_0^c e^{g_{N^c} t} f_e + N_0^c e^{g_{N^c} t} (\phi^s)^{-k} \left[f \left(\frac{k\sigma - (\sigma - 1)}{k - (\sigma - 1)} \right) - \left(\frac{k(\sigma - 1)}{k - (\sigma - 1)} \right) \epsilon \right] \int_0^\infty e^{-\theta(\beta k + g_{N^c})} d\theta \\
&= N_0^c e^{g_{N^c} t} \left\{ f_e + (\phi_0^s)^{-k} \left[f \left(\frac{k\sigma - (\sigma - 1)}{k - (\sigma - 1)} \right) - \left(\frac{k(\sigma - 1)}{k - (\sigma - 1)} \right) \epsilon \right] \left(\frac{1}{\beta k + g_{N^c}} \right) \right\}
\end{aligned}$$

Sur le sentier de croissance équilibrée nous avons :

$$\begin{aligned}
L_t &= L \forall t \text{ puis } N_t^c = N^c \forall t \text{ c'est à dire } g_{N^c} = 0 \text{ ainsi :} \\
L &= N^c \left\{ f_e + (\phi_0^s)^{-k} \left[f \left(\frac{k\sigma - (\sigma - 1)}{k - (\sigma - 1)} \right) - \left(\frac{k(\sigma - 1)}{k - (\sigma - 1)} \right) \epsilon \right] \left(\frac{1}{\beta k} \right) \right\} \\
\text{et } N^c &= L \left\{ f_e + (\phi_0^s)^{-k} \left[f \left(\frac{k\sigma - (\sigma - 1)}{k - (\sigma - 1)} \right) - \left(\frac{k(\sigma - 1)}{k - (\sigma - 1)} \right) \epsilon \right] \left(\frac{1}{\beta k} \right) \right\}^{-1}
\end{aligned}$$

En ce qui concerne N et P_t on a :

$$\begin{aligned}
N &= \int_0^\infty \left(\int_{\phi_t^s}^\infty k \bar{\phi}_{t,\theta}^k \phi^{-k-1} d\phi \right) d\theta = N^c \left(\frac{(\phi_0^s)^{-k}}{\beta k} \right) \\
P_t &= \left[N \int_{\phi_t^s}^\infty \left(\frac{\sigma}{\sigma - 1} \frac{1}{\phi} \right)^{1-\sigma} k (\phi_t^s)^k \phi^{-k-1} d\phi \right]^{\frac{1}{1-\sigma}} \\
&= N^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma - 1} \left[\frac{k}{k - (\sigma - 1)} \right]^{\frac{1}{1-\sigma}} (\phi_t^s)^{-1} \\
\text{sachant que } Q_t &= \frac{1}{\sigma - 1} \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma} D P_t^{\sigma-1} \\
\text{On a } Q_t &= \frac{1}{\sigma - 1} \left(\frac{\sigma}{\sigma - 1} \right)^{-\sigma} D \left\{ N^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma - 1} \left[\frac{k}{k - (\sigma - 1)} \right]^{\frac{1}{1-\sigma}} \right\}^{\sigma-1} (\phi_t^s)^{-(\sigma-1)}
\end{aligned}$$

Par conséquent $g_Q = \frac{\dot{Q}_t}{Q_t} = -\beta(\sigma - 1)$, en le remplaçant dans $\pi_{t,\theta}(\phi_0)$ on obtient :

12. En considérant $c_t(\omega) = \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma}$; $p_t(\omega) = \frac{\sigma}{\phi(\omega)(\sigma-1)}$; $Q_t = \frac{1}{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^{-\sigma} D P_t^{\sigma-1}$ on déduit que $l_t(\omega) = \frac{1}{\phi(\omega)} c_t(\omega) + f = f + Q_t(\sigma - 1)\phi^{\sigma-1}$

$$\phi_{t,\theta}(\phi_0) = Q_0 e^{-\beta(\theta-1)\theta} \phi_0^{\sigma-1} - (f - \epsilon)$$

En posant $\phi_{t,\theta^s}(\phi_0) = 0$, on déduit

$$\theta^s = \text{Max} \left\{ 0; \frac{1}{\beta(\sigma-1)} \ln \left(\frac{Q_0 \phi_0^{\sigma-1}}{f - \epsilon} \right) \right\} \quad (1.3)$$

1.3.4 Discussion

Lemme. *L'environnement spatial impacte la durée de vie de l'entreprise et cette dernière n'entre sur le marché qu'à partir d'une certaine valeur de ϵ .*

Démonstration. Le lien entre la durée de vie de l'entreprise et ϵ est inhérent à l'équation 1.3. De plus, l'entreprise n'entre sur le marché que lorsque $\theta^s > 0$, donc $f - \epsilon < Q_0 \phi_0^{\sigma-1}$. Ainsi, il faut que ϵ soit suffisamment élevé. \square

Proposition 1. *Toute choses égales par ailleurs, pour une entreprise qui est entrée sur le marché, sa productivité impacte positivement sa survie.*

Démonstration. En partant de l'équation 1.3, on obtient $\frac{\partial \theta^s}{\partial \phi_0} = \frac{1}{\beta \phi_0} > 0$ \square

Proposition 2. *La taille de l'entreprise (en termes de chiffre d'affaire) impacte positivement sa survie.*

Démonstration. Le chiffre d'affaire de l'entreprise est $p_t(\omega) c_t(\omega)$. D'après les équations 1.a, 1.1 et 1.2

$$\begin{aligned} \text{Recettes} &= \frac{\sigma}{\phi_t(\omega)(\sigma-1)} \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \\ &= \sigma Q_0 \phi_0^{\sigma-1} \end{aligned}$$

La productivité impacte positivement le chiffre d'affaire, or, d'après la **Proposition 1**, la productivité impacte positivement la survie, par conséquent le chiffre d'affaire impacte positivement la survie. \square

Théorème. *Les avantages liés à la localisation impactent positivement la survie de l'entreprise.*

Démonstration. $\frac{\partial \theta^s}{\partial \epsilon} = \frac{1}{\beta(\sigma-1)(f-\epsilon)} > 0$ \square

Ainsi, notre modèle vient davantage étayer le fait que la durée de vie d'une entreprise est liée aux caractéristiques de son lieu d'implantation.

Conclusion

Le secteur privé par le biais des entreprises qui la composent, joue un rôle prépondérant dans l'économie nationale. Les entreprises sont sources d'innovation, d'emploi et de bien d'autres transformations économiques. Néanmoins, elles ne peuvent avoir véritablement de l'impact que lorsqu'elles sont suffisamment prospères, d'où la nécessité d'analyser les déterminants de leur succès. La littérature sur la survie des entreprises distingue généralement trois groupes de facteurs déterminant la survie des entreprises. Il s'agit des caractéristiques propres à l'entreprise comme l'âge et la taille; des caractéristiques propres à l'industrie comme le degré de concurrence et l'échelle minimale d'efficacité; et des caractéristiques propres à l'environnement géographique qui sont les moins étudiées dans la littérature.

Dans ce chapitre, nous nous sommes intéressés à la survie de 1 651 798 établissements français en fonction de leur environnement géographique. Ces établissements sont observés de 1995 à 2005 et 39,26% d'entre eux ont fermé au cours de la période. Dans un premier temps nous avons, grâce à l'estimation de modèles à hasards proportionnels de Cox, montré que les caractéristiques de l'environnement spatial affectent la survie des établissements de manière significative. Dans un second temps, nous avons montré à travers un modèle théorique que l'environnement spatial peut impacter la survie de l'entreprise. L'une des originalités de notre analyse empirique est que nous avons pris en compte les hétérogénéités du territoire d'implantation de l'établissement en estimant un modèle de Cox avec stratification spatiale.

Nos résultats d'estimation suggèrent que le taux de dépendance économique; le nombre d'équipements dans le commerce; dans l'enseignement supérieur la formation et les services d'éducation; dans les services de santé; dans le sport les loisirs et la culture; dans le transport et le tourisme; le nombre d'établissements par km²; et l'indice de concentration sectorielle de Herfindahl, ont un effet positif sur la survie des établissements. En revanche, le taux de chômage; la densité de population; le nombre d'équipements en actions sociales; et le nombre d'équipements dans les services aux particuliers augmentent le risque de mortalité des établissements. Aussi, conformément à la majorité des travaux dans la littérature (Fotopoulos & Louri, 2000; Staber, 2001; Cefis & Marsili, 2005; Huiban,

2011; F. M. Neffke et al., 2011; Mas-Verdú et al., 2015), nous avons trouvé que la taille influence positivement la survie des établissements.

Ces résultats encouragent vivement les politiques de développement régional. L'augmentation des dotations infrastructurelles peut être un véritable coup d'accélérateur à la prospérité régionale en favorisant notamment l'épanouissement des établissements locaux. L'un des impacts les plus saillants est associé aux équipements dans l'enseignement supérieur la formation et les services d'éducation. Les estimations indiquent qu'un équipement additionnel pour 100 personnes par km² dans une commune réduit le risque instantané de fermeture d'établissements de 87,3%. Cela attire davantage l'attention sur l'importance des infrastructures d'éducation dans les localités.

Par ailleurs, quand nous reconsidérons par exemple la carte 1.4, on peut noter que les fermetures d'établissements (et parallèlement la survie), ont lieu de manière hétérogène sur toute l'étendue du territoire national. Certains établissements vont donc mieux réussir dans les grandes agglomérations alors que d'autres s'en sortiront plus dans des zones moins denses. Les mêmes caractéristiques peuvent rendre une région attractive pour une entreprise et moins attractive pour une autre. En supposant que l'entreprise choisisse le lieu d'implantation qui lui convient le mieux, et en prenant en compte ce choix, qu'en serait-il des impacts de l'environnement spatial sur sa survie? Le chapitre suivant propose d'aborder la question en suivant un échantillon de jeunes entreprises les cinq premières années de leur vie.

Annexes

1.A Catégories d'équipements

- Sport, loisirs et culture
 - ▶ Bassins de natation, Boulodromes, Tennis, Équipements de cyclisme, Domaines skiabiles, Centres équestres, Athlétismes, Terrains de golf, Parcours sportifs/santé, Sports de glace, Plateaux et terrains de jeux extérieurs, Salles spécialisés, Terrains de grands jeux, Salles de combat, Salles non spécialisées, Sports nautiques, Bowling, Salles de remise en forme, Salles multisports (gymnase), Baignade aménagée, Ports de plaisance - Mouillage, Boucles de randonnée, Cinémas, Musées, Conservatoire, Théâtres-Arts de la rue-Pôles Cirque
- Commerce
 - ▶ Hypermarchés, Supermarchés, Grandes surfaces de bricolage, Supérettes, Epiceries, Boulangeries, Boucheries-charcuteries, Produits surgelés, Poissonneries, Librairie papeterie journaux, Magasins de vêtements, Magasins d'équipements du foyer, Magasins de chaussures, Magasins d'électroménager et de matériel audio-vidéo, Magasins de meubles, Magasins d'articles de sports et de loisirs, Magasins de revêtements murs et sols, Drogueries-quincailleries-bricolages, Parfumerie-Cosmétique, Horlogerie Bijouterie, Fleuriste-Jardinerie-Animalerie, Magasins d'optique, Magasins de matériel médical et orthopédique, Stations services
- Services aux particuliers
 - ▶ Police, Gendarmerie, Cours d'appel, Tribunaux de grande instance, Tribunaux d'instance, Conseil des prud'hommes, Tribunaux de commerce, Réseaux spécialisés pôle emploi, Direction Générale des Finances Publiques, Direction Régionale des Finances Publiques, Direction Départementale des Finances Publiques, Réseau de proximité pôle emploi, Réseau partenarial pôle emploi, Maisons de justice et du droit, Antenne de justice Conseil départemental d'accès au droit, Banques-caisses d'épargne, Pompes funèbres, Bureau de poste, Relais poste, Agence postale, Réparation auto et de matériel agricole, Contrôle technique auto, Location auto-utilitaires légers, Écoles de conduite, Maçons, Plâtriers-peintres, Menuisiers-charpentiers-serruriers, Plombiers-couvreurs-chauffagistes, Électriciens, Entreprises générales du bâtiment, Coiffure, Vétérinaires, Agences de travail temporaire, Restaurant-Restoration rapide, Agences immobilières, Pressing

– Laverie automatique, Instituts de beauté – Ongleries

- Actions sociales
 - ▶ Hébergements pour personnes âgées-enfants handicapés-adultes handicapés, Soins à domicile pour personnes âgées, Services d'aide pour personnes âgées-adultes handicapés, Foyers restaurants pour personnes, Services de repas à domicile pour personnes âgées, Crèches, Hébergements pour enfants handicapés, Services à domicile ou ambulatoires pour enfants handicapés, Travail protégé, Aides sociales à l'enfance, Centres d'Hébergement et de Réadaptation Sociale (CHRS), Centres provisoires d'hébergement, Centres d'accueil pour demandeurs d'asile, Autres établissements pour adultes et familles en difficulté
- Services de santé
 - ▶ Établissements santé court séjour, Établissements santé moyen séjour, Établissements santé long séjour, Établissements psychiatriques, Centres lutte cancer, Urgences, Maternités, Centres de santé, Structures psychiatriques en ambulatoire, Centres médecine préventive, Dialyses, Hospitalisation à domicile, Maisons de santé pluridisciplinaires, Pharmacies, Laboratoires d'analyses et de biologie médicales, Ambulances, Transfusion sanguine, Établissement thermal
- Fonctions médicales et paramédicales
 - ▶ Médecins généralistes, Spécialistes (cardiologie, dermatologie vénéréologie, gastro-entérologie hépatologie, psychiatrie, ophtalmologie, oto-rhino-laryngologie, pédiatrie, pneumologie, radiodiagnostic et imagerie médicale, stomatologie, gynécologie), Chirurgien dentiste, Sage-femme, Infirmier, Masseur kinésithérapeute, Orthophoniste, Orthoptiste, Pédicure-podologue, Audio prothésiste, Ergothérapeute, Psychomotricien, Diététicien, Psychologue
- Tourisme et transports
 - ▶ Taxis, Voitures de Transport avec Chauffeur, Aéroports, Gares voyageurs d'intérêt national, Gares voyageurs d'intérêt régional, Gares voyageurs d'intérêt local, Agences de voyage, Hôtels, Campings, Informations Touristique
- Enseignement du premier degré
 - ▶ Écoles maternelles et élémentaires (public et privé)
- Enseignement du second degré
 - ▶ Collèges, Lycées d'enseignement général et/ou technologique, Sections d'enseignement général et technologique, Sections d'enseignement professionnel
- Enseignement supérieur, formation et services de l'éducation
 - ▶ Sections Techniciens Supérieurs-Classes Prépa Grandes Ecoles, Formations santé, Formations commerce, Autres formations post bac non universitaires,

Unités de Formation et de Recherche, Instituts universitaires, Écoles d'ingénieurs, Enseignement général supérieur, Écoles d'enseignement supérieur agricole, Autres enseignements supérieurs, Centres de formation d'apprentis hors agriculture, Les groupements d'établissements (Greta), Centres dispensant de la formation continue agricole, Formations aux métiers du sport, Autres formations continues, Résidences universitaires, Restaurants universitaires.

1.B Tests log-rank

TABLE 1.B.1 – Taux de dépendance économique

Catégories		Événements	
Code	Libellés	Observés	Espérés
0	Valeurs inférieures à la médiane	340029	309298.68
1	Valeurs supérieures à la médiane	308597	339327.32
Total		648626	48626.00
Chi2 (1)		5849.35	
Pr>Chi2		0.0000	

Résultat : Chances de survie significativement différentes pour les deux catégories

TABLE 1.B.2 – Taux de chômage

Catégories		Événements	
Code	Libellés	Observés	Espérés
0	Valeurs inférieures à la médiane	303849	327858.56
1	Valeurs supérieures à la médiane	344777	320767.44
Total		648626	648626.00
Chi2 (1)		3560.51	
Pr>Chi2		0.0000	

Résultat : Chances de survie significativement différentes pour les deux catégories

Catégories		Densité de population		Équipements en actions sociales	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	303952	341402,49	317131	328702,40
1	Valeurs supérieures à la médiane	344674	307223,51	331495	319923,60
Total		648626	648626,00	648626	648626,00
Chi2(1)		8698,85		826,49	
Pr>Chi2		0.0000*		0.0000*	

* Chances de survie significativement différentes pour les deux catégories

Catégories		Équipements dans le commerce		Enseignement au 1er degré	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	319406	329600.52	320645	326330.4
1	Valeurs supérieures à la médiane	329220	319025.48	327981	322295.52
Total		648626	648626.00	648626	648626.00
Chi2(1)			641.71		199.43
Pr>Chi2			0.0000*		0.0000*

* Chances de survie significativement différentes pour les deux catégories

Catégories		Enseignement au 2nd degré		Enseignement supérieur et autres	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	315108	328234.90	311406	335784.83
1	Valeurs supérieures à la médiane	333518	320391.10	337220	312841.17
Total		648626	648626.00	648626	648626.00
Chi2(1)			1063.90		3677.69
Pr>Chi2			0.0000*		0.0000*

* Chances de survie significativement différentes pour les deux catégories

Catégories		Équipements médicaux et paramédicaux		Équipements en santé	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	317557	332153,49	316951	327675,05
1	Valeurs supérieures à la médiane	331069	316472,51	331675	320950,95
Total		648626	648626,00	648626	648626,00
Chi2(1)		8698,85		826,49	
Pr>Chi2		0.0000*		0.0000*	

* Chances de survie significativement différentes pour les deux catégories

Catégories		Services aux particuliers		Sport, loisirs et culture	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	318596	332155,92	328674	318794,72
1	Valeurs supérieures à la médiane	330030	316470,08	319952	329831,28
Total		648626	648626,00	648626	648626,00
Chi2(1)			1135,66		602,40
Pr>Chi2			0,0000*		0,0000*

* Chances de survie significativement différentes pour les deux catégories

Catégories		Transport et tourisme		Distance à la capitale	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	318466	333488,48	340334	307706,95
1	Valeurs supérieures à la médiane	330160	315137,52	308292	340919,05
Total		648626	648626,00	648626	648626,00
Chi2(1)		1394.17		6598.21	
Pr>Chi2		0.0000*		0.0000*	

* Chances de survie significativement différentes pour les deux catégories

Catégories		Nombre d'établissements par km ²		Indice de Herfindahl	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	303578	341432.47	332307	310346.78
1	Valeurs supérieures à la médiane	345048	307193.53	316319	338279.22
Total		648626	648626.00	648626	648626.00
Chi2(1)		8888.71		2983.20	
Pr>Chi2		0.0000*		0.0000*	

* Chances de survie significativement différentes pour les deux catégories

TABLE 1.B.3 – Taille de l'établissement

Catégories		Événements	
Code	Libellés	Observés	Espérés
0	Valeurs inférieures à la médiane	380668	305020.59
1	Valeurs supérieures à la médiane	267958	343605.41
Total		648626	648626.00
Chi2 (1)		35799.53	
Pr>Chi2		0.0000	

Résultat : Chances de survie significativement différentes pour les deux catégories

TABLE 1.B.4 – Secteurs d'activités

Secteurs	Événements	
	Observés	Espérés
AZ	12216	5674.69
BE	73346	79210.80
FZ	79421	81932.70
GI	274229	253565.88
JZ	12818	9760.71
KZ	12666	10504.15
LZ	18462	38461.83
MN	62596	56390.89
OQ	67536	69884.88
RU	35336	43239.47
Total	648626	648626.00
Chi2(9)	23936.95	
Pr>Chi2	0.0000	

Résultat : Chances de survie significativement différentes entre les secteur

1.C Ratios de risques

TABLE 1.C.1 – Ratios de risques COX sans stratification

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Dépendance économique	0.993*** (0.000139)		0.995*** (0.000170)	0.995*** (0.000170)
Chômage	1.034*** (0.000735)		1.031*** (0.000733)	1.032*** (0.000734)
Densité de population	1.000*** (0.000000279)		1.000*** (0.000000280)	1.000*** (0.000000280)
Équipements en actions sociales	1.443*** (0.118)		1.307** (0.108)	1.310** (0.109)
Équipements dans le commerce	0.852*** (0.00752)		0.849*** (0.00747)	0.843*** (0.00742)
Enseignement au 1er degré	0.651** (0.0877)		1.139 (0.142)	0.996 (0.126)
Enseignement au 2nd degré	1.976* (0.566)		1.316 (0.377)	1.522 (0.437)
Enseignement supérieur et autres	0.123*** (0.0218)		0.140*** (0.0249)	0.127*** (0.0227)
Services médicaux et paramédicaux	1.025 (0.0204)		0.997 (0.0196)	1.024 (0.0203)
Services de santé	0.635*** (0.0821)		0.658** (0.0842)	0.625*** (0.0808)
Services aux particuliers	1.156*** (0.00774)		1.158*** (0.00776)	1.158*** (0.00777)
Sport loisirs et culture	0.559*** (0.0163)		0.563*** (0.0157)	0.561*** (0.0157)
Transport et tourisme	0.819*** (0.0104)		0.820*** (0.0104)	0.817*** (0.0104)
Nombre d'établissements par km ²	1.000*** (0.00000375)		1.000*** (0.00000376)	1.000*** (0.00000377)
Indice de Herfindahl	0.0911*** (0.00719)		0.117*** (0.00927)	0.0988*** (0.00781)
Établissement siège		0.588*** (0.00150)	0.590*** (0.00151)	0.589*** (0.00151)
Taille		0.994*** (0.0000848)	0.994*** (0.0000845)	0.994*** (0.0000864)
Distance à la capitale régionale		0.998*** (0.0000239)	0.999*** (0.0000319)	0.999*** (0.0000319)
Dummy Industrie				1.138*** (0.00567)
Dummy Construction				1.208*** (0.00590)
Dummy Commerce, transport, hébergement et restauration				1.283*** (0.00491)
Dummy activités spécialisées				1.213*** (0.00634)
Dummy l'administration publique, enseignement, santé humaine et actions sociales				1.075*** (0.00550)
N	1651788	1651788	1651788	1651788

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 1.C.2 – Ratios de risques COX avec stratification

	Common Baseline hazard	stratified Cox Model
Dépendance économique	0.995*** (0.000170)	0.997*** (0.000175)
Chômage	1.032*** (0.000734)	1.023*** (0.000759)
Densité de population	1.000*** (0.000000280)	1.000*** (0.000000283)
Équipements en actions sociales	1.310** (0.109)	1.356*** (0.114)
Équipements dans le commerce	0.843*** (0.00742)	0.846*** (0.00750)
Enseignement au 1er degré	0.996 (0.126)	0.847 (0.109)
Enseignement au 2nd degré	1.522 (0.437)	0.869 (0.256)
Enseignement supérieur et autres	0.127*** (0.0227)	0.170*** (0.0302)
Services médicaux et paramédicaux	1.024 (0.0203)	1.038 (0.0207)
Services de santé	0.625*** (0.0808)	0.633*** (0.0825)
Services aux particuliers	1.158*** (0.00777)	1.149*** (0.00770)
Sport loisirs et culture	0.561*** (0.0157)	0.582*** (0.0162)
Transport et tourisme	0.817*** (0.0104)	0.815*** (0.0111)
Nombre d'établissements par km ²	1.000*** (0.00000377)	1.000*** (0.00000376)
Indice de Herfindahl	0.0988*** (0.00781)	0.181*** (0.0147)
Établissement siège	0.589*** (0.00151)	0.588*** (0.00151)
Taille	0.994*** (0.0000864)	0.994*** (0.0000862)
Distance à la capitale régionale	0.999*** (0.0000319)	1.000*** (0.0000326)
Dummy Industrie	1.138*** (0.00567)	1.142*** (0.00570)
Dummy Construction	1.208*** (0.00590)	1.210*** (0.00591)
Dummy Commerce , transport, hébergement et restauration	1.283*** (0.00491)	1.289*** (0.00494)
Dummy activités spécialisées	1.213*** (0.00634)	1.210*** (0.00632)
Dummy l'administration publique, enseignement, santé humaine et actions sociales	1.075*** (0.00550)	1.084*** (0.00555)
<i>N</i>	1651788	1651788

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1.D Modèle théorique

1.D.1 Ménages

Maximisation de l'utilité du ménage

Maximisation de la valeur :

$$U_t = \int_t^{\infty} \ln(r_t W_s + \Pi_s + L - \dot{W}_s) - \ln(P_s) ds$$

L'équation d'Euler conduit à :

$$\begin{aligned} \Rightarrow & \frac{\partial [\ln(r_t W_t + \Pi_t + L - \dot{W}_t) - \ln(P_t)]}{\partial W_t} - \frac{d \frac{\partial [\ln(r_t W_t + \Pi_t + L - \dot{W}_t) - \ln(P_t)]}{\partial \dot{W}_t}}{dt} = 0 \\ \Rightarrow & \frac{r_t}{D_t} - \frac{d(-\frac{1}{D_t})}{dt} = 0 \\ \Rightarrow & \frac{r_t}{D_t} - \frac{\dot{D}_t}{D_t^2} = 0 \\ \Rightarrow & r_t = \frac{\dot{D}_t}{D_t} = r \quad \text{qui est constant sur le sentier de croissance équilibrée.} \end{aligned}$$

Dépense constante

$$\begin{aligned} \frac{\dot{W}_t}{W_t} &= r + \frac{\Pi_t + L - D_t}{W_t} = \tau \\ \tau \text{ constant} &\Rightarrow r + \frac{\Pi_t + L - D_t}{W_t} \text{ constant} \\ &\Rightarrow \frac{\Pi_t + L - D_t}{W_t} \text{ constant} \\ &\Rightarrow \Pi_t + L - D_t = 0 \\ &\Rightarrow \dot{\Pi}_t + \dot{L} - \dot{D}_t = 0 \\ &\Rightarrow \dot{D}_t = \dot{\Pi}_t = r D_t \end{aligned}$$

$$\begin{aligned}
& \Rightarrow \frac{\dot{\Pi}_t}{\Pi_t} = r \frac{D_t}{\Pi_t} \\
\frac{\dot{\Pi}_t}{\Pi_t} \text{ étant constant, on doit avoir} & \frac{d\frac{\dot{\Pi}_t}{\Pi_t}}{dt} = 0 \\
& \Rightarrow r \frac{d\frac{D_t}{\Pi_t}}{dt} = 0 \\
& \Rightarrow \frac{\dot{D}_t}{D_t} = \frac{\dot{\Pi}_t}{\Pi_t} \\
& \Rightarrow \dot{D}_t = \dot{\Pi}_t = 0 \text{ du moment où } D_t \neq \Pi_t \\
& \Rightarrow D_t \text{ est constant dans le temps.}
\end{aligned}$$

Indices de prix et de consommation

En utilisant la méthode du lagrangien, on obtient :

$$\left\{ \begin{array}{l}
\frac{\sigma}{\sigma-1} \frac{\sigma-1}{\sigma} c_t^{-\frac{1}{\sigma}}(\omega) \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{1}{\sigma-1}} - \lambda p_t(\omega) = 0 \quad \forall \omega \in \Omega_t \\
(\lambda \text{ est le multiplicateur de lagrange}) \\
D = \int_{\omega \in \Omega_t} p_t(\omega) c_t(\omega) d\omega
\end{array} \right.$$

$$\begin{aligned}
\forall \omega \text{ et } \omega' \quad \text{on a} \quad & \frac{\frac{\sigma}{\sigma-1} \frac{\sigma-1}{\sigma} c_t^{-\frac{1}{\sigma}}(\omega) \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{1}{\sigma-1}}}{\frac{\sigma}{\sigma-1} \frac{\sigma-1}{\sigma} c_t^{-\frac{1}{\sigma}}(\omega') \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{1}{\sigma-1}}} = \frac{\lambda p_t(\omega)}{\lambda p_t(\omega')} \\
& \Rightarrow \left(\frac{c_t(\omega)}{c_t(\omega')} \right)^{-\frac{1}{\sigma}} = \frac{p_t(\omega)}{p_t(\omega')} \\
& \Rightarrow \left(\frac{c_t(\omega)}{C_t} \right)^{-\frac{1}{\sigma}} = \frac{p_t(\omega)}{P_t} \\
& \Rightarrow c_t(\omega) = \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \\
& \Rightarrow D = \int_{\omega \in \Omega_t} p_t(\omega) \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} d\omega \\
& \Rightarrow P_t = \left[\int_{\omega \in \Omega_t} [p_t(\omega)]^{1-\sigma} \right]^{\frac{1}{1-\sigma}}
\end{aligned}$$

En résumé, au niveau du ménage, nous avons : $c_t(\omega) = \frac{D}{P_t} \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma}$; $C_t = \left[\int_{\omega \in \Omega_t} [c_t(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}$ et $P_t = \left[\int_{\omega \in \Omega_t} [p_t(\omega)]^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$

1.D.2 Entreprise

Profit espéré

Le profit espéré $E[\pi(\phi_0)]$ donne :

$$\begin{aligned}
E[\pi(\phi_0)] &= \int_{\phi_0^s}^{\infty} \pi(\phi_0) \frac{dG(\phi_0)}{d\phi_0} d\phi_0 \\
&= \int_{\phi_0^s}^{\infty} \pi(\phi_0) k \phi_0^{-k-1} d\phi_0 \\
&= \int_{\phi_0^s}^{\infty} \left\{ Q_0 e^{[\beta(\sigma-1)+g_Q]t} \phi_0^{\sigma-1} \frac{1}{g_Q} \left[e^{g_Q t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right)} \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{g_Q}{\beta(\sigma-1)}} \phi_0^{\frac{g_Q}{\beta}} - 1 \right] \right. \\
&\quad \left. - (f-\epsilon) \left[t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right) \right] - (f-\epsilon) \ln \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} - \frac{1}{\beta} \ln(\phi_0) \right\} d\phi_0 \\
&= \left[\frac{Q_0 e^{[\beta(\sigma-1)+g_Q]t} \phi_0^{\sigma-1-k}}{g_Q} \left(\frac{k \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{g_Q}{\beta(\sigma-1)}} e^{g_Q t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right)} \phi_0^{\frac{g_Q}{\beta}}}{(\sigma-1 + \frac{g_Q}{\beta}) - k} - \frac{k}{(\sigma-1) - k} \right) \right]_{\phi_0^s}^{\infty} \\
&\quad - \left[-\phi_0^{-k} \left((f-\epsilon) t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right) + (f-\epsilon) \ln \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} \right) - (f-\epsilon) \frac{1}{\beta} \phi_0^{-k} \left(\ln \phi_0 + \frac{1}{k} \right) \right]_{\phi_0^s}^{\infty} \\
&= \left[\frac{k}{k - (\sigma-1 + \frac{g_Q}{\beta})} \left(e^{g_Q t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right)} \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{g_Q}{\beta(\sigma-1)}} \right) (\phi_0^s)^{\frac{g_Q}{\beta}} - \frac{k}{k - (\sigma-1)} \right] \frac{Q_0 e^{[\beta(\sigma-1)+g_Q]t} (\phi_0^s)^{\sigma-1-k}}{g_Q} \\
&\quad - (f-\epsilon) (\phi_0^s)^{-k} \left[t \left(1 + \frac{g_Q}{\beta(\sigma-1)}\right) + \ln \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} + \frac{1}{\beta} \ln(\phi_0^s) \right] - (f-\epsilon) (\phi_0^s)^{-k} \frac{1}{\beta k}
\end{aligned}$$

En remarquant que $\left(e^{g_Q t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right)} \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{g_Q}{\beta(\sigma-1)}} \right) = (\phi_0^s)^{-\frac{g_Q}{\beta}}$
que $\frac{Q_0 e^{[\beta(\sigma-1)+g_Q]t} (\phi_0^s)^{\sigma-1-k}}{g_Q} = \frac{f-\epsilon}{g_Q} (\phi_0^s)^{-k}$
et que $t \left(1 + \frac{g_Q}{\beta(\sigma-1)} \right) + \ln \left(\frac{Q_0}{f-\epsilon} \right)^{\frac{1}{\beta(\sigma-1)}} + \frac{1}{\beta} \ln(\phi_0^s) = 0$

On a finalement $E[\pi(\phi_0)] = (f-\epsilon)(\phi_0^s)^{-k} \left[\frac{1}{g_Q} \left(\frac{k}{k - (\sigma-1 + \frac{g_Q}{\beta})} - \frac{k}{k - (\sigma-1)} \right) - \frac{1}{\beta k} \right]$

Chapitre 2

Choix de localisation et survie des jeunes entreprises françaises

Introduction

La force d'une économie réside dans sa capacité à croître de manière soutenue, sa compétitivité, et surtout sa résilience, qui correspond à sa capacité de résorption des chocs exogènes. Les économies les plus résilientes sont particulièrement diversifiées, et ce grâce à l'entrepreneuriat. En effet, les jeunes entreprises, à travers la création d'emplois, l'introduction de nouveaux procédés de production, et de nouvelles variétés de produits, impulsent une dynamique de développement économique durable dans les régions. Cependant, l'entrée sur le marché est une prise de risque importante et les jeunes entreprises se retrouvent assez souvent en faillites seulement après quelques années d'activités. Indépendamment du pays où elles sont créées, moins de 50% des nouvelles entreprises survivent au delà de cinq années (Cefis & Marsili, 2011; Colombelli et al., 2016). En France, d'après les statistiques de l'INSEE¹, les taux de fermeture au bout de cinq ans étaient respectivement de 40 et 39% pour les entreprises créées au premier semestre 2010 et 2014.

Au début de leurs activités de production, la majorité des jeunes entreprises ne disposent pas suffisamment d'armes internes pour braver les aléas inhérents au marché. Elles sont généralement de petites et moyennes tailles, et rentrent souvent dans l'industrie avec un niveau de production insuffisant pour être profitable, ce qui fragilise considérablement leur chance de succès. De plus, il leur faut du temps pour apprivoiser leur nouvel environnement afin de pouvoir être compétitives et cette période d'adaptation les expose à davantage de risques («aléa de la nouveauté», Stinchcombe, 1965). Au-delà de ces aspects spécifiques à l'entreprise elle-même, l'environnement géographique d'accueil, de par les conditions qui le caractérisent, peut faciliter ou non, la réussite de la jeune entreprise. La structure de l'industrie locale, les infrastructures disponibles, la densité de la population et la densité du tissu économique local, sont des facteurs auxquels la survie des jeunes entreprises peut être sensible.

Pourtant, l'une des décisions les plus importantes au début d'un projet d'entreprise,

1. Insee Première•n° 1852• Avril 2021.

consiste à choisir le lieu où produire, et la littérature en économie spatiale nous apprend que, théoriquement, ce choix résulte d'un arbitrage rationnel visant à maximiser le profit ou à minimiser les coûts de production. L'entrepreneur, conscient des risques associés à son projet d'entreprise, choisit donc où s'implanter après une minutieuse évaluation des caractéristiques de la localisation, et il serait logique de penser que les caractéristiques qui ont motivé son choix, contribueraient de même à sa réussite, mais cela n'est pas systématiquement le cas.

Plusieurs études ([Brixy & Grotz, 2007](#); [Huiban, 2011](#)) ont montré que certaines zones géographiques, en l'occurrence les zones urbaines, enregistrent à la fois un fort taux de création, et de fermeture d'entreprises. Les entreprises, effectuent-elles mal leur choix de localisation? Ou s'agit-il simplement d'un choix par défaut pour certaines, car s'il est possible d'identifier la localisation optimale, s'y installer peut s'avérer difficile en raison de diverses contraintes (financières, par exemple). Ces différentes interrogations, ont motivé le présent chapitre. Nous proposons alors de déterminer la propension de l'entreprise à choisir sa localisation observée, et de prendre ce choix en compte dans l'analyse de sa survie. Ce chapitre, au même titre que les études portant sur les jeunes entreprises, contribue donc à la littérature en étudiant la survie de 45 268 jeunes entreprises en France, suivies sur une période de 5 ans depuis leur naissance. L'accent est mis sur les caractéristiques du territoire d'implantation de l'entreprise. L'originalité de ce travail est qu'il prend en considération la propension des jeunes entreprises à choisir leur localisation observée.

Dans le reste du chapitre, nous présentons d'abord une revue de littérature, sur le choix de localisation des entreprises. Le but de cette revue est de présenter les théories majeures expliquant le choix de localisation des entreprises, et de compléter cela par quelques récents résultats empiriques. Nous avons ensuite revu les études mettant en relation la survie des jeunes entreprises et leur environnement spatial. Pour finir, nous avons présenté les modèles économétriques estimés, les données et variables utilisées, les résultats d'estimation et la conclusion.

2.1 Choix de localisation des entreprises

2.1.1 Littérature théorique

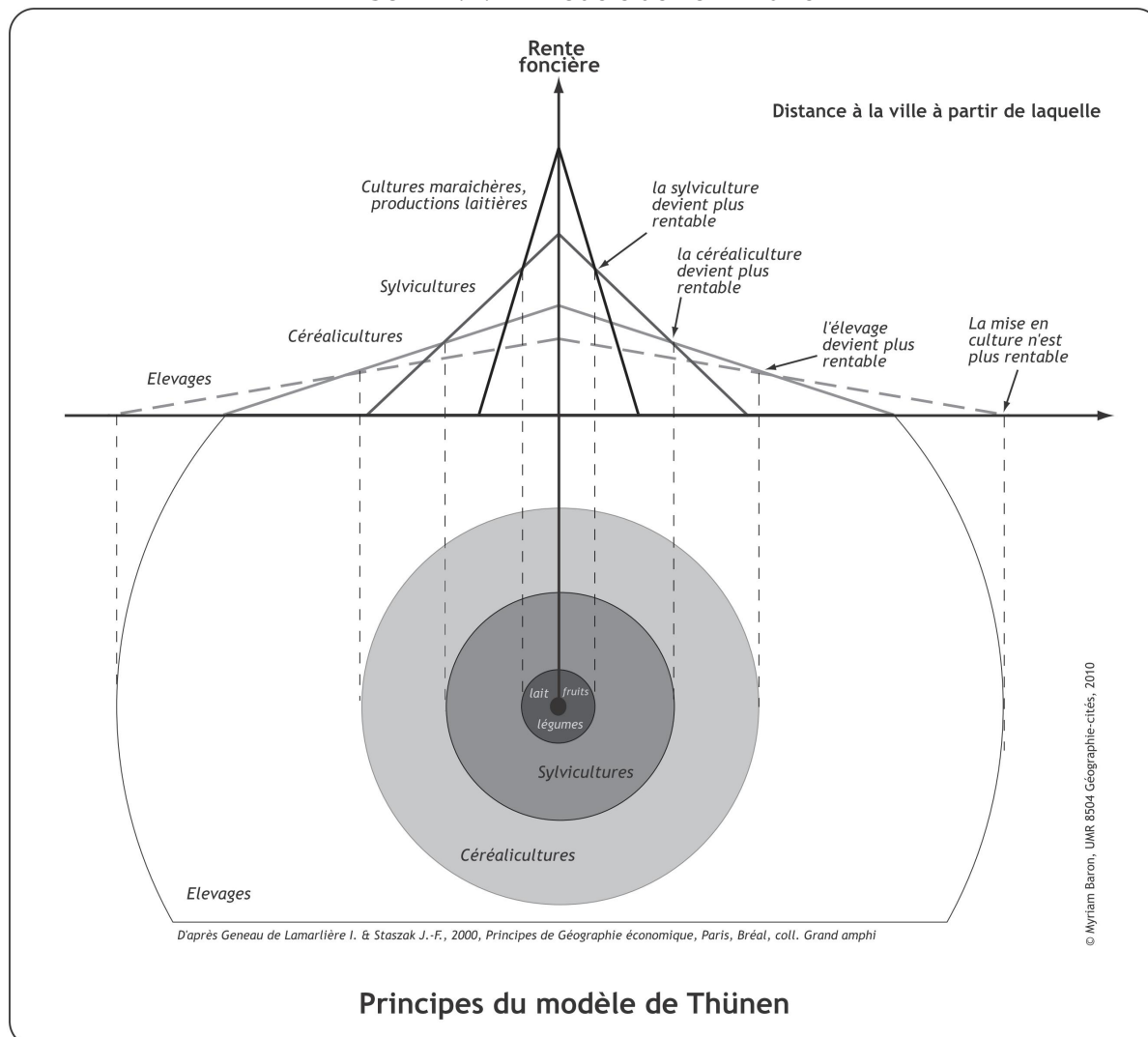
Les développements récents opérés dans l'analyse économique spatiale se sont construits sur la base de quatre grandes approches pionnières que [Ponsard \(1988\)](#) appelle

paradigmes. Ces modèles de référence ne sont autres que les travaux de Von Thünen (1826), Alfred Weber (1909), Hotelling (1929), Christaller (1933) et Lösch (1940).

Généralement appelé «père des théories de la localisation», Von Thünen (1826) est l'un des pionniers de l'analyse économique spatiale. A partir de son expérience pratique dans le Mecklembourg au sud du Rostock (Allemagne), il modélise la location des activités agricoles dans un contexte de ville-marché unique. Les localisations optimales des activités agricoles sont telles qu'en tout point de l'espace la rente foncière est maximisée (Ponsard, 1988). Deux variables principales sont prises en compte pour une localisation optimale : le type de production et la distance au centre (ville-marché) (Myriam Baron, 2013). La rente foncière maximale décroît en fonction de la distance au marché, et cette décroissance varie d'une culture à une autre. Ainsi les cultures pour lesquelles la rente foncière maximale est très élevée et décroît très rapidement en fonction de la distance (cultures maraîchères et les troupeaux de bêtes liés à la production laitière) seront les plus proches du centre, ensuite viendront les cultures pour lesquelles les rentes foncières maximales sont moins élevées que les précédentes mais décroissent moins rapidement en fonction de la distance. Selon un ordre croissant de distance au centre, nous avons le rangement suivant : cultures maraîchères et troupeaux de bêtes liés à la production laitière ; exploitations forestières ; céréales ; troupeaux de bêtes élevées pour de la viande. La rente foncière est obtenue par la formule suivante $R = r(p - c) - r \times T \times m$ où R représente la rente foncière ; r le rendement par unité de surface ; p le prix du marché par unité de produit ; c les charges de fabrication par unité de produit ; T le coût du transport, en unité de compte par unité de produit et par unité de

distance; et m la distance au marché.

FIGURE 2.1.1 – Modèle de Von Thünen



Source : Myriam Baron (2013)

Ainsi, Von Thünen dans son modèle, était principalement intéressé par la question «où produire quoi?», ce qui définit un cadre d'analyse très spécifique. Weber(1909) développe alors un modèle d'une plus grande généralité qui constitue l'une des bases les plus importantes des différentes théories de la localisation des activités de production (Ponsard, 1988).

Dans le modèle de Weber (1909), nous sommes en face d'un établissement industriel qui, à partir de deux intrants disponibles sur deux marchés distincts, produit un bien livré sur un marché donné. Le but de l'établissement consiste à trouver une localisation optimale

par rapport à ces trois marchés, afin de rentabiliser au mieux son activité de production. Weber (1909) modélise ainsi le choix de localisation de cet établissement comme étant le résultat d'une minimisation du coût total de transport.

La résolution du problème d'optimisation posé par le modèle de Weber (1909), repose sur un certain nombre d'hypothèses importantes (Perreur, 1988). La première hypothèse est celle d'un sous-espace euclidien convexe. D'après cette hypothèse, lorsque deux points A et B appartiennent au sous-espace considéré, il est toujours possible d'effectuer le trajet $[AB]$ sans quitter le sous-espace. La seconde hypothèse est celle de l'isotropie de l'espace pour les coûts de transport. Cela suppose qu'indépendamment de la direction, les coûts de transport sont identiques pour une même distance. Cette hypothèse épargne la nécessité de connaître le sens des trajets. Aussi, faut-il préciser que le coût de transport n'est pas seulement fonction de la distance, mais il est également proportionnel aux quantités. L'hypothèse suivante suppose un espace homogène qui présente en tout point les mêmes propriétés. Cette hypothèse est une sorte de *ceteris paribus* (Toutes choses égales par ailleurs), qui permet de faire abstraction des disparités spatiales. Pour finir, le marché est supposé être en situation de concurrence pure et parfaite.

Considérons à présent une entreprise qui produit une quantité m_3 d'output qu'elle commercialise sur un marché M_3 , et que m_1 et m_2 respectivement disponibles sur les marchés M_1 et M_2 sont les quantités d'intrants nécessaires à la production de m_3 . Les points M_i ont pour coordonnées géographiques (a_i, b_i) avec $i = 1, 2, 3$. Selon Weber, le problème du producteur consiste à minimiser par rapport à (x, y) (coordonnées de la localisation optimale) la fonction de coût de transport total :

$$T = tm_1d_1 + tm_2d_2 + tm_3d_3$$

avec t le coût de transport unitaire, d_i est la distance euclidienne entre la localisation optimale de l'entreprise et le marché M_i , $i = 1, 2, 3$. De façon générale on a :

$$T = \sum_{i=1}^3 t_i m_i [(a_i - x)^2 + (b_i - y)^2]^{\frac{1}{2}}$$

Les conditions d'optimalité de premier ordre donnent :

$$\frac{\partial T}{\partial x} = \sum_{i=1}^3 t_i m_i \frac{a_i - x}{[(a_i - x)^2 + (b_i - y)^2]^{\frac{1}{2}}} = 0$$

$$\frac{\partial T}{\partial y} = \sum_{i=1}^3 t_i m_i \frac{a_i - x}{[(b_i - x)^2 + (b_i - y)^2]^{\frac{1}{2}}} = 0$$

On remarque qu'il est impossible de déterminer analytiquement le couple (x,y) à partir des conditions de premier ordre. Plusieurs auteurs (Kulin & Kuenne, 1962; Bellman, 1965; Vergin & Rogers, 1967; Kuenne & Soland, 1972) s'intéresseront donc à la résolution numérique du problème de Weber. Schärli (1973) propose comme solution le centre de gravité à la puissance n , $n = 1, \dots, 6$, en retenant le n pour lequel on obtient le plus faible coût de transport :

$$x = \frac{\sum_{i=1}^3 m_i^n a_i}{\sum_i m_i^n}$$

$$y = \frac{\sum_{i=1}^3 m_i^n b_i}{\sum_i m_i^n}$$

Witzgall (1965) démontre que la localisation optimale dans un problème de Weber est le point M_i (s'il existe) associé à une masse m_i telle que : $m_i > \sum_{j \neq i} m_j$, m_i est dit poids dominant. En d'autres termes, lorsque la quantité d'intrants ou d'output, associée à un marché, est supérieure à la somme des quantités d'intrants ou d'output associées aux autres marchés, ce marché est dit dominant et constitue la localisation optimale.

Bien que Weber dans son approche, ait retenu comme critère, la minimisation du coût de transport, il admet cependant que plusieurs autres facteurs peuvent entrer en compte dans la localisation des activités de production. Il s'intéresse ainsi à l'impact des économies de main d'œuvre, et les économies d'agglomération sur la localisation optimale. En effet, Weber admet d'un côté que l'existence d'une main d'œuvre à faible coût en certains points de l'espace pourrait provoquer une déviation par rapport au point minimum du coût de transport, et d'un autre côté, l'existence d'un avantage lié à la concentration de la production en un seul lieu pourrait entraîner le regroupement des différentes unités de production d'une même firme, auparavant localisées respectivement au point minimum du coût de transport (Perreur, 1988). L'analyse de Weber attire ainsi l'attention sur trois principaux

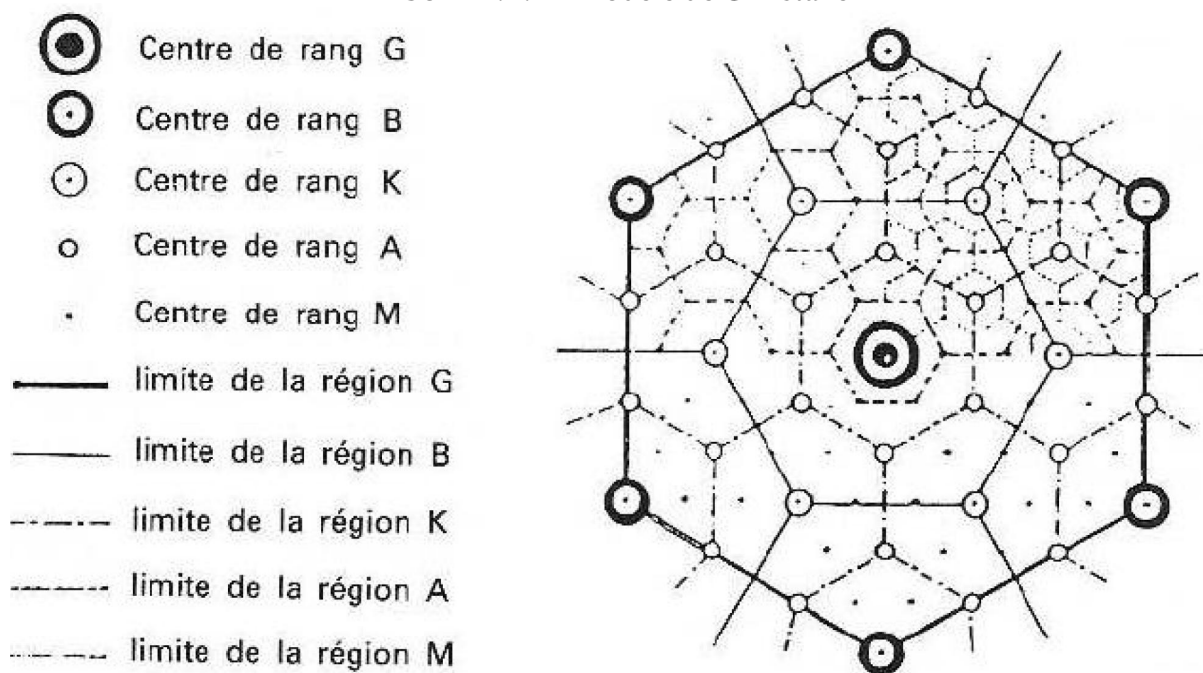
facteurs de localisation : les coûts de transport, les économies de main d'œuvre et les économies d'agglomération. Mais contrairement au modèle de Weber où le raisonnement était mené par rapport à un établissement, Hotelling (1929) propose un cadre d'analyse basé sur deux entreprises en interactions stratégiques (situation de duopole).

Dans son analyse, Hotelling (1929) introduit un aspect spatial dans la concurrence sur le marché des biens et services. Il lève l'hypothèse selon laquelle tous les consommateurs vont systématiquement vers le producteur qui propose le prix le plus faible, et suppose que la demande est une fonction continue de la différence des prix. Ainsi selon Hotelling, même si un producteur augmente son prix, certains clients lui resteront toujours fidèles pour plusieurs raisons possibles : la proximité géographique, des frais de transport abordables, la qualité des produits, l'existence de plusieurs variétés de produits. Il existe donc un cercle de clients qui font de chaque producteur un monopole dans une région bien délimitée (Hotelling, 1929). Pour son modèle, Hotelling suppose un marché de duopole représenté par une droite horizontale sur laquelle les consommateurs sont uniformément distribués ; une demande parfaitement inélastique ; des clients qui choisiront d'aller vers un producteur ou un autre en comparant la somme du prix et des coûts de transport. Le coût de transport inclut donc tous les autres éléments qui conditionnent la préférence des consommateurs pour un tel producteur ou un tel autre. A l'issue de l'analyse, le centre du marché apparaît comme étant la localisation optimale pour les deux entreprises en situation de duopole. L'analyse de Hotelling sera suivie par celles de Christaller et de Losch, qui tentent d'expliquer le paysage économique des villes.

Le géographe Walter Christaller (1933) et l'économiste August Losch (1940) sont connus pour la célèbre théorie des lieux centraux. Bien que les deux auteurs aient travaillé de façon indépendante (Claval, 1966), ils aboutissent à des résultats similaires. Christaller fut le premier à publier ses travaux basés sur une approche non formalisée. Dans son travail, Christaller s'est principalement appesanti sur les centres urbains dits «lieux centraux» qui sont des centres où les régions qui les entourent s'approvisionnent en biens («biens centraux») qu'elles ne produisent pas. Il définit les concepts de «portée d'un bien central» (distance maximale qu'un consommateur est prêt à parcourir pour acquérir le bien dans le centre où il est offert) ; et de «seuil d'un bien» (quantité consommée et produite minimale requise pour assurer un profit positif au producteur) (Beguin, 1988). La portée et le seuil d'un bien sont représentés par des disques centrés sur le lieu d'offre de ce bien, et le bien n'est produit que si le disque de sa portée dépasse celui de son seuil ; l'anneau séparant les deux circonférences est la projection spatiale du profit du producteur (Beguin, 1988). En partant d'un certain nombre de villes distinctes et fournissant un même type de bien central (de

portée 21)² Christaller construit un système de lieux centraux en raisonnant de façon hiérarchique. En effet, à partir d'une ville fournissant un bien de portée 21km, il est impossible de fournir à tous les consommateurs un bien de portée 20 par exemple. Ainsi, pour éviter ce problème, des biens de portée 20km seront offerts au centre de triangle donc les sommets sont les villes où sont fournis les biens de niveaux 21. Le même raisonnement est mené pour les biens de portée 19km, 18km, etc... D'après Christaller, les lieux centraux doivent constituer un treillis hexagonal pour couvrir tout le territoire (voir figure ci-dessous).

FIGURE 2.1.2 – Modèle de Christaller

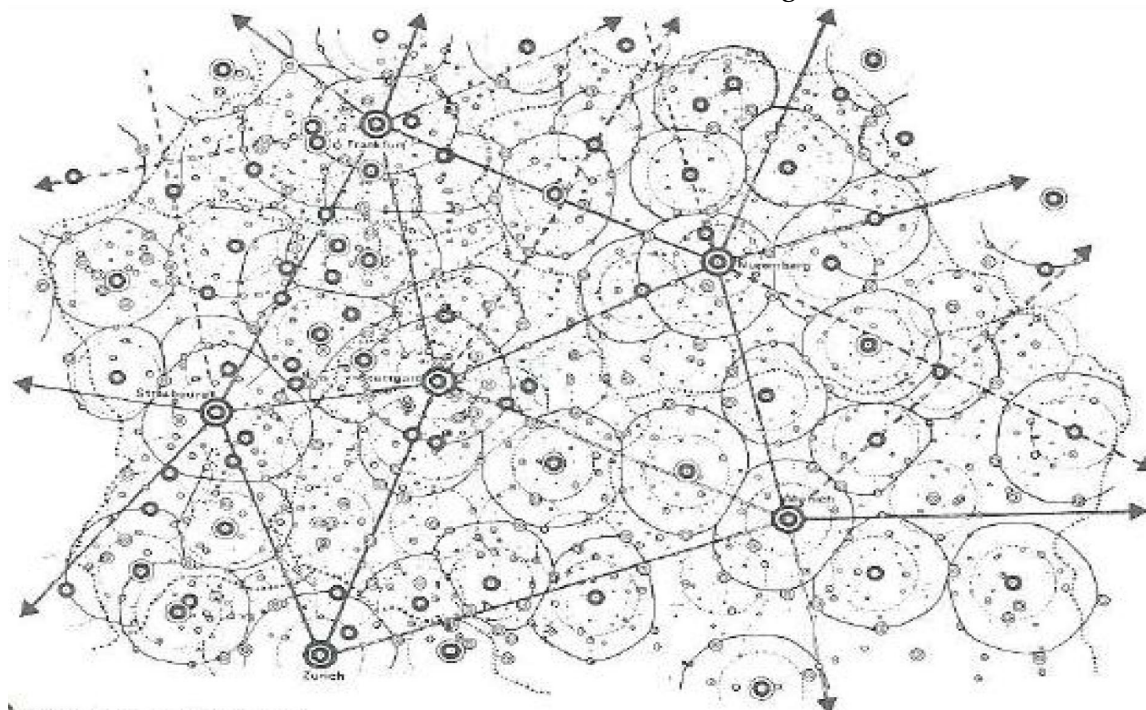


source : Christaller (1933)

En 1966, Christaller effectue une vérification de sa théorie en Allemagne du sud (cf figure ci-dessous) et constate une répartition des villes proche du treillis hexagonal théorique.

2. [Beguin \(1988\)](#)

FIGURE 2.1.3 – Lieux centraux en Allemagne du sud

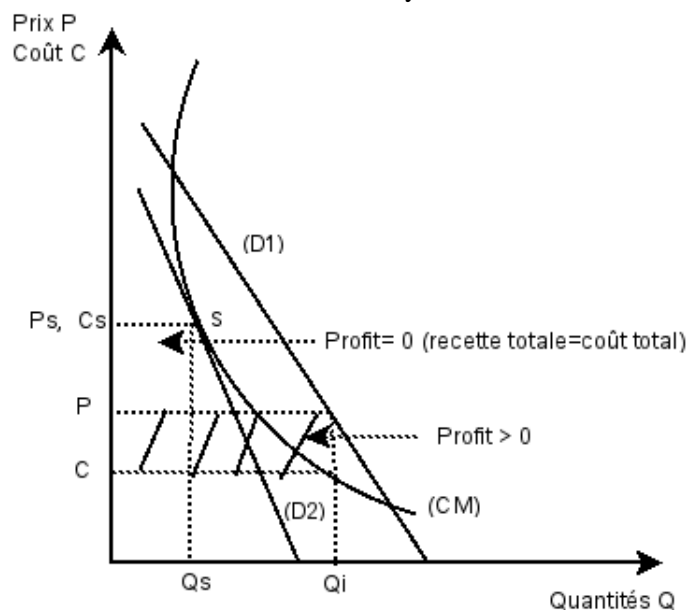


source : Christaller (1966)

Lösch (1940) a une approche beaucoup plus formalisée que celle de Christaller (1933). Il part d'une économie d'autosubsistance sans villes avec un seul bien à production concentrée. Il estime que l'aire de marché d'un bien est conjointement tributaire des forces favorisant la concentration géographique et qui sont sources d'économies d'échelle, puis des forces de dispersion, desquelles découlent des coûts de transport (Beguin, 1988). En reprenant le concept de portée maximale, Lösch estime que la demande individuelle d'un bien sera maximale au lieu de production et nulle à la distance correspondant à la portée.

Il trace sur un même graphique (cf figure ci-dessous) la courbe du coût moyen (courbe CM) et celle de la demande totale (courbes D), le bien ne sera donc produit que si deux courbes se coupent ($D1 \cap CM$) ou sont au moins tangentes ($D2 \cap CM=S$). En ordonnées sont représentés le prix (P) et les valeurs du coût moyen (C), et en abscisse nous avons la quantité.

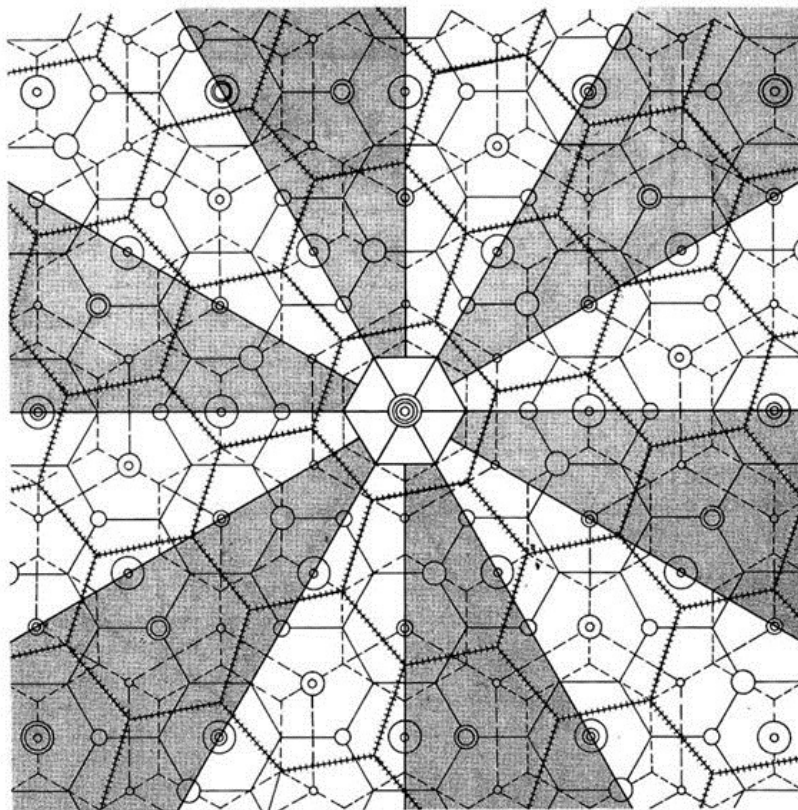
FIGURE 2.1.4 – Demande et coût moyen (Modèle de Lösch (1940))



source : (Beguin, 1988)

Au départ, l'espace sera couvert d'aires de marché circulaires et tangentes dues à l'installation progressive de plusieurs producteurs. Ces aires de marché circulaires conduiront à l'existence d'une demande insatisfaite, celle localisée entre les disques tangents. D'après Lösch, pour satisfaire donc cette demande, des aires hexagonales beaucoup plus rétrécies se substitueront aux aires circulaires, en remplissant la totalité de l'espace. L'entrée de nouveaux producteurs entrainera la baisse de la demande totale adressée à chaque producteur qui passera de D1 à D2. Le nouveau point d'équilibre s'établira au point S, où le profit est nul. Ce point n'est autre que le seuil de Christaller.

Contrairement à Christaller (1933), qui postulait le remplissage de l'espace par des aires de marché hexagonales, Lösch (1940) justifie cela par la libre entrée de producteurs désireux de réaliser un profit (avec l'appui du consommateur qui voit son coût de transport se réduire).



source : A. Lösch(1954), fig. 28.

De ces modèles pionniers, il ressort que la localisation des activités de production est un choix rationnel, tributaire de quatre grandes catégories de facteurs. Il s'agit du coût de transport, du type d'activité, des économies de main d'œuvre, et des économies d'agglomération. Ces différents facteurs peuvent également donner lieu à une concentration spatiale des industries, et cela est parfois la source d'importants avantages pour les entreprises.

2.1.2 Littérature empirique

Les modèles séminaux de Von Thünen (1826), Alfred Weber (1909), Hotelling (1929), Christaller (1933) et Lösch(1940) ont ouvert la voie à une riche littérature qui continue d'améliorer notre compréhension du choix de localisation des entreprises. Depuis l'introduction de la notion d'agglomération par Weber, et les développements apportés par Marshall (1920), plusieurs travaux théoriques ont essayé d'identifier, et d'expliquer les mécanismes qui régissent la concentration spatiale des entreprises. Ces travaux théoriques ont

donné naissance à une littérature empirique qui, d'une part, confronte les conclusions théoriques à la réalité des données, et d'autre part, met en exergue des résultats qui n'apparaissent pas forcément dans la littérature théorique.

T. J. Holmes (1999) s'intéresse à la vérification empirique de deux idées assez complémentaires, qui apparaissent respectivement dans les travaux de Marshall (1920) et de Stigler (1951). Marshall (1920) explique que les entreprises spécialisées dans la production d'inputs (biens intermédiaires) ont plus tendance à se localiser dans les aires de concentration géographique d'entreprises d'une même industrie (district industriel). Quant à Stigler (1951), il émet l'idée selon laquelle les entreprises situées dans des districts industriels, pourraient davantage recourir à la sous-traitance (désintégration verticale). Dans son étude, T. J. Holmes (1999) définit le pourcentage d'inputs achetés par rapport à la recette totale, comme mesure de la désintégration verticale. Ses analyses montrent que les entreprises situées dans un voisinage d'entreprises de la même industrie, sont beaucoup plus verticalement désintégrées. Même si ce résultat n'établit pas une relation de causalité, il pousse à considérer la désintégration verticale comme potentiel facteur de location des entreprises. De manière implicite, les conclusions de T. J. Holmes (1999) indiquent que la possibilité d'avoir recours à la désintégration verticale pourrait motiver la localisation des entreprises dans des districts industriels. En effet, l'internalisation de certaines activités intermédiaires du processus de production étant parfois très coûteuse, il peut être financièrement avantageux pour l'entreprise de déléguer ces activités à des firmes spécialisées. Le travail de T. J. Holmes (1999) nous montre d'une certaine manière que la compréhension des forces à l'origine de la concentration spatiale des entreprises, pourrait permettre de mieux cerner les facteurs qui déterminent leur choix de localisation. Plusieurs travaux théoriques à ce sujet (Helsley & Strange, 1990; Goldstein & Gronberg, 1984; Glaeser, 1999) mettent l'accent sur trois déterminants de la concentration spatiale des entreprises : la mutualisation du marché de travail, le partage d'inputs et la diffusion de la connaissance. Ces trois déterminants sont au centre du papier de S. S. Rosenthal & Strange (2001). Ces derniers étudient empiriquement, dans quelle mesure ces trois facteurs sont prédictifs de la concentration spatiale de l'industrie manufacturière américaine, qui est mesurée par l'indice d'Ellison & Glaeser (1997).

La notion de mutualisation du marché de travail, vient du fait que l'existence d'un grand bassin de travailleurs, faciliterait le recrutement des compétences spécifiques. Les entreprises d'une même industrie, en s'implantant les unes près des autres, peuvent attirer un grand nombre de travailleurs qualifiés, et simplifier ainsi l'appariement employés-employeurs. Quant au concept de partage d'inputs, il rejoint la notion de désintégration

verticale qu'on retrouve chez [T. J. Holmes \(1999\)](#). Les industries manufacturières au sein desquelles l'on note un fort usage d'inputs manufacturiers, peuvent avoir un intérêt financier (réduction des coûts) à se concentrer spatialement. [S. S. Rosenthal & Strange \(2001\)](#) mesure cet aspect en calculant pour chaque industrie, la part d'inputs par dollar de marchandises expédiées. Pour finir, les auteurs calculent la valeur des nouveaux produits introduits (innovation) par dollar de marchandises expédiées, comme indicateur de diffusion de la connaissance. Les entreprises opérant dans des industries à fort potentiel d'innovation, peuvent privilégier une concentration spatiale afin de profiter d'une exposition à de nouvelles idées émanant de l'industrie.

[S. S. Rosenthal & Strange \(2001\)](#) mènent leur analyse à partir de différentes unités géographiques. Ils considèrent la dimension de l'État, du département, et du code postal. Les résultats montrent globalement un effet positif des trois facteurs sur la concentration spatiale des industries. Néanmoins, certains impacts sont spatialement limités, en l'occurrence ceux du partage des inputs et de la diffusion de la connaissance. Le partage d'inputs favorise principalement la concentration des industries à l'échelle de l'État, et n'a qu'un faible impact à des dimensions géographiques plus fines. La diffusion des connaissances n'a d'effets qu'à l'échelle des codes postaux, reflétant ainsi le caractère localisé de l'innovation. Quant à la mutualisation du marché de travail, elle impacte la concentration spatiale industrielle à toutes les échelles géographiques considérées. Cet effet statistiquement robuste, pourrait, selon les auteurs, provenir du fait que les travailleurs qualifiés n'aient pas à quitter leurs zones géographiques pour saisir de nouvelles opportunités d'emplois. Les conclusions de [S. S. Rosenthal & Strange \(2001\)](#), rejoignent celles de plusieurs autres auteurs ayant travaillé sur différentes sources de données.

[Jofre-Monseny et al. \(2011\)](#) étudient via une régression de poisson, le choix de localisation des jeunes entreprises de l'industrie manufacturière espagnole. Ils considèrent l'échelle géographique de la ville et celle des petites municipalités au sein des grandes villes, puis modélisent le nombre de jeunes entreprises en fonction de plusieurs variables approximant principalement la mutualisation du marché du travail, le partage d'inputs, et la diffusion de la connaissance. D'après les résultats, la mutualisation du marché du travail et le partage d'input ont les effets les plus robustes (statistiquement) quel que soit l'échelle géographique. Les jeunes entreprises s'implantent plus dans les lieux où les industries présentes, emploient une main d'œuvre similaire à la leur (jeunes entreprises), et où il y a également un nombre important de producteurs d'inputs. Concernant la diffusion de la connaissance, [Jofre-Monseny et al. \(2011\)](#) trouvent, comme dans le cas de [S. S. Rosenthal & Strange \(2001\)](#), un effet limité dans l'espace. Leurs résultats montrent que la préférence

des jeunes entreprises pour les zones géographiques où les industries utilisent les mêmes technologies qu'elles, n'est significative qu'à l'échelle réduite de la municipalité.

[Lasch et al. \(2013\)](#) travaillent sur des entreprises françaises du secteur des technologies de l'information et de la communication (TIC). Bien que les auteurs trouvent un impact positif de la diffusion de connaissance sur le choix de localisation des jeunes entreprises, ils notent une contribution assez faible de ce facteur, et rejoignent ainsi les conclusions de [S. S. Rosenthal & Strange \(2001\)](#) et de [Jofre-Monseny et al. \(2011\)](#).

La nature de la connaissance, ainsi que son accessibilité, jouent également un rôle important dans le choix de localisation des entreprises. [Audretsch et al. \(2005\)](#) montrent que les jeunes entreprises allemandes du secteur de la technologie ont tendance à s'éloigner des universités produisant beaucoup de connaissances en sciences naturelles (nombre publications), et préfèrent la proximité avec les universités produisant plus de connaissances en sciences sociales. [Baptista et al. \(2011\)](#) montrent à partir de données portugaises, que les régions dotées d'un nouvel établissement d'enseignement supérieur, enregistrent d'un côté, une croissance du nombre d'entreprises créées dans les secteurs faisant un usage intensif de la technologie, et de l'autre côté, une décroissance de créations d'entreprises dans des secteurs à faible usage de technologie. Ces deux papiers soulignent le fait que l'existence d'une source de connaissance ne rend pas systématiquement une région attractive. D'une part, le décryptage de la connaissance peut nécessiter des ressources supplémentaires (compétences) dont l'entreprise ne dispose pas, et d'autre part, cette connaissance peut s'avérer inutile en raison du secteur de l'activité de l'entreprise.

L'accès à des ressources naturelles et la présence d'infrastructures de transport, déterminent aussi le choix de localisation des entreprises. [Naldi et al. \(2021\)](#) travaillent sur des données d'entreprises suédoises et trouvent que les aménités naturelles (réserves naturelles, emplacements côtiers, zones de préservation) et culturelles (patrimoines bâtis) contribuent à la création d'entreprises, particulièrement dans les zones rurales. Concernant les infrastructures de transport, [Chang & Zheng \(2022\)](#) trouvent un effet mitigé en analysant l'impact des lignes à grande vitesse (plus de 250km/h) sur la location des entreprises chinoises. D'après leurs résultats, les zones urbaines pourvues de lignes à grande vitesse attirent les entreprises du secteur service alors que les entreprises du secteur primaire manufacturier préfèrent migrer vers les localités sans lignes à grande vitesse. D'après les auteurs, ce résultat serait dû au fait que les entreprises du secteur service, tirent profit de la croissance démographique et de l'intégration du marché dans les zones pourvues d'infrastructure de transport, tandis que les entreprises du secteur primaire manufacturier ont

du mal à supporter les coûts liés à l'implantation dans de telles zones.

2.2 Environnement spatial et survie des jeunes entreprises

S'il existe une littérature sur la survie des entreprises, peu d'études prennent en compte l'environnement spatial de l'entreprise, et encore moins travaillent sur des jeunes entreprises. Les entreprises transforment aussi bien leur environnement, qu'elles sont influencées par ce dernier, mais seules les entreprises suffisamment établies ont des impacts conséquents (Shane, 2009). Les jeunes entreprises quant à elles, sont exposées à des risques d'échec très élevés dès leurs premières années, et vont plus avoir tendance à prendre leur environnement spatial comme donnée et s'y adapter, car elles sont moins en mesure de la transformer. Analyser donc l'influence de l'environnement spatial sur la survie à partir d'un échantillon d'entreprises très âgées peut masquer une éventuelle causalité inverse.

L'environnement spatial idéal pour une jeune entreprise, est celui où elle peut facilement accéder aux différentes ressources nécessaires à sa prospérité. Les jeunes entreprises opèrent généralement sur les marchés locaux et ont besoin d'une demande conséquente pour croître. Il leur faut également une main d'œuvre de qualité pour mieux satisfaire la demande, des infrastructures de transport, et surtout de bonnes informations relatives à leur secteur d'activité. Les disparités des régions quant à la dotation en ces ressources, peuvent conduire à une hétérogénéité spatiale des risques d'échec des jeunes entreprises.

Fotopoulos & Louri (2000) travaillent sur un échantillon de jeunes entreprises grecques et trouvent que les zones urbaines sont favorables à la survie des entreprises. De même, Wennberg & Lindqvist (2010) sur la base d'un échantillon d'entreprises suédoises, montrent que la densité de population impacte positivement la survie. Ces résultats paraissent intuitifs quand on pense aux externalités (infrastructures, marchés locaux, informations) positives associées à l'urbanisation. Dans ces zones, les jeunes entreprises pourront d'une part côtoyer d'autres entreprises du même secteur et bénéficier des économies d'agglomération, et d'autre part profiter de la diversité industrielle qui est une source d'innovation (Jacobs, 1969).

Les résultats de Wennberg & Lindqvist (2010) suggèrent que les jeunes entreprises implantées dans des régions où leur industrie d'appartenance est concentrée, ont plus de chances de survivre. Les auteurs montrent qu'au-delà du fait de survivre, ces entreprises ont de meilleures performances économiques en termes de création d'emplois. Renski (2011) travaille sur de jeunes entreprises américaines. Pour prendre en compte les hété-

rogénéités entre les différentes industries, il considère d'un côté les entreprises au sein de cinq industries manufacturières, et de l'autre côté trois secteurs de service. L'auteur calcule la proportion d'entreprises de la même industrie comme mesure de la localisation industrielle, et l'inverse de la différence absolue entre la part locale de l'emploi dans chaque industrie et la part nationale, comme une approximation de la diversité industrielle. Les résultats de régressions par secteurs, montrent dans plus de la moitié des industrie étudiées, que les externalités de localisation et la diversité industrielle réduisent significativement le risque d'échec des jeunes entreprises.

Les travaux de [Howell et al. \(2018\)](#) confirment l'impact positif de la diversité industrielle sur la base d'un échantillon de 135 000 entreprises chinoises observées entre 1998-2007. [Ebert et al. \(2019\)](#) eux, ne trouvent aucun impact significatif de la diversité industrielle. Quant aux externalités de localisation, elles ne favorisent, d'après leurs résultats, que les jeunes entreprises non-technologiques et non innovantes. Les avantages liés à la localisation dans des zones urbaines bien développées, viennent souvent avec des coûts qui peuvent être de nature financière (prix du logement, niveau de salaire) ou de nature stratégique (niveau de concurrence). Les jeunes entreprises qui ne peuvent pas supporter ces coûts, font faces à des chances de survie réduites. Les travaux de [Huiban \(2011\)](#) et de [Brixy & Grotz \(2007\)](#), portant respectivement sur des échantillons d'entreprises françaises et allemandes, montrent que les entreprises ont de faibles chances de survivre en zone urbaine. S'éloigner de ces zones peut s'avérer être une stratégie bénéfique pour ces entreprises. [Manzato et al. \(2011\)](#) se focalisent sur les infrastructures de transport, et travaillent sur des données hollandaises. L'estimation de modèles de survie sur 286 874 entreprises montrent qu'être éloigné de certaines infrastructures telles que l'aéroport, gare de train intercity, autoroute, augmente les chances de survie.

Ces études montrent d'une manière ou d'une autre que l'environnement spatial n'est pas neutre dans la survie des jeunes entreprises. De plus, certains résultats mettent en évidence l'hétérogénéité de l'impact de l'environnement spatial. Il conviendrait donc de prendre cet élément en considération lorsqu'on étudie la survie des jeunes entreprises dans leur environnement spatial, et c'est justement ce que propose ce chapitre.

2.3 Analyse économétrique

Ce chapitre s'inscrit dans la lignée des travaux empiriques portant sur la survie des jeunes entreprises en fonction de leur environnement spatial. Ces travaux, contrairement, à

ceux portant sur des entreprises âgées, permettent de mieux cerner certaines réalités d'autant plus que la trajectoire de vie des entreprises âgées résultent de routines bien établies et de plusieurs années d'expérience parfois difficile à mesurer. Notre approche méthodologique consiste à estimer plusieurs modèles économétriques sur des données de jeunes entreprises françaises observées pendant leurs cinq premières années de vie. L'apport original de ce chapitre est que nous estimons la propension de l'entreprise à choisir sa localisation observée, et de prendre cela en compte dans l'analyse de sa survie.

Cette section est constituée de deux parties dans lesquelles nous présentons respectivement notre stratégie d'analyse de la survie, et la stratégie d'estimation du choix de localisation.

2.3.1 Stratégie d'analyse de la survie

La stratégie économétrique dans cette partie consiste à étudier la sensibilité de la probabilité de survie des jeunes entreprises aux caractéristiques de leur environnement spatial. Pour ce faire, nous avons choisi les modèles de survie car ils s'avèrent convenir le mieux à nos données. En effet, ces modèles, contrairement aux modèles de choix binaires (probit et logit), permettent de prendre en compte les problèmes de censure et la durée de vie des entreprises étudiées (information dont nous disposons).

Nous avons construit progressivement notre analyse en commençant par l'estimation des fonctions de survie de Kaplan-Meier (voir chapitre 1). Ensuite, nous avons estimé des modèles à hasards proportionnels Exponentiel et de Weibull, puis des modèles à hasards proportionnels de Cox (voir chapitre 1).

L'estimation des fonctions de survie de Kaplan-Meier nous permet d'avoir une première idée des hétérogénéités de durée de vie entre les entreprises de notre échantillon. Elle reste cependant une analyse descriptive permettant de visualiser les différences de survie pour une variable donnée, sans prendre en considération l'impact d'autres variables. Il est donc important de compléter l'analyse avec des régressions.

Les modèles de survie (paramétriques) nécessitent de choisir une loi de probabilité pour la durée de vie des individus. Ce choix peut être difficilement considéré comme étant objectif car il n'existe pas de méthode permettant de conjecturer de manière précise la distribution de la durée de vie. C'est pourquoi l'on a souvent recours à l'approche semi-paramétrique de Cox qui ne requiert pas de choisir une loi de probabilité pour la durée de vie des individus (voir chapitre 1).

Dans notre cas, nous avons étudié plusieurs possibilités (approche paramétrique et semi-paramétrique) afin de pouvoir choisir la modélisation qui convient le mieux à nos données. Nous avons testé l'éventualité que la durée de vie des entreprises de notre échantillon soit distribuée selon les lois Exponentiel et de Weibull, qui font partie des lois les plus simples et les plus utilisées dans les modèles de survie paramétriques. Nous avons ensuite estimé les modèles de survie associés à ces deux lois car les tests ne permettent pas d'aboutir à des conclusions totalement objectives sur le choix du modèle. Pour lever toute ambiguïté, nous avons complété notre analyse par l'estimation de modèles de Cox. Les sous-sections suivantes présentent un peu plus en détail notre stratégie.

2.3.1.1 Modèle à hasards proportionnels Exponentiel

Dans le modèle à hasards proportionnels exponentiel, on suppose que la durée de vie T suit une loi exponentielle de paramètre λ . Ainsi, les fonctions de répartition F et de densité f s'écrivent respectivement $1 - \exp(-\lambda t)$ et $\lambda \exp(-\lambda t)$. On sait que $S(t) = 1 - F(t)$, alors $S(t) = \exp(-\lambda t)$, or $h(t) = \frac{f(t)}{S(t)}$ donc $h(t) = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda$. Le modèle à hasards proportionnels exponentiel, sous forme matricielle, s'écrit :

$$h(t | X) = \lambda \exp(X\beta) = \exp(X\beta + \log(\lambda))$$

Lorsque la première colonne de la matrice $X_{n \times K}$ est composée de 1, et que la première composante du vecteur $\beta_{K \times 1}$ est $\log(\lambda)$, on peut écrire

$$h(t | X) = \exp(X\beta)$$

Sachant que $h(t) = -\frac{1}{S(t)} \frac{dS(t)}{dt}$ on peut écrire $S(t) = \exp(-\int_0^t h(u) du)$ d'où $S(t | X) = \exp(-t \exp(X\beta))$. Partant de $h(t) = \frac{f(t)}{S(t)}$ on obtient $f(t) = h(t)S(t)$, ainsi

$$f(t | X) = \exp(X\beta) \exp(-t \exp(X\beta))$$

Soit x_i le vecteur ligne de X correspondant aux variables explicatives de l'individu i . En notant $f(t_i; \beta; x_i)$ et $S(t_i; \beta; x_i)$ les fonctions de densité et de survie de la durée de vie de l'individu i , et en supposant que les données censurées sont indépendantes de celles non censurées, la vraisemblance $L(\beta)$ de l'échantillon des n individus peut s'écrire :

$$L(\beta) = \prod_{i=1}^n [f(t_i; \beta; x_i)]^{\delta_i} [S(t_i; \beta; x_i)]^{1-\delta_i}$$

où $\delta_i = 0$ si l'observation i est censurée et 1 sinon. La log vraisemblance $\log L(\beta)$ s'écrit

$$\log L(\beta) = \sum_{i=1}^n \delta_i x_i \beta - \sum_{i=1}^n t_i \exp(x_i \beta)$$

Les K éléments du score $S(\beta) = \frac{\partial \log L(\beta)}{\partial \beta}$ s'obtiennent par

$$S_k(\beta) = \sum_{i=1}^n x_{ki} [\delta_i - t_i \exp(x_i \beta)]$$

avec k allant de 1 à K . La matrice variances-covariances, soit $\Omega(\hat{\beta})$ de dimensions $K \times K$, est estimée en inversant la matrice d'information de Fisher $I(\hat{\beta})$:

$$\begin{aligned} \Omega(\hat{\beta}) &= [I(\hat{\beta})]^{-1} \\ &= \left[-\frac{\partial^2 \log L(\hat{\beta})}{\partial \hat{\beta}_k \partial \hat{\beta}'_k} \right]^{-1} \\ \Omega(\hat{\beta}) &= \left[\sum_{i=1}^n x_{ki} x'_{ki} t_i \exp(x_i \beta) \right]^{-1} \end{aligned}$$

La significativité statistique des paramètres $\hat{\beta}_k$ peut être testée de diverses manières. Dans la littérature empirique, il est assez fréquent de rencontrer le test de significativité de Wald et celui du ratio de vraisemblance.

Lorsque que l'estimateur $\hat{\beta}$ est convergent et suit asymptotiquement une distribution normale, et que $\Omega(\hat{\beta})$ est un estimateur convergent de la variance asymptotique de $\hat{\beta}$, la statistique de Wald $W = (\hat{\beta} - \beta_h)' \Omega^{-1}(\hat{\beta}) (\hat{\beta} - \beta_h)$ suit une distribution de Khi-deux à K degrés de liberté ($\chi^2_{(K)}$) sous l'hypothèse $H_0 : \hat{\beta} = \beta_h$, où β_h est une valeur hypothétique.

Quant au test du ratio de vraisemblance, il est basé sur la statistique $\Lambda = 2 \log L(\hat{\beta}) - 2 \log L(\beta_h)$ qui suit également une distribution de Khi-deux à K degrés de liberté ($\chi^2_{(K)}$) sous l'hypothèse $H_0 : \hat{\beta} = \beta_h$.

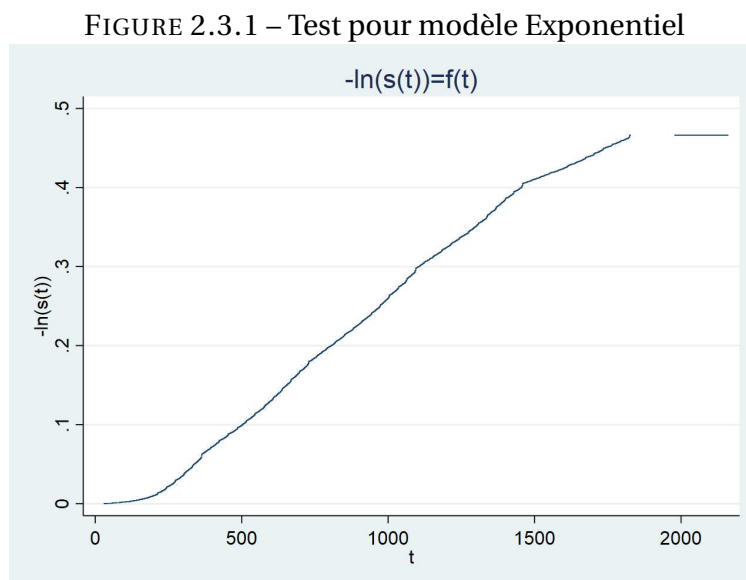
2.3.1.1.1 Choix d'un modèle Exponentiel

Lorsque la durée de vie suit une loi exponentielle de paramètre λ , on a $S(t) = \exp(-\lambda t)$.

$$\begin{aligned} S(t) &= \exp(-\lambda t) \\ \Rightarrow -\ln(S(t)) &= \lambda t \end{aligned}$$

La représentation graphique de la courbe $-\ln(S(t)) = f(t)$ aurait donc la forme d'une fonction affine si le modèle est exponentiel.

A partir de nos données, nous avons estimé et récupéré les fonctions de survie par la méthode de Kaplan-Meier. Nous avons ensuite représenté la courbe $-\ln(S(t)) = f(t)$ (figure 2.3.1)



Le graphique ci-dessous nous montre une fonction qui n'est pas totalement affine comme il est souvent le cas dans la pratique. Il revient ainsi au chercheur de décider que les courbures observées sont suffisamment négligeables pour choisir un modèle exponentiel. La courbe sur la figure 2.3.1 ne saurait, certes, pas être parfaitement assimilée à la représentation d'une fonction affine, mais peut toutefois servir de base pour l'estimation d'un modèle exponentiel.

2.3.1.2 Modèle à hasards proportionnels de Weibull

Dans le modèle à hasards proportionnels de Weibull, la durée de vie T suit une distribution de Weibull. Considérons $f(t) = \frac{\mathcal{K}}{\lambda} \left(\frac{t}{\lambda}\right)^{\mathcal{K}-1} \exp\left[-\left(\frac{t}{\lambda}\right)^{\mathcal{K}}\right]$ et $F(t) = 1 - \exp\left[-\left(\frac{t}{\lambda}\right)^{\mathcal{K}}\right]$ respectivement les fonctions de densité et de répartition d'une distribution de Weibull de paramètres \mathcal{K} (paramètre de forme) et λ (paramètre d'échelle).

$$S(t) = 1 - F(t)$$

$$\text{or } h(t) = \frac{f(t)}{S(t)} \text{ alors}$$

$$h(t) = \frac{\frac{\mathcal{K}}{\lambda} \left(\frac{t}{\lambda}\right)^{\mathcal{K}-1} \exp\left[-\left(\frac{t}{\lambda}\right)^{\mathcal{K}}\right]}{\exp\left[-\left(\frac{t}{\lambda}\right)^{\mathcal{K}}\right]}$$

$$\text{si on pose } \tilde{\lambda} = \frac{1}{\lambda}$$

$$\text{on peut écrire } h(t) = \tilde{\lambda} \mathcal{K} (\tilde{\lambda} t)^{\mathcal{K}-1}$$

Le modèle à hasards proportionnels de Weibull s'écrit :

$$h(t | X) = \tilde{\lambda} \mathcal{K} (\tilde{\lambda} t)^{\mathcal{K}-1} \exp(X\beta)$$

$$S(t | X) = \exp\{-\tilde{\lambda} t^{\mathcal{K}} \exp(X\beta)\}$$

La vraisemblance pour les n individus de l'échantillon s'écrit :

$$L(\beta; \tilde{\lambda}; \mathcal{K}) = \prod_{i=1}^n \left[\tilde{\lambda} \mathcal{K} (\tilde{\lambda} t)^{\mathcal{K}-1} \exp(x_i \beta) \right]^{\delta_i} \exp\{-\tilde{\lambda} t^{\mathcal{K}} \exp(x_i \beta)\}$$

La vraisemblance est maximisée par rapport aux paramètres $\beta; \tilde{\lambda}; \mathcal{K}$. Les tests de Wald et de ratio de vraisemblance peuvent être également construits de la même manière que dans le modèle à hasards proportionnels exponentiel précédemment décrit.

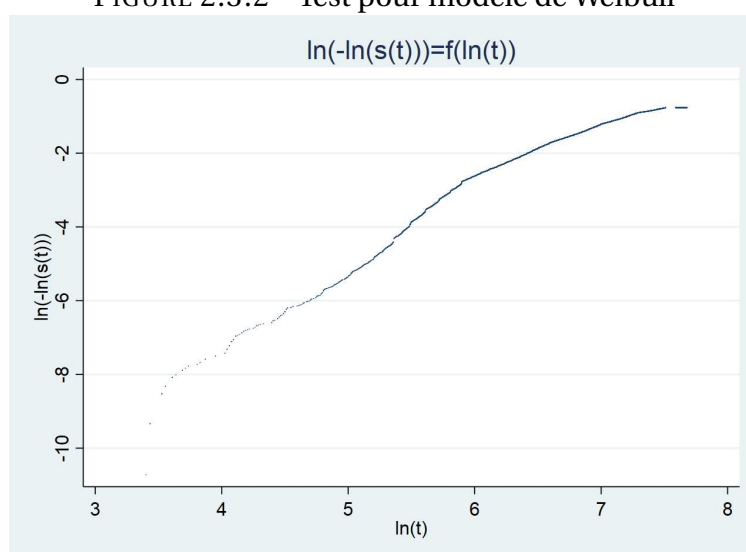
2.3.1.2.1 Choix d'un modèle de Weibull

Dans le cas d'un modèle de Weibull, on a :

$$\begin{aligned} S(t) &= \exp\left[-\left(\frac{t}{\lambda}\right)^{\mathcal{K}}\right] \\ \ln[-\ln(S(t))] &= -\mathcal{K} \ln(\lambda) + \mathcal{K} \ln(t) \end{aligned}$$

La courbe $\ln(-\ln(S(t))) = f(\ln(t))$ serait donc celle d'une fonction affine si la fonction de survie suit une distribution de Weibull. Comme dans le cas du modèle exponentiel (paragraphe 2.3.1.1.1), ce résultat théorique n'est jamais obtenu en pratique. La figure 2.3.2 nous montre une courbe qui n'est pas tout à fait celle d'une fonction affine. Nous avons tout de même estimé les deux modèles (Exponentiel et Weibull) et celui de Cox pour lever toute ambiguïté.

FIGURE 2.3.2 – Test pour modèle de Weibull



2.3.2 Stratégie d'estimation du choix de localisation

Lorsqu'une entreprise i_1 choisit une localisation j par défaut, certains attributs X_j de cette localisation peuvent constituer des obstacles pour l'entreprise et réduire ses chances de réussite. En revanche, ces mêmes attributs X_j peuvent être des atouts et augmenter les chances de survie d'une autre entreprise i_2 pour qui la localisation j n'est pas un choix par défaut, mais un choix optimal. Par exemple, l'urbanisation et la densité de la population peuvent d'un côté augmenter les chances de survie (Fotopoulos & Louri, 2000; Wennberg & Lindqvist, 2010) et de l'autre côté les réduire (Stearns et al., 1995; Brixy & Grotz, 2007; Huiban, 2011). Nous proposons dans ce travail d'estimer la propension de chaque entreprise à choisir sa localisation observée et de prendre cela en compte dans notre analyse de la survie. Nous supposons que cette propension sera d'autant élevée que l'entreprise considèrera son choix de localisation comme étant optimal. Pour estimer cette propension, nous utilisons un cadre d'analyse basé sur la maximisation d'un profit aléatoire espéré (voir Jofre-Monseny et al. (2011)).

Considérons une zone d'emploi j caractérisée par des attributs X_j , où l'entreprise i espère réaliser un niveau de profit π_{ij} . Nous retenons dans notre modèle la densité du tissu économique comme déterminant de π_{ij} car elle informe aussi bien sur le niveau d'activité économique (nombre d'entreprises) que sur le degré d'urbanisation (population) de la zone d'emploi. En supposant que π_{ij} est une combinaison linéaire de X_j et d'un facteur aléatoire non observé δ_{ij} , on a $\pi_{ij} = X_j\beta + \delta_{ij}$.

Soit Z_i la variable aléatoire désignant la zone d'emploi choisie par l'entreprise i et $P(Z_i = j/X_j) = P_{ij}$ la probabilité la probabilité que l'entreprise i choisisse la zone d'emploi j . Nous avons 321 zones d'emploi donc Z_i prend ses valeurs j dans $\{1, 2, \dots, 321\}$. En supposant que l'entreprise choisira la localisation où le profit espéré est le plus élevé et que δ_{ij} est distribué selon une loi d'extrémum généralisée de type 2, P_{ij} s'écrit sous forme de logit conditionnel :

$$P_{ij} = \frac{e^{X_j\beta}}{\sum_{j=1}^{321} e^{X_j\beta}}$$

La vraisemblance L pour un échantillon de taille n donne :

$$L = \prod_{i=1}^n \prod_{j=1}^{321} \left(\frac{e^{X_j\beta}}{\sum_{j=1}^{321} e^{X_j\beta}} \right)^{d_{ij}}$$

où $d_{ij} = \begin{cases} 1 & \text{si } i \text{ choisit } j \\ 0 & \text{sinon} \end{cases}$

La maximisation de la vraisemblance permet d'avoir une estimation $\hat{\beta}$ du coefficient qu'on utilisera ensuite pour calculer les prédictions \hat{P}_{ij} qui représentent ici la propension des entreprises à choisir leur localisation.

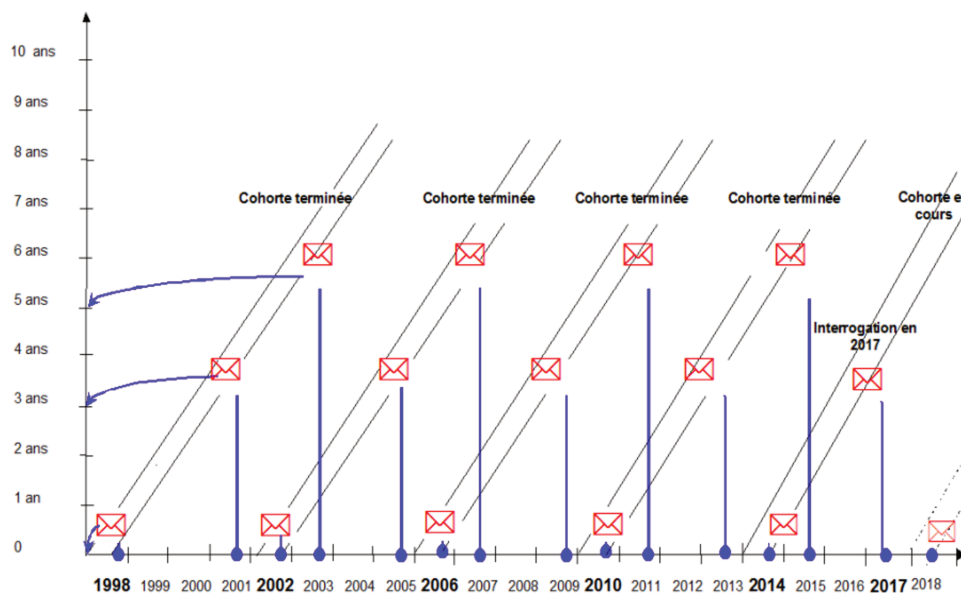
2.3.3 Données et variables

2.3.3.1 Données

Les données mobilisées dans ce chapitre proviennent du Système d'Information sur les Nouvelles Entreprises (SINE)³ et des bases; BPE, ADMIN EXPRESS et GEOFLA, atlas des zones d'emploi; décrites dans le chapitre 1.

Le SINE est un dispositif de suivi des jeunes entreprises sur une période de cinq années mis en place depuis 1994. Le but est d'analyser les conditions de développement et les problèmes rencontrés par les jeunes entreprises lors des cinq premières années de leur existence. Le suivi commence dès la naissance des entreprises et se déroule en trois étapes. La cohorte d'entreprises ciblées est successivement enquêtée dès la création, trois années après puis cinq années plus tard (pour celles qui ont survécu). L'enquête se déroule directement par voie postale et ne cible pas l'intégralité des entreprises créées en France. En 2010, l'échantillon interrogé comprenait 52 053 entreprises soit environ 42% du total des créations.⁴

FIGURE 2.3.3 – Calendrier SINE



Source : Descriptif du dispositif SINE

3. Accès via le CASD.

4. Source : Descriptif du dispositif SINE, accès via CASD.

Dans le cadre de ce chapitre, les entreprises étudiées sont créées en 2010 et donc suivies jusqu'en 2015. Elles opèrent dans les secteurs de l'industrie, de la construction, du commerce et des services, en France métropolitaine, dans les départements et régions d'outre-mer. L'enquête de 2010 exclue les entreprises ayant vécu moins d'un mois; les entreprises saisonnières, ou ayant une activité occasionnelle; les personnes morales de droit étranger; les entreprises ayant une activité de holding; les groupements de droit privé non dotés de la personnalité morale; les Coopératives d'Utilisation du Matériel Agricole en commun; les groupements d'intérêts économiques; les sociétés civiles immobilières; les loueurs; les collaborateurs, vacataires et stagiaires. Après la suppression des entreprises hors du champ de l'enquête et celles ayant cessé au démarrage, nous avons obtenu un échantillon final composé de 45 268 entreprises avec un taux de fermeture de 37,26%.

2.3.3.2 Variables explicatives

Le tableau 2.3.1 résume les variables utilisées. Elles mesurent les caractéristiques individuelles des entreprises, les caractéristiques des communes d'implantation et celles des zones d'emploi.

Les variables dénombrant les équipements sont calculées de la même façon que dans le chapitre 1 grâce aux informations apportées par la base permanente des équipements. La seule différence ici est que nous n'avons pas pu utiliser la superficie des communes en raison de plusieurs valeurs manquantes. Ainsi, pour chaque commune c , le nombre d'équipements pour 100 personnes (N_c) est calculé de la façon suivante :

$$N_c = 100 \times \left(\frac{\text{Nombre d'équipements}}{\text{Population communale}} \right)$$

Ces variables nous permettent de proposer une description plus détaillée de l'environnement spatial de l'entreprise.

Le taux de dépendance économique et le taux de chômage sont décrits dans le chapitre 1 et sont issus de l'atlas des zones d'emploi 2010. Ils représentent des indicateurs du climat économique dans la zone d'emploi de l'entreprise.

Pour mesurer la concurrence à laquelle les entreprises font face dans leur zone d'emploi, nous nous inspirons de Perraud et al. (2008). Nous calculons ainsi la densité du tissu économique en rapportant le nombre d'entreprises dans la zone d'emploi à sa population, puis en multipliant le résultat par 1000 afin d'éviter de très faibles valeurs et faciliter l'inter-

prétation.

$$\text{dnbre_ze} = 1000 \times \frac{\text{Nombre d'entreprises dans la zone d'emploi}}{\text{Population de la zone d'emploi}}$$

L'indice de [Maurel & Sédillot \(1999\)](#) est utilisé pour mesurer la concentration spatiale des secteurs. Nous avons interverti les secteurs et les zones géographiques afin que l'indice soit une mesure de la concentration sectorielle des zones d'emploi ([Houdebine, 1999](#)). En utilisant la nomenclature des activités agrégées en 21 postes, nous obtenons pour chaque zone d'emploi z :

$$\text{Mau_Sed} = \frac{\frac{\sum_s (\alpha_z^s)^2 - \sum_s (\lambda_s)^2}{1 - \sum_s (\lambda_s)^2} - \sum_{\text{ent} \in z} \varepsilon_{\text{ent}}^2}{1 - \sum_{\text{ent} \in z} \varepsilon_{\text{ent}}^2}$$

α_z^s représente la part de l'emploi du secteur s dans la zone d'emploi z ; λ_s correspond à la part du secteur s dans les effectifs nationaux; $\varepsilon_{\text{ent}}^2$ est la part de l'entreprise *ent* dans les effectifs du secteur auquel elle appartient. L'indice varie entre -1 et 1. Il vaut -1 lorsqu'il y a une répartition équitable des secteurs dans la zone d'emploi; 0 lorsque la répartition est aléatoire et 1 lorsqu'il y a une parfaite concentration d'un secteur dans la zone d'emploi. Les autres variables de contrôle proviennent de la base SINE ou sont construites à partir d'une variable préexistante dans la base. Il conviendrait toutefois de clarifier la définition de quelques unes d'entre elles.

La variable *filial* permet de prendre en compte le caractère filiale ou non d'une entreprise. D'après l'enquête SINE 2010, une entreprise est considérée comme filiale lorsqu'une ou plusieurs sociétés participent à plus de 50% de son capital. Les filiales arrivent donc sur le marché avec un soutien que les autres jeunes entreprises n'ont pas, ce qui est susceptible d'augmenter leur chance de survie. Il est donc important de prendre ce facteur en considération.

Les entreprises n'ont pas les mêmes chances de survie selon les secteurs dans lesquels elles opèrent ([Mata & Portugal, 1994](#); [Audretsch, 1995](#); [Baumöhl et al., 2020](#)). De plus, l'impact de l'environnement spatial (économies ou déséconomies d'agglomération par exemple) peut être hétérogène à travers les secteurs. [Ebert et al. \(2019\)](#) mettent cela en évidence en montrant que les entreprises non-technologiques et non innovantes sont les

seules qui bénéficient des externalités de localisation. Pour prendre en compte cet aspect, nous avons créé la variable indicatrice *sec_inov* qui permet distinguer les entreprises qui appartiennent au secteur innovant. La notion de secteur innovant est établie par l'enquête SINE, et regroupe les activités liées aux technologies de l'information et de la communication, aux produits pharmaceutiques, aux biotechnologies et aux nouveaux matériaux.

La variable *ambitieux* reflète l'objectif du créateur de l'entreprise. À la question de savoir quel est son objectif principal, l'entrepreneur a le choix entre trois réponses : «essentiellement assurer votre propre emploi»; «développer fortement votre entreprise en termes d'emploi» puis «développer fortement votre entreprise en termes d'investissement». Les deux dernières réponses sont celles pour lesquelles nous qualifions l'entrepreneur d'ambitieux.

TABLE 2.3.1 – Codes et libellés des variables

Codes	Libellés dans les tableaux de résultat	Libellés
agea	Age du créateur	Age du créateur en 2010
totald	Taille au démarrage	Total des emplois au démarrage
d_actsoc	Équipements en actions sociales	Nombre d'équipements en action social, pour 100 personnes
d_commerce	Équipements dans le commerce	Nombre d'équipements dans le commerce pour 100 personnes
d_ens2nddeg	Enseignement au 2nd degré	Nombre d'équipements dans l'enseignement au 2nd degré pour 100 personnes
d_ens1erdeg	Enseignement au 1er degré	Nombre d'équipements dans l'enseignement au 1er degré pour 100 personnes
d_ensupformautr	Enseignement supérieur et autres	Nombre d'équipements dans l'enseignement supérieur formation et services d'éducation pour 100 personnes
d_medparmed	Équipements médicaux et paramédicaux	Nombre d'équipements, dans les services médicaux et paramédicaux pour 100
d_sante	Équipements en santé	Nombre d'équipements dans les services de santé pour 100
d_servparticul	Services aux particuliers	Nombre d'équipements dans les services aux particuliers pour 100 personnes
d_sprtslrscult	Sport loisirs et culture	Nombre d'équipements dans le sport et la culture pour 100 personnes
d_trsprtsm	Transport et tourisme	Nombre d'équipements dans le transport et le tourisme pour 100 personnes
taux_dep_eco_09	Taux de dépendance économique	Taux de dépendance économique de la zone d'emploi
tx_chom_t1_03	Taux de chômage	Taux de chômage de la zone d'emploi
dnbre_ze	Nbre d'entreprises pour 1000	Nombre d'entreprises pour 1000 personnes dans la zone d'emploi
Mau_Sed	Indice Maurel-Sédillot	Indice Maurel-Sédillot
dum_creav	Expérience en création	Variante indicatrice égale à 1 lorsque l'entrepreneur avait créé ou repris une entreprise autre que celle actuelle, avant 2010
bac_plus	Bac et plus	Variante indicatrice égale à 1 lorsque l'entrepreneur a au moins le BAC
filial	Filiale	Variante indicatrice égale à 1 lorsque l'entreprise est une filiale
sec_inov	Secteur innovant	Variante indicatrice égale à 1 lorsque l'entreprise appartient au secteur innovant
inov_proced	Innovation de procédé	Variante indicatrice égale à 1 lorsque l'entreprise a réalisé une innovation de procédé
inov_mark	Innovation de marketing	Variante indicatrice égale à 1 lorsque l'entreprise a réalisé une innovation de marketing
inov_organ	Innovation organisation	Variante indicatrice égale à 1 lorsque l'entreprise a réalisé une innovation en matière d'organisation
inov_prod	Innovation de produit	Variante indicatrice égale à 1 lorsque l'entreprise a réalisé une innovation de produits ou services
français	Français	Variante indicatrice égale à 1 lorsque l'entreprise est créée par un français
ambitieux	Ambitieux	Variante indicatrice égale à 1 lorsque le créateur de l'entreprise a pour objectif principal de développer fortement l'entreprise en termes d'emploi ou d'investissement, et 0 lorsqu'il est essentiellement question pour le créateur d'assurer son propre emploi
homme	homme	Variante indicatrice égale à 1 lorsque l'entreprise est créée par un homme

2.3.4 Analyse descriptive

2.3.4.1 Statistiques descriptives

Les tableaux 2.3.2 et 2.3.3 présentent respectivement les statistiques descriptives des variables indicatrices pour l'intégralité de l'échantillon et pour les entreprises ayant fermé.

Dans le tableau 2.3.2, concernant l'entrepreneur, on note que 92,66% sont des citoyens français; 71,42% sont des hommes; 33,49% avaient déjà eu à créer ou reprendre une entreprise; 59,88% ont au moins le BAC et 37,34% ont pour objectif principal de développer fortement l'entreprise en termes d'emploi ou d'investissement. Au niveau des entreprises, nous avons 6,73% qui sont des filiales et 3,62% qui appartiennent au secteur innovant. En ce qui concerne l'innovation, nous avons 10,3% des entreprises qui ont réalisé une innovation de procédé; 13,98% ont réalisé une innovation de marketing; 17,12% ont réalisé une innovation en matière d'organisation et 38,32% ont réalisé une innovation de produit ou service.

Lorsqu'on ne considère que les 16 868 entreprises ayant fermé, le tableau 2.3.3 montre que 32,40% des entrepreneurs avaient déjà eu à créer ou reprendre une entreprise; 56,66% ont au moins le BAC; 70,86% sont des hommes; 91,65% sont des citoyens français et 34,52% ont pour objectif principal de développer fortement l'entreprise en termes d'emploi ou d'investissement. On note que 5,48% des entreprises sont des filiales et 3,61% des entreprises appartiennent au secteur innovant. Quant à l'innovation, nous avons 9,84% des entreprises qui ont réalisé une innovation de procédé; 14,10% ont réalisé une innovation de marketing; 16,23% ont réalisé une innovation en matière d'organisation. et 38,02% ont réalisé une innovation de produit ou service.

TABLE 2.3.2 – Description des variables qualitatives

	dum_creav	bac_plus	filial	sec_inov	inov_proced	inov_mark	inov_organ	inov_prod	français	ambitieux	homme
	pct	pct	pct	pct	Proportions	pct	pct	pct	pct	pct	pct
0	66.51	40.12	93.27	96.38	89.70	86.02	82.88	61.68	7.34	62.65	28.58
1	33.49	59.88	6.73	3.62	10.30	13.98	17.12	38.32	92.66	37.35	71.42
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Nombre d'observations : 45 268											

TABLE 2.3.3 – Entreprises fermées

	dum_creav	bac_plus	filial	sec_inov	inov_proced	inov_mark	inov_organ	inov_prod	francais	ambitieux	homme
	pct	pct	pct	pct	pct	pct	pct	pct	pct	pct	pct
0	67.60	43.34	94.52	96.39	90.16	85.90	83.77	61.98	8.35	65.48	29.14
1	32.40	56.66	5.48	3.61	9.84	14.10	16.23	38.02	91.65	34.52	70.86
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Nombre d'observations : 16 868											

TABLE 2.3.4 – Description des variables quantitatives

	Statistiques			
	moyenne	écart-type	min	max
Age du créateur	40.08	10.61	18.00	88.00
Taille au démarrage	1.93	4.59	1.00	370.00
Équipements en actions sociales	0.09	0.08	0.00	2.76
Équipements dans le commerce	0.74	0.81	0.00	16.67
Enseignement au 2nd degré	0.02	0.02	0.00	0.51
Enseignement au 1er degré	0.07	0.06	0.00	1.72
Enseignement supérieur et autres	0.02	0.03	0.00	0.91
Équipements médicaux et paramédicaux	0.68	0.42	0.00	5.84
Équipements en santé	0.09	0.06	0.00	1.77
Services aux particuliers	2.41	1.62	0.00	55.56
Sport loisirs et culture	0.27	0.28	0.00	10.99
Transport et tourisme	0.20	0.27	0.00	5.86
Taux de dépendance économique	90.98	10.70	70.51	130.18
Taux de chômage	9.95	5.82	4.00	46.60
Nbre d'entreprises pour 1000	1.02	0.58	0.16	3.28
Indice Maurel-Sédillot	0.00	0.01	-0.01	0.03
Nombre d'observations : 45 268				

Le tableau 2.3.4 présente les statistiques descriptives des variables quantitatives pour tout l'échantillon. Les résultats indiquent que les entrepreneurs ont en moyenne 40 ans au début de l'étude. La taille moyenne des entreprises au démarrage est de 2 employés. Les entreprises sont localisées dans des communes dotées en moyenne de moins d'un équipement pour 100 personnes, quelque soit le type d'équipement. Les seuls équipements qui dérogent à cette observation sont ceux dans les services aux particuliers où on dénombre en moyenne 2 pour 100 personnes dans les communes. Le taux de dépendance économique et le taux de chômage dans les zones d'emploi valent respectivement 90,08% et 9,95% en moyenne. Concernant la densité du tissu économique dans les zones d'emploi, nous avons en moyenne une entreprise pour 1000 personnes. L'indice de concentration spatiale Maud_Sed a une valeur moyenne de 0,00139, cela suggère une répartition approximativement aléatoire des secteurs d'activité dans les zones d'emploi.

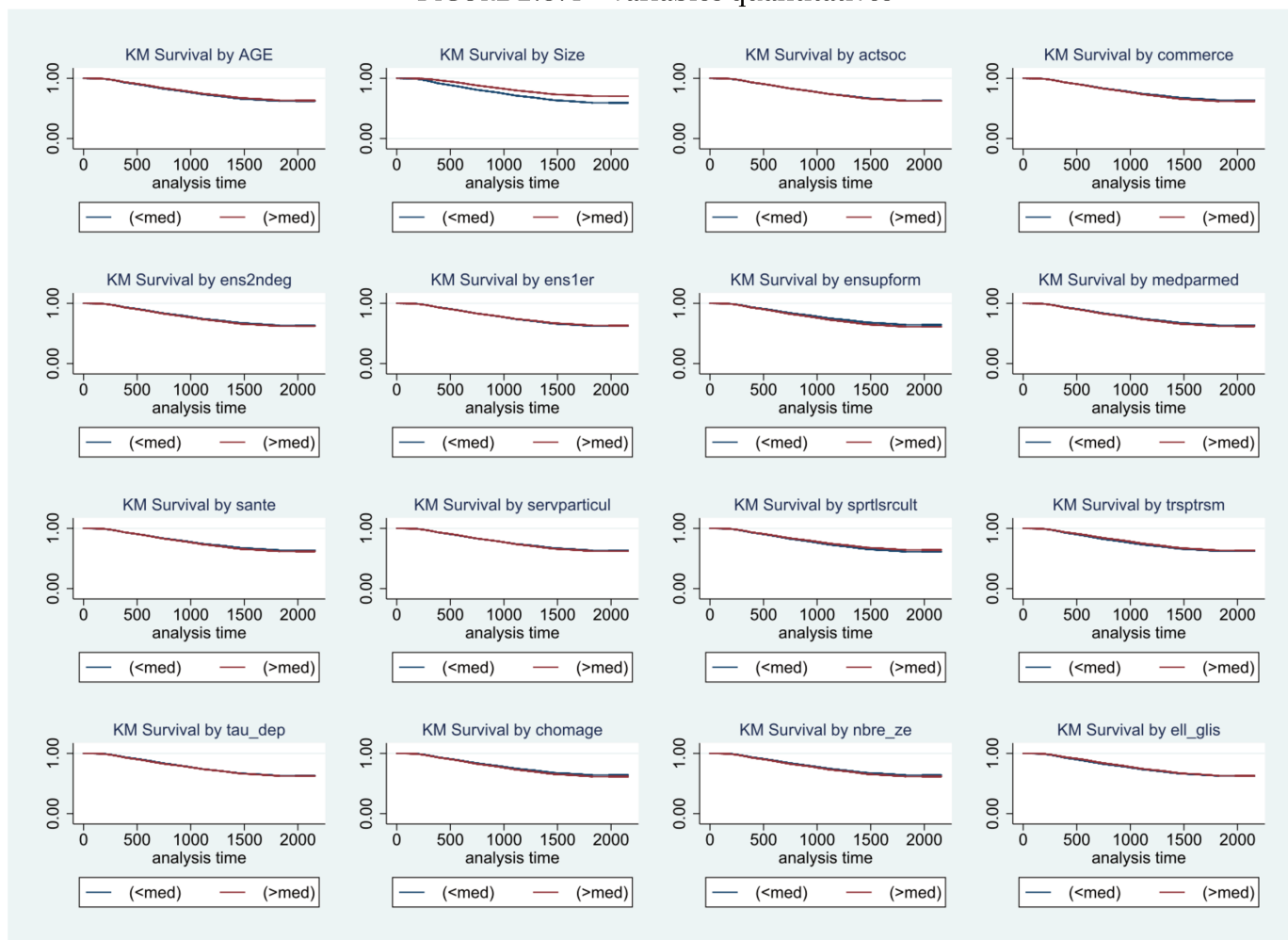
Le tableau 2.3.5 reprend les mêmes statistiques sur les entreprises ayant fermé et celles n'ayant pas fermé.

TABLE 2.3.5 – Description par catégories

	Censurées				Fermées			
	moyenne	écart-type	min	max	moyenne	écart-type	min	max
Age du créateur	40.21	10.44	18.00	88.00	39.88	10.88	18.00	82.00
Taille au démarrage	2.11	5.49	1.00	370.00	1.61	2.35	1.00	103.00
Équipements en actions sociales	0.09	0.08	0.00	2.61	0.09	0.08	0.00	2.76
Équipements dans le commerce	0.73	0.81	0.00	10.35	0.74	0.80	0.00	16.67
Enseignement au 2nd degré	0.02	0.02	0.00	0.51	0.02	0.02	0.00	0.51
Enseignement au 1er degré	0.08	0.06	0.00	1.72	0.07	0.05	0.00	1.27
Enseignement supérieur et autres	0.02	0.03	0.00	0.91	0.02	0.03	0.00	0.79
Équipements médicaux et paramédicaux	0.68	0.43	0.00	5.84	0.69	0.41	0.00	5.84
Équipements en santé	0.09	0.06	0.00	1.77	0.09	0.06	0.00	0.75
Services aux particuliers	2.40	1.62	0.00	15.14	2.42	1.63	0.00	55.56
Sport loisirs et culture	0.27	0.28	0.00	10.99	0.26	0.28	0.00	8.79
Transport et tourisme	0.20	0.28	0.00	5.86	0.20	0.26	0.00	5.56
Taux de dépendance économique	90.94	10.71	70.51	130.18	91.04	10.68	70.51	130.18
Taux de chômage	9.82	5.65	4.00	46.60	10.17	6.10	4.00	46.60
Nbre d'entreprises pour 1000	1.01	0.58	0.16	3.28	1.04	0.59	0.16	3.28
Indice Maurel-Sédillot	0.00	0.01	-0.01	0.03	0.00	0.01	-0.01	0.03
Nombre d'observations : 28 000 censurées et 16 868 fermées								

2.3.4.2 Analyse non paramétrique de survie de Kaplan Meier

FIGURE 2.3.4 – Variables quantitatives

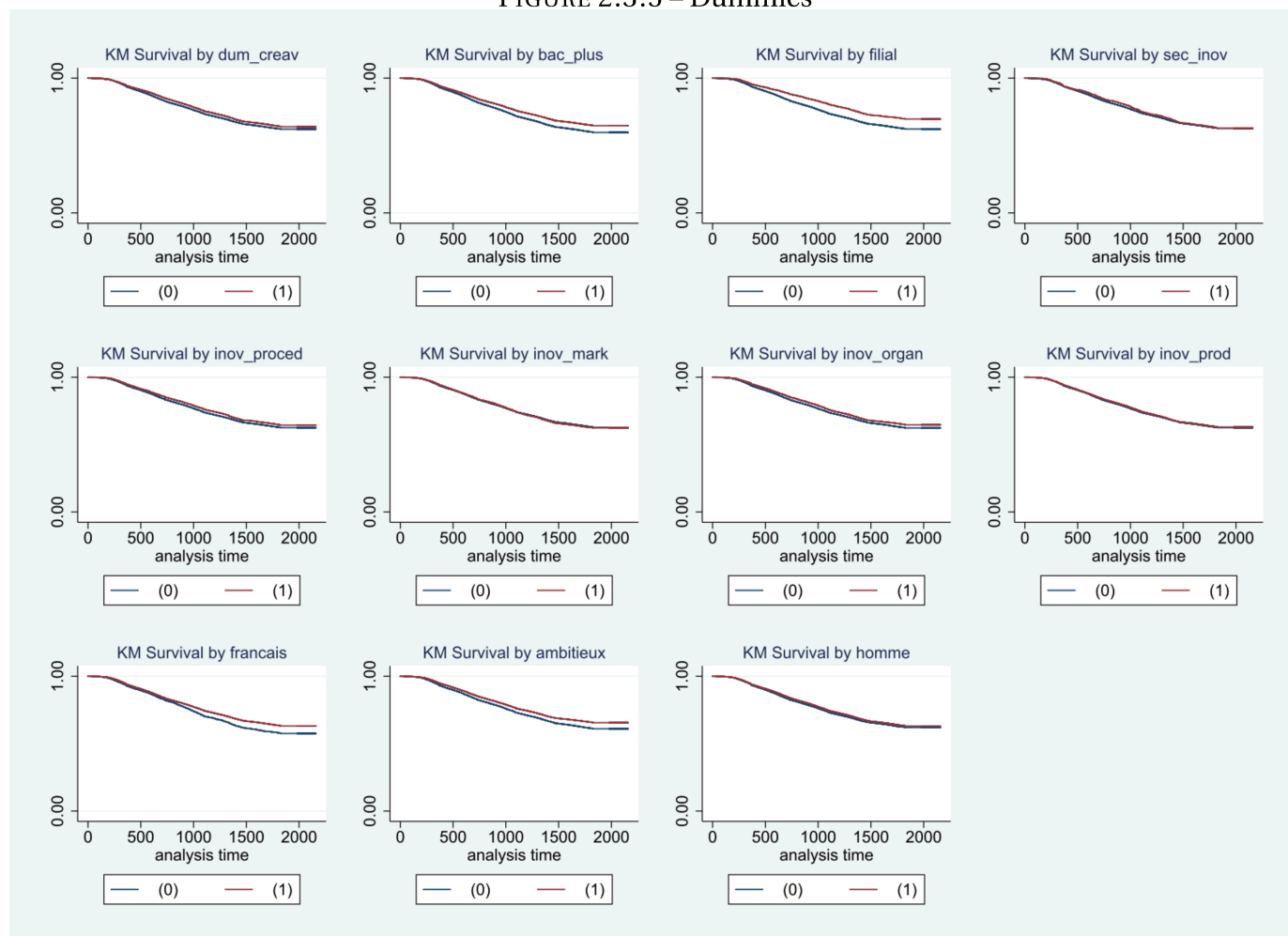


Dans cette section nous analysons graphiquement les fonctions de survie de Kaplan Meier. Les figures 2.3.4 et 2.3.5 présentent respectivement les courbes pour les variables quantitatives continues et les variables indicatrices. Au niveau des variables quantitatives continues, on distingue les valeurs à gauche de la médiane (courbes bleues) et celles à droite de la médiane (courbes rouges). Les tests log-rank sont disponibles dans l'annexe 2.A.

D'après les tests log-rank, les différences observées entre les fonctions de survie sont significatives pour toutes les variables sauf d_{actsoc} ; $d_{ens1erdeg}$; $d_{servparticul}$; $taux_dep_eco_09$; Mau_Sed ; sec_inov ; $inov_mark$ et $inov_prod$. Cependant, les fonctions de survie ne se démarquent ostensiblement les unes des autres que dans le cas de peu de

variables. C'est notamment le cas pour les variables comme *totd* (KM Survival by Size); *dum_creav*; *bac_plus*; *filial*; *inov_proced*; *inov_organ*; *français*; et *ambitieux*.

FIGURE 2.3.5 – Dummies



Les tendances constatées au niveau de ces variables paraissent assez intuitives et ne sont pas particulièrement en contraste avec les faits stylisés. D'après les courbes de survie, les entreprises ayant démarré avec plus d'un employé (médiane= 1) ont une durée de vie plus longue que les entreprises ayant démarré sans employés. De même, les entreprises ayant réalisé une innovation (de procédé ou en matière d'organisation) ont une plus longue durée de vie que celles n'en ayant pas réalisé. La taille au démarrage et l'innovation apparaissent donc sans surprise comme des facteurs favorables à la survie de l'entreprise.

Le parcours académique, l'expérience passée, et l'ambition du créateur de l'entreprise jouent également un rôle positif. Les graphiques nous montrent que les entreprises créées par des entrepreneurs ayant eu à créer ou reprendre une autre entreprise, vivent plus long-

temps que celles créées par des entrepreneurs sans expérience. Les entreprises créées par des entrepreneurs ayant au moins le BAC sont de même plus pérennes que celles créées par des entrepreneurs ayant moins que le BAC. Concernant l'ambition du créateur, les courbes de survie nous révèlent que les entreprises créées par des entrepreneurs désireux de développer fortement l'entreprise en termes d'emploi ou d'investissement, vivent plus longtemps que celles créées par des entrepreneurs qui souhaitent simplement assurer leur propre emploi.

Enfin, on note que les filiales et les entreprises créées par des citoyens français vivent plus longtemps que les entreprises indépendantes et celles créées par des étrangers.

2.3.5 Résultats d'estimations

Cette section s'articule autour de deux sous-sections dans lesquelles nous présentons et commentons nos résultats d'estimation. La première sous-section présente et discute les résultats de l'analyse de la survie lorsque la propension à choisir la localisation n'est pas prise en compte. La seconde section complète l'analyse en prenant le choix de localisation en considération.

2.3.5.1 Résultats d'analyse de la survie

Nous retenons dans notre analyse trois spécifications (modèle à hasards proportionnels Exponentiel, de Weibull et de Cox) et estimons trois modèles pour chacune des spécifications. Le premier modèle inclut les caractéristiques de l'entreprise, le second inclut les caractéristiques de l'environnement spatial de l'entreprise et le troisième modèle inclut les deux. Cette procédure d'estimation en plusieurs étapes nous permet de voir dans un premier temps l'effet isolé de chaque groupe de variables.

Ensuite, nous avons effectué un test de rapport de vraisemblance pour le modèle 3, en prenant le modèle 1 comme modèle contraint. Les résultats du test montrent pour toutes les spécifications que les caractéristiques de l'environnement spatial de l'entreprise sont globalement significatives. Aussi, remarque-t-on que les résultats d'estimation sont pratiquement identiques pour toutes les spécifications. Nous nous baserons donc sur le modèle de Cox pour la discussion.

TABLE 2.3.6 – Modèle de Cox

	Modèle 1	Modèle 2	Modèle 3
Age du créateur	-0.00233** (0.000771)		-0.00220** (0.000771)
Taille au démarrage	-0.0690*** (0.00514)		-0.0701*** (0.00516)
Expérience en création	-0.0211 (0.0172)		-0.0245 (0.0172)
Bac et plus	-0.163*** (0.0160)		-0.157*** (0.0161)
Filiale	-0.158*** (0.0343)		-0.164*** (0.0344)
Secteur innovant	0.0211 (0.0416)		0.485*** (0.0755)
Innovation de procédé	-0.0356 (0.0285)		-0.0336 (0.0285)
Innovation de marketing	0.0752** (0.0248)		0.0614* (0.0249)
Innovation organisation	-0.0366 (0.0236)		-0.0289 (0.0236)
Innovation de produit	0.00793 (0.0177)		0.00105 (0.0177)
Français	-0.154*** (0.0280)		-0.152*** (0.0282)
Ambitieux	-0.0844*** (0.0169)		-0.0881*** (0.0169)
homme	-0.0344* (0.0173)		-0.0245 (0.0174)
Équipements en actions sociales		0.00843 (0.116)	0.0310 (0.115)
Équipements dans le commerce		0.00122 (0.0196)	0.00963 (0.0195)
Enseignement au 2nd degré		-0.217 (0.425)	-0.256 (0.424)
Enseignement au 1er degré		-0.412** (0.159)	-0.462** (0.160)
Enseignement supérieur et autres		0.697* (0.303)	0.835** (0.303)
Équipements médicaux et paramédicaux		-0.0279 (0.0283)	-0.0124 (0.0283)
Équipements en santé		0.198 (0.180)	0.192 (0.179)
Services aux particuliers		0.000622 (0.0111)	-0.00138 (0.0111)
Sport loisirs et culture		-0.104** (0.0354)	-0.116** (0.0356)
Transport et tourisme		-0.0428 (0.0480)	-0.0543 (0.0481)
Taux de dépendance économique		0.00133 (0.000839)	0.00102 (0.000842)
Taux de chômage		0.00649*** (0.00155)	0.00569*** (0.00155)
Nbre d'entreprises pour 1000		0.0370* (0.0166)	0.0246 (0.0166)
Indice Maurel-Sédillot		-5.901*** (1.380)	-17.78*** (2.442)
<i>N</i>	45268	45268	45268
lrtest_df			14
lrtest_chi2			146.0
lrtest_p			4.55e-24

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Discussion

Dans ce paragraphe, nous commentons les résultats du tableau 2.3.6. Les ratios de risque sont disponibles dans l'annexe 2.B. Les résultats du modèle 3 montrent que la prise en compte des caractéristiques de l'environnement spatial n'a pas perturbé l'impact des variables spécifiques à l'entreprise. Inversement, le contrôle des caractéristiques de l'entreprise n'affecte guère les variables décrivant l'environnement spatial de l'entreprise. Nous discuterons donc des résultats du modèle 3.

D'après les résultats d'estimation, l'âge du créateur affecte positivement les chances de survie de l'entreprise. Lorsque le créateur est âgé d'une année supplémentaire, l'entreprise a, 0,2% de risque de moins de fermer. Cet effet positif est confirmé par les travaux de [Cressy \(1996\)](#) et de [Boyer & Blazy \(2014\)](#) qui ont respectivement travaillé sur des données anglaises et françaises. L'âge du créateur de l'entreprise peut être, toutes choses égales par ailleurs, considéré comme un indicateur d'expérience professionnelle ou académique, susceptible de bénéficier à l'entreprise. Les résultats de [Boyer & Blazy \(2014\)](#) ont justement mis en exergue l'importance de l'expérience professionnelle. Les auteurs ont trouvé que les entrepreneurs qui occupaient un emploi avant de créer leur entreprise ont presque 20% de risque de moins de déposer le bilan. Quant à l'expérience académique, elle s'est avérée significative dans nos estimations. Nos résultats montrent que lorsque l'entrepreneur a au moins le BAC, les risques de fermeture de l'entreprise sont réduits de 14,5%.

Le nombre d'employés au démarrage a un effet positif sur la survie des entreprises. Un employé additionnel baisse le risque instantané d'échec de 6,8%. Ce résultat récurrent dans la littérature ([Fotopoulos & Louri, 2000](#); [Staber, 2001](#); [Cefis & Marsili, 2005](#); [Huiban, 2011](#); [F. M. Neffke et al., 2011](#); [Mas-Verdú et al., 2015](#)), peut être expliqué par le fait que les entreprises ayant peu d'employés ne produisent pas assez pour bénéficier des économies d'échelle. Elles deviennent ainsi moins compétitive avec temps et quittent le marché plus tôt.

La probabilité pour les entreprises filiales de fermer est de 15,1% inférieure à celle des entreprises indépendantes. D'après l'enquête SINE, plus de la moitié du capital des entreprises filiales proviennent d'autres sociétés. Cet apport extérieur conséquent peut augmenter la productivité de l'entreprise et réduire ainsi les risques de mortalité.

Les entreprises du secteur innovant ont un taux instantané de mortalité de 62,4% supérieur autres entreprises. Le secteur innovant est composé d'activités liées aux technologies de l'information et de la communication, aux produits pharmaceutiques, aux biotechnologies et aux nouveaux matériaux. Le risque élevé de fermeture constaté dans ce secteur peut

être dû à la concurrence au sein de ce secteur et surtout au fait qu'il est indispensable d'innover pour s'en sortir. Cette situation peut s'avérer très difficile pour les jeunes entreprises.

L'introduction d'une innovation de marketing sur le marché, augmente la probabilité d'échec des jeunes entreprises de 6,3%. Ce résultat qui semble contre-intuitif, existe dans la littérature. En effet, [Hyytinen et al. \(2015\)](#), à partir d'un échantillon d'entreprises finlandaises, trouvent que les entreprises innovantes ont un taux de survie de six à sept points de pourcentage inférieur à celui des entreprises non-innovantes. L'activité d'innovation requiert un effort d'investissement considérable et a souvent très peu de chances de réussir. Certaines entreprises, en l'occurrence de jeunes entreprises, peuvent avoir du mal à se remettre de cet effort surtout lorsque les retours sur investissement ne sont pas conséquents.

Les entreprises créées par des citoyens français ont 14,1% de risque de moins de fermer comparativement aux entreprises créées par des étrangers. Ce résultat rejoint les travaux de [Cooper et al. \(1994\)](#) et de [Boyer & Blazy \(2014\)](#), et peut s'expliquer par plusieurs facteurs. Les citoyens français ont, toutes choses égales par ailleurs, une plus ample connaissance du marché, et cela peut réduire leurs risques d'échec. De plus, les entrepreneurs étrangers peuvent difficilement développer ou intégrer des réseaux d'entreprises, ils peuvent aussi avoir du mal à accéder à certaines ressources importantes comme le crédit bancaire, et cela peut entraver leur chances de succès.

Les entreprises créées par des entrepreneurs visant développer fortement leur projet en terme d'emploi ou d'investissement (variable *ambitieux*), ont plus de chances de survivre. Leur risque de mortalité est de 4,1% inférieur à celui des entreprises créées pour simplement assurer l'emploi de l'entrepreneur. Un entrepreneur très ambitieux a ainsi plus de chance de porter son projet au succès qu'un entrepreneur peu ambitieux. Par ailleurs, ce résultat pourrait révéler une réalité autre que la question d'ambition. [Andersson & Wadensjö \(2007\)](#) trouvent que les personnes inactives qui se mettent à l'entrepreneuriat, ont des revenus moins importants que celles qui étaient actives sur le marché de l'emploi avant la création de leur entreprise. [Muñoz-Bullón & Cueto \(2011\)](#) montrent que le risque d'échec est plus grand lorsque l'entrepreneuriat est précédé d'une longue période de chômage. Le risque de mortalité élevé au niveau des entreprises créées pour assurer l'emploi de l'entrepreneur, peut venir du fait que ces derniers entreprennent juste pour sortir d'une longue période de chômage.

En ce qui concerne les variables explicatives décrivant l'environnement spatial de l'entreprise, seulement cinq ont d'impacts significatifs. Il s'agit du nombre d'équipements dans l'enseignement au 1er degré; du nombre d'équipements dans l'enseignement supérieur la

formation et les services d'éducation; du nombre d'équipements dans le sport et la culture; du taux de chômage de la zone d'emploi; et de la concentration spatiale de secteurs dans une zone d'emploi.

Le nombre d'équipements dans l'enseignement au 1er degré; le nombre d'équipements dans le sport et la culture; et la concentration spatiale affectent positivement la survie des jeunes entreprises. Plus précisément, un équipement additionnel pour 100 personnes dans l'enseignement au 1er degré au niveau communal réduit le risque instantané de fermeture de 37%. La baisse de risque est de 10,9% lorsqu'il s'agit d'équipements dans le sport et la culture. Ces équipements destinés au bien-être des habitants des communes, peuvent susciter des flux migratoires vers ces communes, créant ainsi une demande potentielle supplémentaire pour les entreprises locales. Ces équipements peuvent également contribuer au confort de l'entrepreneur et lui permettre de s'investir davantage dans son projet. Quant à la concentration spatiales de secteurs, son effet positif sur la survie peut être compris comme la présence d'externalités MAR (Marshall, 1890; Arrow, 1962; Romer, 1986).

Le taux de chômage et le nombre d'équipements dans l'enseignement supérieur, la formation et les services d'éducation, accroissent les risques de mortalité des entreprises. Un taux de chômage élevé implique, toutes choses égales par ailleurs, un déséquilibre du marché de travail et une faible demande adressée aux entreprises. Brixy & Grotz (2007) et Huggins et al. (2017) ont également mis en évidence l'effet négatif du chômage sur la survie des entreprises.

L'impact négatif des équipements dans l'enseignement supérieur, formation et services d'éducation est quant à lui contre-intuitif, mais peut s'expliquer. Les équipements de l'enseignement supérieur, en particulier les universités, constituent des sources d'externalités de connaissance, généralement prisées par les entreprises (Audretsch et al., 2005). Cependant, les connaissances produites par ces institutions semblent n'intéresser que les entreprises opérant dans certains secteurs. En effet, Baptista et al. (2011) ont mis en évidence une baisse des créations d'entreprises des secteurs à faible usage de technologie, dans les régions dotées d'un nouvel établissement d'enseignement supérieur. Dans notre échantillon, seul 3,62% des entreprises opèrent dans des domaines liés aux technologies de l'information et de la communication, aux produits pharmaceutiques, aux biotechnologies et aux nouveaux matériaux. L'impact négatif des équipements dans l'enseignement supérieur, formation et services d'éducation, pourrait donc traduire le fait que les entreprises étudiées ne tirent aucun profit de ces équipements.

2.3.5.2 Choix de localisation

Les résultats de la section précédente montrent que la survie des jeunes entreprises dépend aussi, toutes choses égales par ailleurs, de l'environnement spatial. Cependant le choix de localisation des entreprises n'a pas été pris en compte. Dans cette sous-section, nous nous sommes basés sur la stratégie décrite dans la section 2.3.2 pour estimer la propension de chaque entreprise à choisir sa zone d'implantation. Le tableau ci-dessous présente les résultats d'estimation du modèle de prédiction de la propension.

Résultats estimés des paramètres					
Paramètre	DDL	Estimation	Écart type	Valeur du test t	Approx Pr > t
dnbre_ze	1	0.5708	0.006692	85.29	<.0001

Cette propension est ensuite incluse dans la régression de la section précédente. Les résultats d'estimation figurent dans le tableau 2.3.8 (tableau 2.3.9 pour les ratios de risque).

Discussion

Le tableau 2.3.8 présente les résultats du modèle précédemment interprété (Modèle 1) et ceux du modèle qui contrôle le choix de localisation des entreprises (Modèle 2). D'un point de vue du signe des coefficients, les résultats n'ont pas changé, les interprétations précédentes tiennent donc toujours. Toutefois, lorsqu'on considère l'ampleur des impacts, des différences peuvent être soulignées au niveau de certaines variables. Ces différences ont notamment plus de sens lorsqu'on raisonne en ratios de risque, et montrent de manière générale qu'une négligence du choix de localisation peut conduire à de légères surestimations de certains effets.

Ainsi, la réduction du risque de fermeture associée à un équipement additionnel pour 100 personnes dans l'enseignement au 1er degré passe 37% à 35,8% lorsque le choix de localisation est pris en compte. On passe de 10,9% de risque de moins à 11,4% de risque de moins pour les équipements dans le sport et la culture. Dans l'enseignement supérieur la formation et les services d'éducation, un équipement en plus pour 100 personnes augmentait le risque de fermeture des entreprises de 130,5%, mais lorsqu'on contrôle le choix de localisation de l'entreprise, la hausse du risque passe à 105,7%.

Un autre résultat qu'il convient de mettre en exergue est le coefficient négatif associé à la probabilité prédite. Cela signifie que les entreprises les plus enclines à choisir leur localisation actuelle, ont des risques de fermeture plus faibles. Ce résultat sous-entend clairement

qu'il est possible pour une entreprise de s'implanter dans une localisation qui ne lui est pas optimale en terme de profit espéré. En effet, la localisation optimale est susceptible d'impliquer des coûts au delà des capacités de certaines entreprises. Ces dernières vont donc se retrouver à devoir choisir d'autres localisations, et ce, parfois au péril de leur survie.

TABLE 2.3.8 – Résultats d'estimation : coefficients

	Modèle 1	Modèle 2
Age du créateur	-0.00220** (0.000771)	-0.00226** (0.000772)
Taille au démarrage	-0.0701*** (0.00516)	-0.0699*** (0.00516)
Expérience en création	-0.0245 (0.0172)	-0.0240 (0.0172)
Bac et plus	-0.157*** (0.0161)	-0.157*** (0.0161)
Filiale	-0.164*** (0.0344)	-0.163*** (0.0344)
Secteur innovant	0.485*** (0.0755)	0.485*** (0.0755)
Innovation de procédé	-0.0336 (0.0285)	-0.0325 (0.0285)
Innovation de marketing	0.0614* (0.0249)	0.0614* (0.0249)
Innovation organisation	-0.0289 (0.0236)	-0.0290 (0.0236)
Innovation de produit	0.00105 (0.0177)	0.000995 (0.0177)
Français	-0.152*** (0.0282)	-0.152*** (0.0282)
Ambitieux	-0.0881*** (0.0169)	-0.0879*** (0.0169)
homme	-0.0245 (0.0174)	-0.0242 (0.0174)
Équipements en actions sociales	0.0310 (0.115)	0.0282 (0.115)
Équipements dans le commerce	0.00963 (0.0195)	0.0104 (0.0195)
Enseignement au 2nd degré	-0.256 (0.424)	-0.224 (0.425)
Enseignement au 1er degré	-0.462** (0.160)	-0.443** (0.160)
Enseignement supérieur et autres	0.835** (0.303)	0.721* (0.306)
Équipements médicaux et paramédicaux	-0.0124 (0.0283)	-0.0122 (0.0283)
Équipements en santé	0.192 (0.179)	0.195 (0.179)
Services aux particuliers	-0.00138 (0.0111)	-0.00208 (0.0111)
Sport loisirs et culture	-0.116** (0.0356)	-0.121*** (0.0357)
Transport et tourisme	-0.0543 (0.0481)	-0.0386 (0.0483)
Taux de dépendance économique	0.00102 (0.000842)	0.000617 (0.000852)
Taux de chômage	0.00569*** (0.00155)	0.00717*** (0.00162)
Nbre d'entreprises pour 1000	0.0246 (0.0166)	0.251*** (0.0685)
Indice Maurel-Sédillot	-17.78*** (2.442)	-17.80*** (2.443)
p		-97.16*** (28.70)
N	45268	45268
lrtest_df	14	1
lrtest_chi2	146.0	11.88
lrtest_p	4.55e-24	0.000568

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 2.3.9 – Ratios de risque

	Modèle 1	Modèle 2
Age du créateur	0.998** (0.000770)	0.998** (0.000770)
Taille au démarrage	0.932*** (0.00481)	0.933*** (0.00481)
Expérience en création	0.976 (0.0168)	0.976 (0.0168)
Bac et plus	0.855*** (0.0137)	0.855*** (0.0137)
Filiale	0.849*** (0.0292)	0.850*** (0.0292)
Secteur innovant	1.624*** (0.123)	1.625*** (0.123)
Innovation de procédé	0.967 (0.0276)	0.968 (0.0276)
Innovation de marketing	1.063* (0.0265)	1.063* (0.0265)
Innovation organisation	0.972 (0.0229)	0.971 (0.0229)
Innovation de produit	1.001 (0.0177)	1.001 (0.0177)
Français	0.859*** (0.0242)	0.859*** (0.0242)
Ambitieux	0.916*** (0.0155)	0.916*** (0.0155)
homme	0.976 (0.0169)	0.976 (0.0170)
Équipements en actions sociales	1.031 (0.119)	1.029 (0.118)
Équipements dans le commerce	1.010 (0.0197)	1.010 (0.0197)
Enseignement au 2nd degré	0.774 (0.329)	0.799 (0.339)
Enseignement au 1er degré	0.630** (0.101)	0.642** (0.103)
Enseignement supérieur et autres	2.305** (0.698)	2.057* (0.629)
Équipements médicaux et paramédicaux	0.988 (0.0279)	0.988 (0.0280)
Équipements en santé	1.212 (0.217)	1.215 (0.217)
Services aux particuliers	0.999 (0.0111)	0.998 (0.0111)
Sport loisirs et culture	0.891** (0.0317)	0.886*** (0.0316)
Transport et tourisme	0.947 (0.0456)	0.962 (0.0465)
Taux de dépendance économique	1.001 (0.000843)	1.001 (0.000852)
Taux de chômage	1.006*** (0.00156)	1.007*** (0.00163)
Nbre d'entreprises pour 1000	1.025 (0.0170)	1.285*** (0.0880)
Indice Maurel-Sédillot	1.89e-08*** (4.61e-08)	1.86e-08*** (4.53e-08)
p		6.40e-43*** (1.84e-41)
N	45268	45268

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conclusion

Une fois sur le marché, les jeunes entreprises se trouvent confrontées à certaines difficultés spécifiques qui amenuisent leur perspective de réussite. Très souvent de petites tailles, elles sont difficilement rentables en raison d'un faible niveau de production. En général, moins de 50% d'entre elles arrivent à vivre plus de cinq ans (Cefis & Marsili, 2011; Colombelli et al., 2016). L'environnement spatial constitue l'un des facteurs déterminants de la survie des jeunes entreprises. Certaines caractéristiques de l'environnement spatial qui motivent le choix de localisation des entreprises, s'avèrent parfois néfastes à leur survie. Pourtant, d'après les différentes théories de localisation, l'entreprise est supposée choisir rationnellement sa localisation dans le but de maximiser son profit.

Dans ce chapitre nous avons étudié la survie de jeunes entreprises françaises en prenant en compte leur choix de localisation. Dans un premier temps, nous estimons plusieurs modèles de survie incluant les caractéristiques de l'entreprise et celles de son environnement spatial. Ils suggèrent d'un côté que la survie des jeunes entreprises est positivement affectée par : l'âge du créateur; le nombre d'employés au démarrage de l'entreprise; le niveau d'éducation du créateur; le fait que l'entreprise soit une filiale; la citoyenneté française du créateur; l'ambition du créateur; le nombre d'équipements dans l'enseignement au 1er degré dans la commune; le nombre d'équipements dans le sport et la culture dans la commune; et la concentration spatiale des secteurs dans la zone d'emploi. De l'autre côté, la survie des entreprises est négativement affectée par : le fait que l'entreprise opère dans le secteur innovant; le fait que l'entreprise réalise une innovation de marketing; le taux de chômage dans la zone d'emploi; et le nombre d'équipements dans l'enseignement supérieur la formation et les services d'éducation.

Dans un second temps, par l'estimation d'un logit conditionnel, nous récupérons la probabilité prédite que l'entreprise choisisse sa localisation actuelle. Cette probabilité qui est une mesure de la prédisposition de l'entreprise à occuper son emplacement observé, est incluse dans la régression. Les résultats montrent des variations au niveau des impacts de certaines variables comme les équipements dans l'enseignement au 1er degré; les équipements dans le sports et la culture; puis les équipements dans l'enseignement supérieur, la formation et les services d'éducation. Ces variations de plusieurs points en pourcentage dans les ratios de risques, montrent que le choix de localisation n'est pas à négliger. Nous notons également que la probabilité prédite impacte positivement la survie de l'entreprise. Autrement dit, plus il est probable pour une entreprise de choisir sa localisation actuelle,

moins il est risqué pour elle de fermer.

Cependant, lorsqu'on considère l'entreprise comme un organisme composé de plusieurs entités, la notion de localisation optimale peut être relative à chacune des entités. Une entreprise ayant plusieurs établissements n'étudiera pas de la même manière le choix de localisation des unités de production et celui du quartier général. Les facteurs géographiques qui influencent la performance des unités de production n'ont pas nécessairement le même effet sur le quartier général d'autant plus que les deux entités n'ont pas les mêmes missions. Le précédent chapitre et le présent, ayant respectivement travaillé sur des établissements et des jeunes entreprises majoritairement indépendantes (93%), n'offrent pas forcément une vision complète du rapport entre la survie de l'entreprise dans sa globalité et l'environnement spatial. Le chapitre suivant propose ainsi d'étudier l'influence de l'environnement spatial du quartier général de l'entreprise sur sa survie.

Annexes

2.A Tests log-rank

TABLE 2.A.1 – Age du créateur de l'entreprise

Catégories		Événements	
Code	Libellés	Observés	Espérés
0	Valeurs inférieure à la médiane	8613	8426.75
1	Valeurs supérieure à la médiane	8255	8441.25
Total		16868	16868.00
Chi2 (1)		8.23	
Pr>Chi2		0.0041	

Résultat : Chances de survie significativement différentes pour les deux catégories

TABLE 2.A.2 – Taille de l'entreprise au démarrage

Catégories		Événements	
Code	Libellés	Observés	Espérés
0	Valeurs inférieure à la médiane	12785	11392.57
1	Valeurs supérieure à la médiane	4083	5475.43
Total		16868	16868.00
Chi2 (1)		524.69	
Pr>Chi2		0.0000	

Résultat : Chances de survie significativement différentes pour les deux catégories

TABLE 2.A.3 – Équipements en actions sociales et dans le commerce

Catégories		Équipements en actions sociales		Équipements dans le commerce	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8355	8407.95	8228	8479.31
1	Valeurs supérieures à la médiane	8513	8460.05	8640	8388.69
Total		16868	16868.00	16868	16868.00
Chi2(1)		0.67		14.98	
Pr>Chi2		0.4148 ^a		0.0001 ^b	

^a Différences de survie non significatives entre les catégories

^b Différences de survie significatives à 1% entre les catégories

TABLE 2.A.4 – Équipements dans l'enseignement au premier et au second degré

Catégories		Enseignement au 1er degré		Enseignement au 2nd degré	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8537	8421.48	8294	8465.59
1	Valeurs supérieures à la médiane	8331	8446.52	8574	8402.41
Total		16868	16868.00	16868	16868.00
Chi2(1)		3.17		6.98	
Pr>Chi2		0.0752 ^a		0.0082 ^b	

^a Différences de survie significatives à 10% entre les catégories

^b Différences de survie significatives à 1% entre les catégories

TABLE 2.A.5 – Équipements médicaux et paramédicaux, enseignement supérieur et autres

Catégories		Enseignement supérieur et autres		Équipements médicaux et paramédicaux	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8092	8551.72	8268	8490.94
1	Valeurs supérieures à la médiane	8776	8316.28	8600	8377.06
Total		16868	16868.00	16868	16868.00
Chi2(1)		50.15		11.79	
Pr>Chi2		0.0000*		0.0006*	

* Différences de survie significatives entre les catégories

TABLE 2.A.6 – Équipements en Santé, Services aux particuliers

Catégories		Équipements en santé		Services aux particuliers	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8225	8474.52	8318	8450.33
1	Valeurs supérieures à la médiane	8643	8393.48	8550	8417.67
Total		16868	16868.00	16868	16868.00
Chi2(1)		14.77		4.15	
Pr>Chi2		0.0001****		0.0415*	

**** Différences de survie significatives à 1/10000 entre les catégories

* Différences de survie significatives à 5% entre les catégories

TABLE 2.A.7 – Équipements en Sport, loisirs, culture, Transport et Tourisme

Catégories		Sport loisirs et culture		Transport et tourisme	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8755	8342.27	8573	8364.09
1	Valeurs supérieures à la médiane	8113	8525.73	8295	8503.91
Total		16868	16868.00	16868	16868.00
Chi2(1)			40.42		10.35
Pr>Chi2			0.0000*		0.0013*

* Différences de survie significatives entre les catégories

TABLE 2.A.8 – Taux de dépendance économique et Chômage

Catégories		Taux de dépendance économique		Taux de chômage	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8412	8492.50	8083	8449.38
1	Valeurs supérieures à la médiane	8456	8375.50	8785	8418.62
Total		16868	16868.00	16868	16868.00
Chi2(1)		1.54		31.85	
Pr>Chi2		0.2150 ^a		0.0000 ^b	

^a Différences de survie non significatives entre les catégories

^b Différences de survie significatives entre les catégories

TABLE 2.A.9 – Nbre d'entreprises pour 1000 et Indice Maurel-Sédillot

Catégories		Nbre d'entreprises pour 1000		Indice Maurel-Sédillot	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	Valeurs inférieures à la médiane	8135	8455.56	8581	8514.76
1	Valeurs supérieures à la médiane	8733	8412.44	8287	8353.24
Total		16868	16868.00	16868	16868.00
Chi2(1)		24.38		1.04	
Pr>Chi2		0.0000 ^a		0.3076 ^b	

^a Différences de survie significatives entre les catégories

^b Différences de survie non significatives entre les catégories

TABLE 2.A.10 – Expérience en création d'entreprise, et niveau de diplôme

Catégories		Expérience en création		Bac et plus	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	-Pas d'expérience -Diplôme inférieur au Bac	11402	11137.53	7310	6612.88
1	-Expérience -Bac et plus	5466	5730.47	9558	10255.12
Total		16868	16868.00	16868	16868.00
Chi2(1)		18.49		120.94	
Pr>Chi2		0.0000*		0.0000*	

* Différences de survie significatives entre les catégories

TABLE 2.A.1.1 – Filiale et type de secteur

Catégories		Filiale		Secteur innovant	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	-Non Filiale -Secteur non innovant	15944	15666.71	16259	16249.78
1	-Filiale -Secteur innovant	924	1201.29	609	618.22
Total		16868	16868.00	16868	16868.00
Chi2(1)		68.95		0.14	
Pr>Chi2		0.0000 ^a		0.7056 ^b	

^a Différences de survie significatives entre les catégories

^b Différences de survie non significatives entre les catégories

TABLE 2.A.12 – Innovations de procédé et de marketing

Catégories		Innovation de procédé		Innovation de marketing	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	N'a pas innové	15208	15099.15	14489	14509.21
1	A innové	1660	1768.85	2379	2358.79
Total		16868	16868.00	16868	16868.00
Chi2(1)		7.49		0.20	
Pr>Chi2		0.0062 ^a		0.6536 ^b	

^a Différences de survie significatives entre les catégories

^b Différences de survie non significatives entre les catégories

TABLE 2.A.13 – Innovations en organisation et de produit

Catégories		Innovation organisation		Innovation de produit	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	N'a pas innové	14130	13927.04	10454	10380.28
1	A innové	2738	2940.96	6414	6487.72
Total		16868	16868.00	16868	16868.00
Chi2(1)		16.97		1.36	
Pr>Chi2		0.0000 ^a		0.2432 ^b	

^a Différences de survie significatives entre les catégories

^b Différences de survie non significatives entre les catégories

TABLE 2.A.14 – Citoyenneté et Ambition

Catégories		Citoyenneté		Ambition	
Code	Libellés	Événements observés	Événements espérés	Événements observés	Événements espérés
0	-Non Français -Non Ambitieux	1409	1197.31	11046	10436.23
1	-Français -Ambitieux	15459	15670.69	5822	6431.77
Total		16868	16868.00	16868	16868.00
Chi2(1)			40.31		93.48
Pr>Chi2			0.0000*		0.0000*

* Différences de survie significatives entre les catégories

TABLE 2.A.15 – Sexe du créateur

Sexe		Événements	
Code	Libellés	Observés	Espérés
0	Femme	4916	4775.59
1	Homme	11952	12092.41
Total		16868	16868.00
Chi2 (1)		5.76	
Pr>Chi2		0.0164	

Résultat : Chances de survie significativement différentes pour les deux catégories

2.B Ratios de risques

TABLE 2.B.1 – Modèle Exponentiel

	Modèle 1	Modèle 2	Modèle 3
$_t$			
Age du créateur	0.998** (0.000769)		0.998** (0.000770)
Taille au démarrage	0.932*** (0.00480)		0.931*** (0.00482)
Expérience en création	0.980 (0.0169)		0.977 (0.0168)
Bac et plus	0.846*** (0.0135)		0.852*** (0.0137)
Filiale	0.854*** (0.0293)		0.849*** (0.0292)
Secteur innovant	1.026 (0.0427)		1.637*** (0.123)
Innovation de procédé	0.965 (0.0275)		0.967 (0.0276)
Innovation de marketing	1.082** (0.0269)		1.067** (0.0265)
Innovation organisation	0.964 (0.0227)		0.972 (0.0229)
Innovation de produit	1.009 (0.0179)		1.002 (0.0178)
Français	0.853*** (0.0239)		0.855*** (0.0241)
Ambitieux	0.918*** (0.0155)		0.915*** (0.0155)
homme	0.967 (0.0167)		0.977 (0.0170)
Équipements en actions sociales		1.014 (0.118)	1.038 (0.120)
Équipements dans le commerce		1.000 (0.0196)	1.009 (0.0197)
Enseignement au 2nd degré		0.809 (0.344)	0.779 (0.331)
Enseignement au 1er degré		0.656** (0.104)	0.625** (0.0999)
Enseignement supérieur et autres		1.992* (0.603)	2.283** (0.689)
Équipements médicaux et paramédicaux		0.971 (0.0275)	0.987 (0.0279)
Équipements en santé		1.231 (0.221)	1.224 (0.219)
Services aux particuliers		1.001 (0.0111)	0.999 (0.0111)
Sport loisirs et culture		0.898** (0.0319)	0.888*** (0.0316)
Transport et tourisme		0.962 (0.0461)	0.950 (0.0456)
Taux de dépendance économique		1.001 (0.000840)	1.001 (0.000843)
Taux de chômage		1.006*** (0.00156)	1.005*** (0.00156)
Nbre d'entreprises pour 1000		1.041* (0.0173)	1.028 (0.0171)
Indice Maurel-Sédillot		0.00278*** (0.00384)	1.61e-08*** (3.94e-08)
N	45268	45268	45268

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 2.B.2 – Modèle de Weibull

	Modèle 1	Modèle 2	Modèle 3
<i>_t</i>			
Age du créateur	0.998** (0.000770)		0.998** (0.000770)
Taille au démarrage	0.931*** (0.00481)		0.930*** (0.00483)
Expérience en création	0.980 (0.0169)		0.976 (0.0168)
Bac et plus	0.844*** (0.0135)		0.849*** (0.0136)
Filiale	0.852*** (0.0293)		0.847*** (0.0292)
Secteur innovant	1.026 (0.0427)		1.650*** (0.124)
Innovation de procédé	0.964 (0.0275)		0.966 (0.0276)
Innovation de marketing	1.083** (0.0269)		1.068** (0.0266)
Innovation organisation	0.963 (0.0227)		0.971 (0.0229)
Innovation de produit	1.009 (0.0179)		1.002 (0.0178)
Français	0.850*** (0.0238)		0.853*** (0.0241)
Ambitieux	0.918*** (0.0155)		0.914*** (0.0155)
homme	0.966* (0.0167)		0.976 (0.0170)
Équipements en actions sociales		1.014 (0.118)	1.040 (0.120)
Équipements dans le commerce		1.000 (0.0197)	1.009 (0.0197)
Enseignement au 2nd degré		0.806 (0.342)	0.777 (0.329)
Enseignement au 1er degré		0.653** (0.104)	0.620** (0.0992)
Enseignement supérieur et autres		2.012* (0.609)	2.311** (0.698)
Équipements médicaux et paramédicaux		0.970 (0.0275)	0.987 (0.0279)
Équipements en santé		1.235 (0.222)	1.228 (0.219)
Services aux particuliers		1.001 (0.0111)	0.999 (0.0111)
Sport loisirs et culture		0.897** (0.0318)	0.886*** (0.0316)
Transport et tourisme		0.961 (0.0460)	0.950 (0.0456)
Taux de dépendance économique		1.001 (0.000840)	1.001 (0.000843)
Taux de chômage		1.006*** (0.00156)	1.006*** (0.00156)
Nbre d'entreprises pour 1000		1.041* (0.0173)	1.028 (0.0171)
Indice Maurel-Sédillot		0.00250*** (0.00346)	1.18e-08*** (2.88e-08)
<i>l</i>			
ln_p	1.106*** (0.00795)	1.102*** (0.00793)	1.108*** (0.00795)
<i>N</i>	45268	45268	45268

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 2.B.3 – Modèle de Cox

	Modèle 1	Modèle 2	Modèle 3
Age du créateur	0.998** (0.000769)		0.998** (0.000770)
Taille au démarrage	0.933*** (0.00479)		0.932*** (0.00481)
Expérience en création	0.979 (0.0169)		0.976 (0.0168)
Bac et plus	0.850*** (0.0136)		0.855*** (0.0137)
Filiale	0.853*** (0.0293)		0.849*** (0.0292)
Secteur innovant	1.021 (0.0425)		1.624*** (0.123)
Innovation de procédé	0.965 (0.0275)		0.967 (0.0276)
Innovation de marketing	1.078** (0.0268)		1.063* (0.0265)
Innovation organisation	0.964 (0.0227)		0.972 (0.0229)
Innovation de produit	1.008 (0.0178)		1.001 (0.0177)
Français	0.857*** (0.0240)		0.859*** (0.0242)
Ambitieux	0.919*** (0.0155)		0.916*** (0.0155)
homme	0.966* (0.0167)		0.976 (0.0169)
Équipements en actions sociales		1.008 (0.117)	1.031 (0.119)
Équipements dans le commerce		1.001 (0.0197)	1.010 (0.0197)
Enseignement au 2nd degré		0.805 (0.342)	0.774 (0.329)
Enseignement au 1er degré		0.662** (0.105)	0.630** (0.101)
Enseignement supérieur et autres		2.008* (0.609)	2.305** (0.698)
Équipements médicaux et paramédicaux		0.972 (0.0276)	0.988 (0.0279)
Équipements en santé		1.219 (0.219)	1.212 (0.217)
Services aux particuliers		1.001 (0.0111)	0.999 (0.0111)
Sport loisirs et culture		0.901** (0.0319)	0.891** (0.0317)
Transport et tourisme		0.958 (0.0460)	0.947 (0.0456)
Taux de dépendance économique		1.001 (0.000840)	1.001 (0.000843)
Taux de chômage		1.007*** (0.00156)	1.006*** (0.00156)
Nbre d'entreprises pour 1000		1.038* (0.0172)	1.025 (0.0170)
Indice Maurel-Sédillot		0.00274*** (0.00378)	1.89e-08*** (4.61e-08)
<i>N</i>	45268	45268	45268

Ratios de risques; Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.C Coefficients estimés

TABLE 2.C.1 – Modèle Exponentiel

	Modèle 1	Modèle 2	Modèle 3
$_t$			
Age du créateur	-0.00233** (0.000771)		-0.00221** (0.000772)
Taille au démarrage	-0.0703*** (0.00515)		-0.0714*** (0.00517)
Expérience en création	-0.0203 (0.0172)		-0.0236 (0.0172)
Bac et plus	-0.167*** (0.0160)		-0.161*** (0.0161)
Filiale	-0.158*** (0.0343)		-0.163*** (0.0344)
Secteur innovant	0.0254 (0.0416)		0.493*** (0.0755)
Innovation de procédé	-0.0357 (0.0285)		-0.0335 (0.0285)
Innovation de marketing	0.0786** (0.0248)		0.0646** (0.0249)
Innovation organisation	-0.0368 (0.0236)		-0.0288 (0.0236)
Innovation de produit	0.00881 (0.0177)		0.00202 (0.0177)
Français	-0.160*** (0.0280)		-0.157*** (0.0282)
Ambitieux	-0.0851*** (0.0169)		-0.0889*** (0.0169)
homme	-0.0336 (0.0173)		-0.0235 (0.0174)
Équipements en actions sociales		0.0135 (0.116)	0.0374 (0.115)
Équipements dans le commerce		0.000112 (0.0196)	0.00860 (0.0195)
Enseignement au 2nd degré		-0.212 (0.425)	-0.249 (0.424)
Enseignement au 1er degré		-0.421** (0.159)	-0.470** (0.160)
Enseignement supérieur et autres		0.689* (0.303)	0.825** (0.302)
Équipements médicaux et paramédicaux		-0.0298 (0.0283)	-0.0136 (0.0283)
Équipements en santé		0.208 (0.179)	0.202 (0.179)
Services aux particuliers		0.00105 (0.0111)	-0.00104 (0.0111)
Sport loisirs et culture		-0.107** (0.0355)	-0.119*** (0.0356)
Transport et tourisme		-0.0391 (0.0479)	-0.0508 (0.0480)
Taux de dépendance économique		0.00129 (0.000839)	0.000974 (0.000842)
Taux de chômage		0.00625*** (0.00155)	0.00542*** (0.00155)
Nbre d'entreprises pour 1000		0.0398* (0.0166)	0.0274 (0.0166)
Indice Maurel-Sédillot		-5.886*** (1.382)	-17.94*** (2.442)
$_cons$	-7.857*** (0.0439)	-8.541*** (0.0767)	-7.982*** (0.0875)
N	45268	45268	45268
$lrtest_df$			14
$lrtest_chi2$			147.5
$lrtest_p$			2.26e-24

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 2.C.2 – Modèle de Weibull

	Modèle 1	Modèle 2	Modèle 3
<hr/>			
_t			
Age du créateur	-0.00238** (0.000772)		-0.00226** (0.000772)
Taille au démarrage	-0.0715*** (0.00517)		-0.0726*** (0.00519)
Expérience en création	-0.0206 (0.0172)		-0.0240 (0.0172)
Bac et plus	-0.170*** (0.0160)		-0.164*** (0.0161)
Filiale	-0.160*** (0.0343)		-0.166*** (0.0344)
Secteur innovant	0.0259 (0.0416)		0.501*** (0.0754)
Innovation de procédé	-0.0365 (0.0285)		-0.0344 (0.0285)
Innovation de marketing	0.0801** (0.0248)		0.0659** (0.0249)
Innovation organisation	-0.0372 (0.0236)		-0.0291 (0.0236)
Innovation de produit	0.00897 (0.0177)		0.00203 (0.0177)
Français	-0.162*** (0.0280)		-0.159*** (0.0282)
Ambitieux	-0.0860*** (0.0169)		-0.0899*** (0.0169)
homme	-0.0342* (0.0173)		-0.0238 (0.0174)
Équipements en actions sociales		0.0141 (0.116)	0.0388 (0.115)
Équipements dans le commerce		0.0000602 (0.0197)	0.00876 (0.0195)
Enseignement au 2nd degré		-0.216 (0.425)	-0.253 (0.424)
Enseignement au 1er degré		-0.427** (0.159)	-0.478** (0.160)
Enseignement supérieur et autres		0.699* (0.303)	0.838** (0.302)
Équipements médicaux et paramédicaux		-0.0302 (0.0283)	-0.0136 (0.0283)
Équipements en santé		0.211 (0.179)	0.205 (0.179)
Services aux particuliers		0.00110 (0.0111)	-0.00108 (0.0111)
Sport loisirs et culture		-0.109** (0.0355)	-0.121*** (0.0357)
Transport et tourisme		-0.0397 (0.0479)	-0.0518 (0.0480)
Taux de dépendance économique		0.00130 (0.000839)	0.000987 (0.000842)
Taux de chômage		0.00634*** (0.00155)	0.00551*** (0.00155)
Nbre d'entreprises pour 1000		0.0403* (0.0166)	0.0278 (0.0166)
Indice Maurel-Sédillot		-5.990*** (1.383)	-18.26*** (2.440)
_cons	-8.645*** (0.0736)	-9.308*** (0.0972)	-8.781*** (0.106)
<hr/>			
/			
ln_p	0.101*** (0.00718)	0.0971*** (0.00720)	0.102*** (0.00718)
N	45268	45268	45268
lrtest_df			14
lrtest_chi2			152.4
lrtest_p			2.38e-25

Erreurs types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapitre 3

Environnement spatial du quartier général et survie de l'entreprise

Introduction

L'entreprise est une organisation économique régie par une architecture bien structurée. Les différentes entités qui composent l'entreprise, sont caractérisées par des missions bien précises et complémentaires qui concourent à la réalisation de ses objectifs de production et de distribution de biens et services. De manière générale, le fonctionnement d'une entreprise s'articule autour de deux axes principaux : le management et la production. Le management est assuré par le quartier général de l'entreprise et inclut les fonctions comme la gestion de la chaîne de valeur, la formulation et le déploiement de la stratégie de développement (Soparnot, 2012). Le quartier général, pour efficacement exercer ses fonctions managériales, doit parfois s'implanter dans un environnement géographique particulier. Même s'il n'est pas rare que l'unité de production d'une entreprise soit confondue avec son quartier général, beaucoup d'entreprises optent pour une séparation entre ces deux entités car les facteurs géographiques susceptibles de favoriser la réussite des unités de production n'ont pas forcément le même impact sur les quartiers généraux. Pour implanter une unité de production, l'entreprise peut par exemple privilégier une zone géographique rurale facilitant l'accès à divers intrants, et une main d'œuvre manufacturière peu couteuse. En revanche, pour son quartier général, l'entreprise aura tendance à opter pour une grande ville (A. W. Evans, 1973; Aarland et al., 2007; Zhang et al., 2018) où elle peut avoir un meilleur accès aux informations, une main d'œuvre qualifiée, et peut facilement externaliser certaines fonctions (Adler & Florida, 2020).

Dans les deux précédents chapitres, nous avons travaillé sur des établissements et des jeunes entreprises indépendantes, en observant ces derniers dans leur environnement géographique. Les résultats d'estimation ont montré dans les deux cas que les caractéristiques de la zone d'implantation impactent significativement la survie des établissements et des jeunes entreprises indépendantes. Néanmoins, les établissements d'une entreprise sont souvent des unités de production, et lorsque l'entreprise en dispose de plusieurs, la fermeture d'un n'est pas synonyme d'une cessation définitive de ses activités. Aussi, les jeunes entreprises sont-elles souvent de petites entreprises qui n'ont généralement pas les moyens

de séparer unité de production et quartier général.

Ce chapitre trois prolonge ainsi, dans un premier temps, les deux précédents, en mobilisant des sources de données qui renseignent sur l'état administratif (active ou fermée) de l'entreprise elle-même. La notion de fermeture ici équivaut donc à une cessation définitive de toute activité économique. Dans un second temps, vu que cette thèse s'intéresse globalement à l'impact de l'environnement spatial sur la survie de l'entreprise, le présent chapitre s'inscrit dans sa continuité en complétant les deux premiers par une étude des effets de la localisation du quartier général de l'entreprise sur sa survie, d'autant plus que le quartier général est une entité de premier plan.

Les études qui s'intéressent à l'environnement spatial des quartiers généraux d'entreprise s'arrêtent souvent à l'analyse des facteurs déterminant leur choix de localisation. Les quelques travaux qui mettent en relation la localisation du quartier général et la survie de l'entreprise, étudient généralement l'impact de l'éloignement du centre de commande sur la survie des établissements (Landier et al., 2009; Kalnins & Lafontaine, 2013). Dans ce chapitre, nous proposons une approche plus générale consistant à caractériser l'environnement spatial du quartier général et évaluer son influence sur la survie de l'entreprise. Pour ce faire, nous sommes partis de diverses sources de données et avons élaboré un échantillon de 761 108 entreprises observées entre 2005 et 2015. Nous avons ensuite utilisé une approche empirique reposant sur l'estimation de plusieurs modèles à hasards proportionnels de Cox.

Dans la suite de ce chapitre, nous commençons par présenter la littérature sur les quartiers généraux d'entreprises. Nous proposons ensuite une description de la localisation des quartiers généraux d'entreprises françaises. Nous poursuivons avec une présentation des données et variables utilisées, des modèles estimés, des résultats d'estimation, puis nous finissons par une conclusion.

3.1 Littérature

Les quartiers généraux d'entreprises en raison de leurs responsabilités managériales, ne s'implantent pas toujours dans le même environnement spatial que les filiales, les établissements ou les unités de production. Alors que l'entreprise peut privilégier la proximité à des sources d'intrants pour l'implantation de ses unités de production, la localisation du quartier général va beaucoup plus être déterminée par l'accès à de l'information. Les différents travaux empiriques sur la question aboutissent souvent à la conclusion selon

laquelle les quartiers généraux d'entreprises sont particulièrement enclins à se concentrer entre eux, et ce, dans les grandes métropoles.

Ono (2006) met justement l'emphase sur certaines raisons susceptibles de motiver la concentration spatiale des quartiers généraux d'entreprises. Dans son papier, l'auteur s'intéresse à la localisation des quartiers généraux d'entreprises dans les principales zones métropolitaines aux États-Unis. À l'issue de ses analyses, il explique que les besoins d'échange d'informations et de transactions directes peuvent conduire les entreprises à implanter leur quartier général dans une agglomération d'autres quartiers généraux. Cette conclusion rejoint celle de Davis & Henderson (2008) qui s'intéressent aux déterminants de l'agglomération des quartiers généraux d'entreprises. Leurs résultats montrent, entre autres, que les quartiers généraux sont plus souvent créés dans un voisinage de quartiers généraux. Les auteurs expliquent ce résultat en soulignant l'importance des économies de localisation. Cette tendance à s'agglomérer sera également mise en évidence par Strauss-Kahn & Vives (2009) qui travaillent sur des données américaines. Ces derniers trouvent que le nombre de quartiers généraux (issus d'une même industrie ou non) dans une zone, augmente la probabilité pour cette zone d'être choisie comme lieu d'implantation de quartiers généraux. Strauss-Kahn & Vives (2009) expliquent ce résultat en évoquant le fait que la proximité géographique avec d'autres quartiers généraux facilite les interactions et permet de profiter des externalités qui peuvent en découler, en l'occurrence les échanges informels d'informations. Il s'agit là d'un point d'importance cruciale surtout lorsque l'entreprise opère à l'étranger.

Lovely et al. (2005) étudient la concentration spatiale des quartiers généraux d'entreprises américaines exportatrices. En utilisant l'indice d'Ellison & Glaeser (1997) comme mesure de la concentration, les auteurs trouvent que les entreprises exportatrices ont tendance à se concentrer dans l'espace à un degré d'autant plus élevé que le marché étranger comporte des barrières commerciales. Les activités d'exportation requièrent parfois une connaissance spécifique du marché étranger et cette connaissance peut être difficile d'accès lorsque le pays n'est pas très ouvert au commerce international. Ainsi, les entreprises exportatrices optent pour une concentration spatiale de leurs quartiers généraux afin d'optimiser leurs chances d'accéder à certaines informations spécifiques relatives aux marchés étrangers où elles opèrent. Il est également important de noter que cette concentration de quartiers généraux d'entreprises a souvent lieu dans les grandes métropoles. Celles-ci doivent leur attractivité à la diversité (en termes d'activités) qui les caractérise, et aux possibilités de sous-traitance qu'elles offrent. Cet aspect a été mis en évidence dans plusieurs travaux empiriques.

En effet, les résultats de [Strauss-Kahn & Vives \(2009\)](#) ont montré que les quartiers généraux ont une forte prédilection pour les zones métropolitaines, en particulier celles dotées de bonnes infrastructures aéroportuaires et d'offres conséquentes de services aux entreprises. De même, [Davis & Henderson \(2008\)](#) ont trouvé que lorsque le nombre de fournisseurs de services intermédiaires divers augmentent de 10% dans une région, l'implantation de quartiers généraux dans la région augmente de 3,6%. [Aarland et al. \(2007\)](#) concluent que la perspective de sous-traiter certains services emmène les entreprises localisées dans les petites villes à davantage implanter leur quartier général dans les grandes villes. Cette préférence pour les grandes villes sera également mise en exergue par [A. W. Evans \(1973\)](#) et [Zhang et al. \(2018\)](#) qui travaillent respectivement sur des données anglaises et chinoises.

Mais dans un contexte post-industriel où le capitalisme cognitif occupe désormais une place importante, les entreprises n'implantent pas seulement leurs quartiers généraux sur la base d'aménités géographiques. [Adler & Florida \(2020\)](#) partent de l'hypothèse que le capital humain serait désormais le premier facteur de choix de localisation des quartiers généraux des grandes entreprises. Pour tester leur hypothèse, ils suivent sur la période 1966-2017 les 500 premières entreprises américaines en termes de chiffre d'affaire. Les estimations de plusieurs modèles de comptages suggèrent que le capital humain (part des adultes avec un baccalauréat +3 ou plus) est toutes choses égales par ailleurs un prédicteur fortement significatif de la localisation des quartiers généraux des grandes entreprises. Ces derniers (quartiers généraux) privilégient de plus en plus les régions richement dotées de capital humain.

L'un des constats qui émerge de ces études est que les facteurs de localisation les plus valorisés par les quartiers généraux, semblent majoritairement caractériser les grandes aires urbaines. Pourtant, implanter le quartier général dans une grande ville peut parfois signifier l'éloigner des différents établissements de l'entreprise. Cet éloignement entre quartier général et établissements, est susceptible d'engendrer des difficultés de gestion pouvant conduire à des conséquences dommageables pour l'entreprise. À ce propos, plusieurs études ont montré que les établissements localisés à proximité de leur quartier général ont de meilleures performances.

[Audia & Rider \(2010\)](#) travaillent sur l'industrie manufacturière américaine de chaussures. Ils montrent que les économies d'agglomérations ne profitent qu'aux établissements localisés dans le même voisinage que leur quartier général. Ces derniers ont un risque de mortalité plus faible que les établissements éloignés de leur quartier général. [Audia & Rider \(2010\)](#) expliquent ce résultat en se référant à des travaux ([Corporation et al., 1946](#); [Friedland](#)

& Palmer, 1984; Stern & Aldrich, 1980) qui montrent que, contrairement aux organisations éloignées de leur siège social, les organisations localisées dans le même voisinage que leur siège social s'engagent plus dans les interactions avec leur communauté locale. Ces interactions peuvent être de l'ordre de l'engagement politique ou de la participation à la vie sociale de la communauté locale. Cela garantit à ces organisations un soutien des acteurs locaux, et leur permet de jouir de meilleures perspectives de survie. Ce résultat de Audia & Rider (2010) apparaissait déjà dans les travaux de Landier et al. (2009) qui ont aussi travaillé sur des données américaines.

Landier et al. (2009) trouvent que les établissements proches de leur quartier général sont moins susceptibles de fermer. D'après les auteurs, ces établissements proches de leur siège social sont mieux gérés car les managers ont une meilleure connaissance de l'environnement de l'établissement et prennent en conséquence des décisions plus adéquates. Lorsque l'établissement est trop éloigné du centre de commande, le décideur est déconnecté de certaines réalités de l'établissement et est susceptible de prendre des décisions inappropriées, préjudiciables à l'établissement. Kalnins & Lafontaine (2013) ont montré que les établissements éloignés de leur quartier général ont des niveaux de recette plus faibles.

La distance entre le centre de commande et les établissements, peut également engendrer des problèmes de discipline au niveau des employés. Ces derniers, en absence de supervision suffisante, peuvent être tentés de travailler moins. Dans le but de responsabiliser leurs établissements, certaines entreprises ont recours au franchisage. Cette pratique du franchisage est beaucoup plus fréquente chez les entreprises dont les établissements sont éparpillés. Lafontaine (1992) a étudié le comportement de 548 franchiseurs américains opérant dans divers domaines d'activité en 1986. L'un des principaux résultats est que la dispersion géographique (proportion de points de vente étrangers dans chaque chaîne et le nombre d'États dans lesquels l'entreprise a des établissements) des établissements d'une entreprise augmente ses chances d'avoir recours au franchisage. En outre, les établissements éloignés sont les plus susceptibles d'être franchisés (Brickley & Dark, 1987).

Tous ces travaux nous montrent d'un côté comment les entreprises choisissent la localisation de leur quartier général, et de l'autre côté les conséquences d'un éloignement du quartier général des unités de production. Nous proposons dans ce chapitre de compléter cette littérature en étudiant l'impact de l'environnement spatial du quartier général de l'entreprise sur sa survie. Mais avant, nous menons une analyse descriptive de la localisation des quartiers généraux d'entreprises françaises.

3.2 Les quartiers généraux d'entreprises françaises

Dans cette section, qui précède notre analyse de la survie, nous proposons une description de la localisation des quartiers généraux d'entreprises françaises. Cette analyse préliminaire nous permettra d'une part, d'avoir une idée des facteurs motivant le choix du lieu d'implantation des quartiers généraux d'entreprises en France, et d'autre part de voir si nos observations sur le territoire français s'inscrivent dans les tendances mises en exergue dans des études ayant porté sur d'autres échantillons. Les éventuelles similitudes avec des travaux antérieurs peuvent nous aider à mieux comprendre et mieux justifier nos résultats.

3.2.1 Données et échantillons

Les données d'entreprises mobilisées dans cette section proviennent de la base SIRENE. L'échantillon est constitué de 2 909 grandes entreprises françaises qui opèrent dans les secteurs de l'industrie (extractive et manufacturière) de la construction, du commerce et réparation, de l'hébergement et de la restauration. Ces entreprises sont créées entre 1990 et 2005, et ont toutes leur quartier général en France métropolitaine. Nous avons particulièrement ciblé les entreprises créées entre 1990 et 2005, car celles étudiées dans la partie économétrique (prochaine section) sont en moyenne âgées de près de 13 ans en 2005, les plus jeunes étant créées cette même année. La construction de l'échantillon est effectuée sur le site <https://www.sirene.fr> où plusieurs options nous ont permis de directement choisir les catégories d'entreprises qui nous intéressaient. Nous avons ensuite ajouté à cet échantillon, des données sur les départements français provenant de l'Observatoire des Territoires. L'Observatoire des Territoires est créé en 2004 et a pour vocation d'analyser et de diffuser des données liées aux dynamiques et disparités des territoires. Les données proposées sont disponibles sur le site internet : <https://www.observatoire-des-territoires.gouv.fr>.

La combinaison de ces deux sources de données nous a permis d'avoir une base de 96 départements français au sein desquels sont répartis les 2909 quartiers généraux d'entreprises dont nous avons également les coordonnées géographiques.

3.2.2 Variables

Le tableau 3.2.1 ci-dessous, présente les différentes variables considérées. Elles décrivent toutes la situation socio-économique et infrastructurelle des départements. Excepté

le nombre de quartiers généraux, les variables (et leur description) proviennent toutes de l'Observatoire des Territoires. L'unité d'observation est le département.

La variable *Part des 25-34 ans titulaires d'un diplôme de l'enseignement supérieur*, représente la proportion des personnes âgées de 25 à 34 ans, titulaires d'un diplôme universitaire de 1er cycle, BTS, DUT, diplôme des professions sociales ou de santé, d'infirmier(ère), ou d'un diplôme universitaire de 2ème ou 3ème cycle (y compris médecine, pharmacie, dentaire), diplôme d'ingénieur, d'une grande école, doctorat. Cette variable permet d'avoir pour chaque département, une idée du niveau de capital humain, qui est un déterminant important du choix de localisation des quartiers généraux d'entreprises.

TABLE 3.2.1 – Variables

Variables	Libellés
Part des 25-34 ans titulaires d'un diplôme de l'enseignement supérieur en 2018	Part des 25-34 ans diplômée du sup
Longueur des routes en service en 2010	Longueur des routes
Densité de population en 2018	Densité de population
Taux d'activité des 15-64 ans en 2018	Taux d'activité des 15-64 ans
Part des "cadres et professions intellectuelles supérieures" dans la population en 2018	Part des cadres et intellectuels
Médiane du revenu disponible par Unité de Consommation en 2018	Revenu median
Nombre de quartiers généraux d'entreprises dans le département	Nombre de QG

La variable *Longueur des routes en service*, mesure en kilomètre les autoroutes, les routes nationales, départementales et communale. Seules celles ouvertes à la circulation sont prises en compte. Cet indicateur pourrait permettre aux quartiers généraux d'entreprises d'avoir une idée des potentielles difficultés liées aux échanges et à la mobilité au sein des départements.

La *Densité de population* nous donne le nombre d'habitant par kilomètre carré. Elle permet de distinguer les zones urbaines (forte densité) des zones rurales (faible densité). La variable *Taux d'activité des 15-64 ans* est le nombre total d'actifs (en emploi ou au chômage), rapporté à la population en âge de travailler (15-64 ans). La *Part des "cadres et professions intellectuelles supérieures"* est le pourcentage (sur l'ensemble de la population) d'individus rattachés à cette catégorie socio-professionnelle. Cette variable, avec le *Revenu médian*, fournissent des informations sur le niveau de vie des habitant du département.

La variable d'intérêt de cette analyse descriptive est le *Nombre de quartiers généraux*

d'entreprises. Nous l'avons calculée en comptant pour chaque département, le nombre de quartiers généraux. Nous mettons ainsi cette variable en relation avec les autres, afin de comprendre les facteurs susceptibles de faire d'un département, une zone prisée par les quartiers généraux d'entreprises.

3.2.3 Analyse

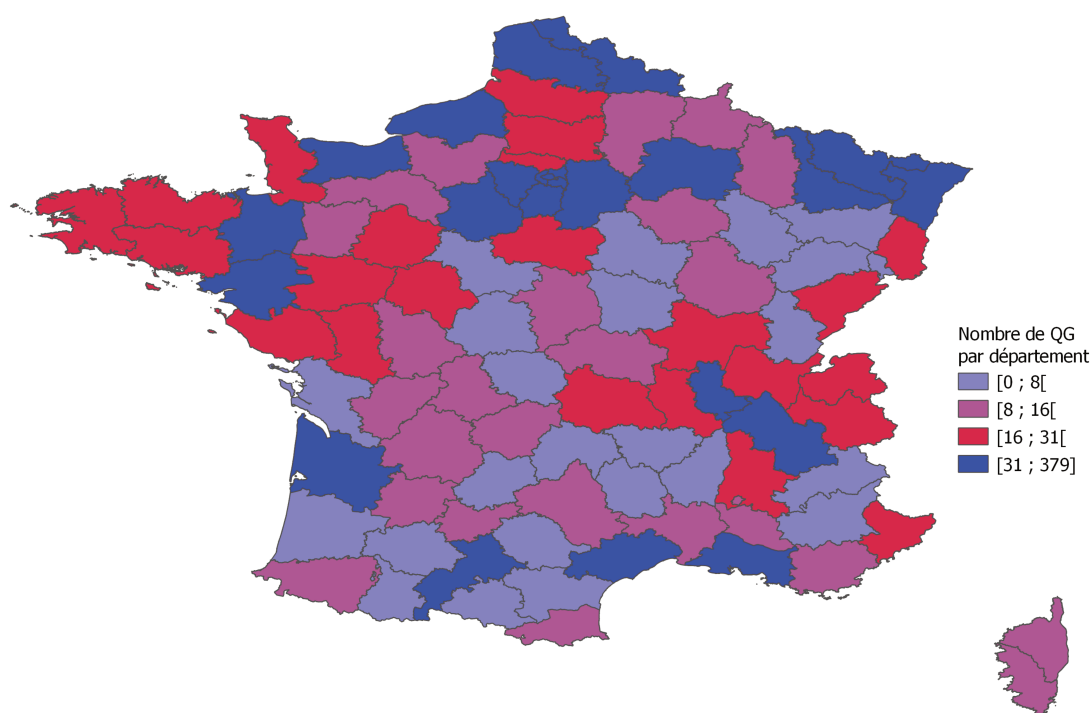
Le tableau 3.2.2 contient les statistiques descriptives des variables. Les différentes valeurs font état d'une forte dispersion, ce qui suggère l'idée certaine hétérogénéité entre les départements.

TABLE 3.2.2 – Statistiques descriptives

	Moyenne	Écart-type	Minimum	Maximum
Part des 25-34 ans diplômée du sup	39.14	6.99	29.10	72.10
Longueur des routes	10938.75	4278.06	1301.00	20662.00
Densité de population	565.85	2437.22	14.81	20641.38
Taux d'activité des 15-64 ans	74.18	1.91	69.60	80.30
Part des cadres et intellectuels	7.43	4.10	4.00	29.70
Revenu median	21395.42	1672.54	17740.00	28270.00
Nombre de QG	30.30	52.04	0.00	379.00
Nombre d'observations : 96 départements				

Dans le but de mieux comprendre la localisation des quartiers généraux, nous avons procédé à une analyse cartographique. Dans un premier temps, nous avons produit une carte choroplèthe (carte 3.2.1) montrant le nombre de quartiers généraux dans chaque département.

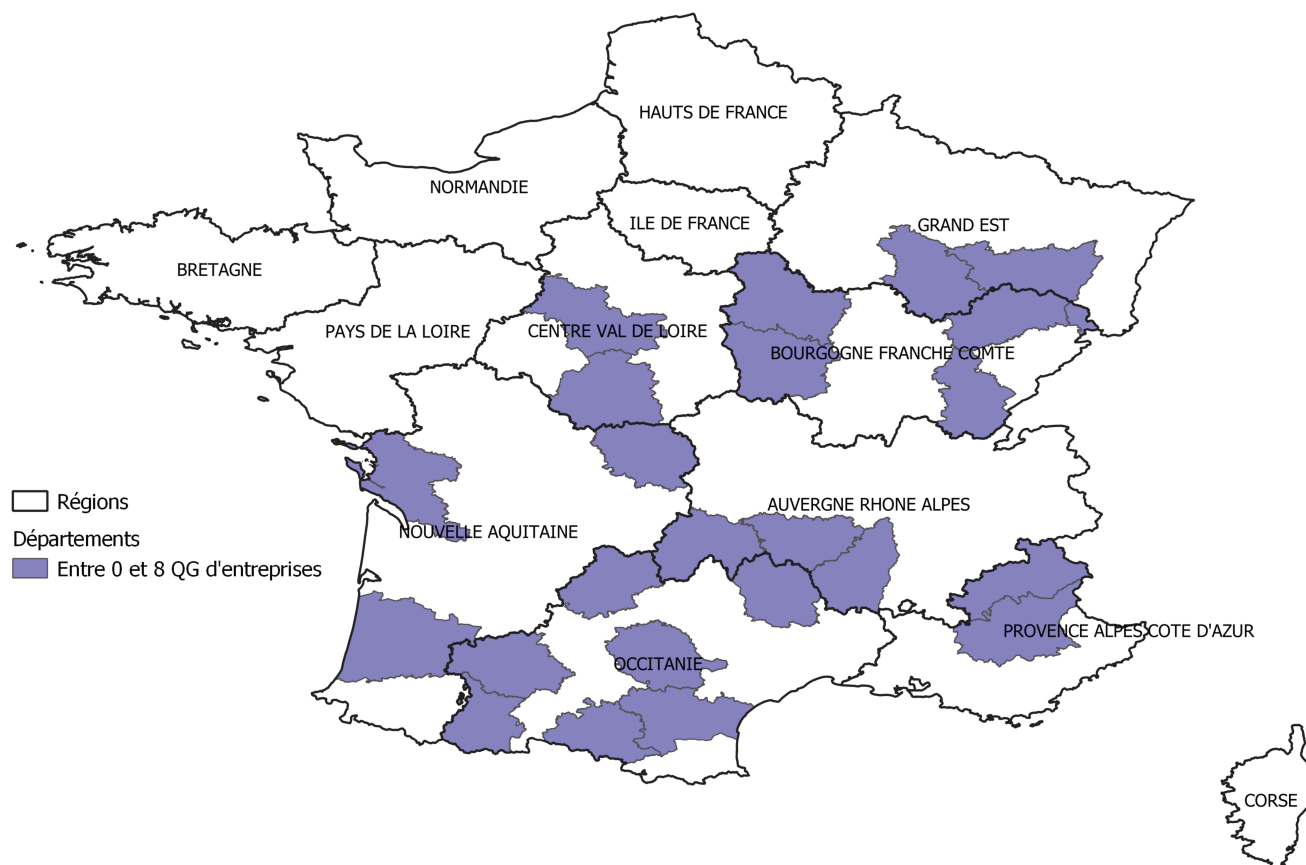
FIGURE 3.2.1 – Choroplèthe du nombre de QG



Nous notons, à partir de la carte 3.2.1, une hétérogénéité très prononcée du nombre de quartiers généraux d'entreprises dans les départements. La valeur médiane du nombre de quartier général par département est de 16, ce qui est presque deux fois moins que la moyenne qui vaut 30. Nous sommes donc en présence d'une distribution fortement asymétrique où un petit groupe de départements comptent un nombre très élevé de quartier généraux. La carte 3.2.1 nous montre également que les départements français à gauche du premier quartile de la distribution du nombre quartiers généraux, abritent entre 0 et 8 quartiers généraux alors que ceux situés à droite du troisième quartile en dénombrent entre 31 et 379.

Pour mieux comprendre ces disparités, en termes d'attractivité, entre départements, nous avons essayé de caractériser d'un côté les départements les plus prisés, et de l'autre côté, les moins prisés.

FIGURE 3.2.2 – Les départements les moins attractifs



La carte 3.2.2 montre les 24 départements en deçà du premier quartile de la distribution du nombre de quartiers généraux. Ces départements comptent en moyenne 4 quartiers généraux d'entreprises et sont dans les régions comme le Centre Val de Loire, la Bourgogne Franche-Comté, le Grand Est, la Nouvelle Aquitaine, l'Auvergne Rhone Alpes, l'Occitanie et la Provence-Alpes-Côte d'Azur. Le tableau ci-dessous présente les statistiques descriptives des différentes variables caractérisant les départements concernés.

TABLE 3.2.3 – Caractéristiques des départements

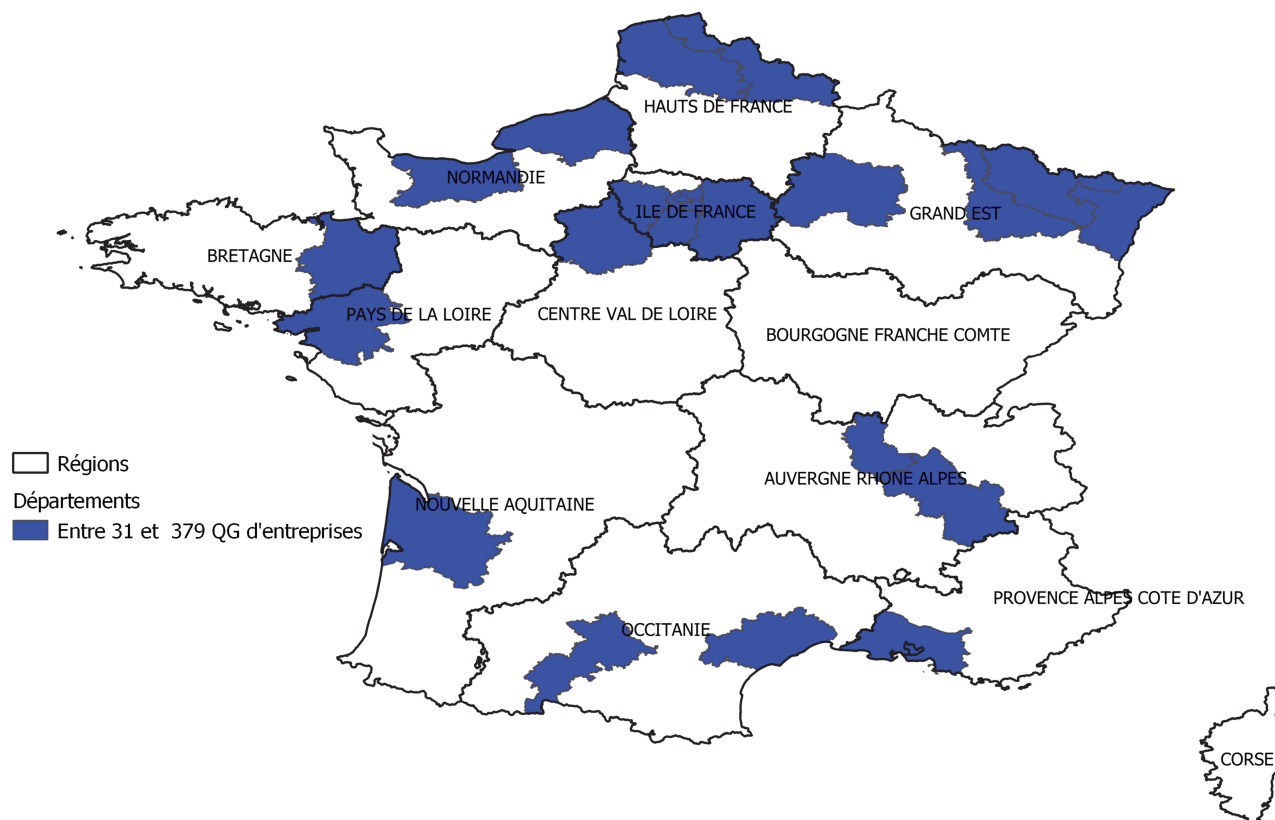
Variabes	Part des 25-34 ans diplômée du sup	Longueur des routes	Densité de population	Taux d'activité des 15-64 ans	Part des cadres et intellectuels	Revenu median	Nombre de QG
Moyenne	35,7917	9528,5	50,3233	74,15	5,19583	20616,7	3,83333

La valeur moyenne des variables au niveau de ces départements est inférieure à la va-

leur moyenne des variables lorsqu'on considère l'ensemble des départements. La part des 25-34 ans titulaires d'un diplôme de l'enseignement supérieur est en moyenne inférieure de plus de trois points de pourcentage comparativement à l'ensemble des départements. Les routes en services sont en moyenne longues de moins de 1400 kilomètres lorsqu'on compare à la moyenne de l'ensemble des départements. La densité de la population est 11 fois inférieure à sa moyenne sur l'ensemble des départements. Le taux d'activité des 15-64 ans et la part des intellectuels sont respectivement inférieurs de moins d'un point et de plus de deux points de pourcentage par rapport à la totalité des départements. Quant au revenu médian, il vaut 700 euros de moins dans ces départements. Le principal constat que l'on pourrait tirer de cette observation, c'est que les départements les moins choisis comme localisation de quartiers généraux, semblent limités en termes d'opportunités pour les entreprises. Le niveau du capital humain, les infrastructures routières, et le niveau de vie des habitants (demande potentielle), sont moins avantageux dans ces départements.

Intéressons-nous à présent aux départements les plus prisés par les quartiers généraux. La carte 3.2.3 montre les 24 départements situés au delà du troisième quartile de la distribution des quartiers généraux.

FIGURE 3.2.3 – Les départements les plus prisés



Alors que l'ensemble des départements comptent en moyenne une trentaine de quartiers généraux d'entreprises, on en dénombre en moyenne 85 dans ces 25% supérieurs. Les régions concernées sont les Hauts de France, la Bretagne, les Pays de la Loire, la Normandie, l'Île-de-France, le Grand Est, La Nouvelle Aquitaine, l'Auvergne Rhone Alpes, l'Occitanie et la Provence Alpes Côte d'Azur. Le tableau ci-dessous présente les valeurs moyennes des différentes variables au niveau de ces départements.

TABLE 3.2.4 – Caractéristiques des départements

Variabes	Part des 25-34 ans diplômée du sup	Longueur des routes	Densité de population	Taux d'activité des 15-64 ans	Part des cadres et intellectuels	Revenu median	Nombre de QG
Moyenne	45,9667	11159	1992,75	74,2375	11,7333	22519,6	85,25

Nous constatons que ces valeurs moyennes sont toutes supérieures à celles obtenues

sur l'ensemble des départements. L'on peut donc déduire, toutes choses égales par ailleurs, que les départements les plus prisés par les quartiers généraux, se distinguent des autres par leur niveau élevé de capital humain, une forte demande potentielle (niveau de vie des habitants), et de bonnes infrastructures routières. Ces départements sont également de grandes aires urbaines, caractérisés par une densité de population de plus de trois fois supérieure à la moyenne. Ono (2006) fait une observation similaire quant aux entreprises américaines. En effet, les statistiques descriptives de l'auteur montrent que les villes comme Chicago, New York, Washington D.C., Dallas et Philadelphie, arrivent en tête du classement des aires métropolitaines en termes de nombre de quartiers généraux. Cette réalité n'est donc pas spécifique aux régions françaises.

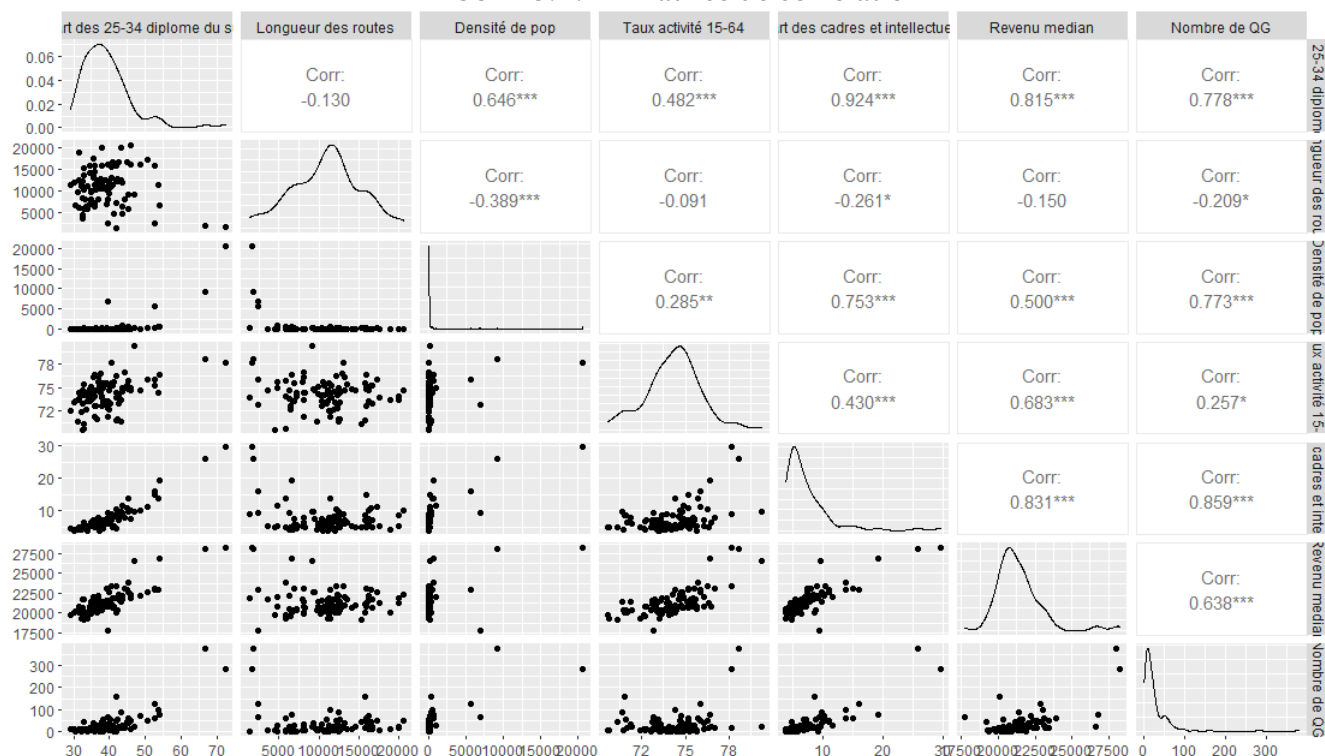
Nous retenons de cette analyse que les variables considérées, font potentiellement partie des facteurs susceptibles de motiver le choix d'un département comme lieu d'implantation du quartier général d'une entreprise. Une analyse du lien statistique entre le nombre de quartiers généraux et les différentes variables pourrait nous apporter plus de précision. A cette fin, la figure 3.2.4 présente la matrice de corrélation entre les variables. La diagonale principale affiche les distributions de chaque variable, en dessous nous avons les nuages de points et au-dessus nous avons les coefficients de corrélation de Pearson. Les étoiles indiquent le degré de significativité¹.

D'après les informations de la dernière colonne de la matrice, le nombre de quartiers généraux d'entreprises est significativement corrélé à toutes les autres variables. Toutes les corrélations sont positives à l'exception de celle avec la longueur des routes, qui est également la moins significative. D'une manière générale, ces résultats corroborent les observations précédentes selon lesquelles les départements les plus dotés comptent davantage de quartiers généraux.

Le coefficient le plus élevé vaut 0.859 et correspond à la corrélation entre le nombre de quartiers généraux et, la part des cadres et professions intellectuelles supérieures dans la population. Ce résultat, bien que n'impliquant pas une causalité, se rapproche dans une certaine mesure de l'étude de Adler & Florida (2020) qui place le capital humain au centre des facteurs déterminant la localisation des quartiers généraux d'entreprises.

1. * * * significatif à 1‰; ** significatif à 1%; * significatif à 5%

FIGURE 3.2.4 – Matrice de corrélation



Cette section nous a permis d'avoir une idée des facteurs déterminant le choix des entreprises pour la localisation de leur quartier général. Nous avons également pu remarquer que la tendance observée en France, ressortait déjà dans des études antérieures sur d'autres pays.

Par ailleurs, dans un monde de plus en plus changeant d'un point de vue technologique, le paysage spatial n'est plus une donnée fixe. L'environnement spatial du quartier général d'une entreprise peut imprévisiblement subir des transformations profondes sous l'influence de plusieurs forces endogènes. Le climat politico-institutionnel, le développement technologique, et les caractéristiques des agents économiques; peuvent, au sein d'une localisation, influencer le niveau de concurrence, modifier les frontières et la conception de la distance (Essletzbichler & Rigby, 2010). Les dotations en infrastructures peuvent également évoluer au point où des régions, au départ peu convoitées, se transforment et deviennent des pôles attractifs aux dépens des zones préalablement prisées. Les quartiers généraux d'entreprises doivent donc fréquemment ajuster leur stratégie pour faire face aux divers challenges associés à un environnement inconstant.

Cependant, cette adaptabilité n'est pas systématique. En effet, certains quartiers généraux réussiront à s'adapter, d'autres changeront de localisation (en particulier, les plus

grands en termes de recettes (Strauss-Kahn & Vives, 2009)), et une partie échouera dans sa mission et conduira l'entreprise à la faillite.

Dans la section qui suit, nous étudions les chances de survie d'une entreprise en fonction des caractéristiques de la zone d'implantation de son quartier général.

3.3 Localisation du quartier général et la survie de l'entreprise

Cette section propose une analyse économétrique de la sensibilité du risque de fermeture des entreprises, aux conditions de l'environnement spatial des quartiers généraux. Elle s'articule autour de deux sous-sections qui présentent respectivement les données et variables utilisées, puis les modèles économétriques estimés.

3.3.1 Données et variables

Plusieurs sources de données ont servi à la construction de l'échantillon des 761 108 entreprises étudiées dans ce chapitre. Nous avons profité des informations fournies par le Fichier de Comptabilité Unifié dans SUSE² (FICUS), le Fichier Approché des Résultats d'ESANE³ (FARE), et le Répertoire des Entreprises et des Établissements (REE). Les données SIRENE (Système Informatique pour le Répertoire des Entreprises et leurs Établissements) nous ont également été utiles. Les variables relatives à l'environnement spatial du quartier général de l'entreprise, décrivent la population locale, et informent sur la dotation en infrastructures des communes. Il s'agit des mêmes variables que nous avons précédemment détaillées dans le chapitre 1.

Les entreprises observées opèrent dans le secteur marchand non agricole et sont suivies sur la période 2005-2015 grâce au Répertoire des Entreprises et des Établissements (REE). L'échantillon obtenu a ensuite été joint aux bases FICUS et FARE afin d'avoir plus d'informations sur les entreprises. Le suivi des entreprises, et les fusions des différentes bases sont possibles grâce au numéro unique d'identification des entreprises (SIREN).

Créées en 1994, les bases FICUS sont constituées d'entreprises imposées aux bénéficiaires

2. Système Unifié de Statistiques d'Entreprises : Base de données individuelles et statistiques sur les entreprises.

3. Élaboration des Statistiques Annuelles d'Entreprises : Successeur du dispositif SUSE depuis 2009.

industriels commerciaux et aux bénéfiques non commerciaux, hors micro-entreprises. Les exploitations agricoles soumises aux bénéfiques industriels commerciaux et les entreprises imposées aux bénéfiques agricoles ne sont pas prises en considérations. Jusqu'en 2007, les bases FICUS concernaient les unités légales ayant déposé une liasse fiscale ou d'autres types de déclarations. En 2008, FICUS est remplacé par FARE qui au même titre que son prédécesseur, fournit d'amples informations sur la comptabilité des entreprises. Cependant, au niveau des bases FARE, le champ de la statistique d'entreprise est défini ex-ante, à partir du répertoire (SIRENE, puis SIRUS⁴), en fonction du caractère marchand des unités, leur caractère productif, leur activité principale, leur taille et leur catégorie juridique. La raison principale de ce changement est la qualité statistique. En effet, l'INSEE estime que la statistique d'entreprise serait mieux maîtrisée et de meilleure qualité si l'on s'appuyait sur les répertoires, plutôt que sur les différents champs de déclarations fiscales. Ainsi de 2005 à 2007 nous avons utilisé les bases FICUS et de 2008 à 2015 nous avons utilisé les bases FARE. Des tables de correspondances nous ont permis de retrouver les variables dont les noms ont changé entre les deux fichiers.

Dans ce chapitre, nous avons décidé de travailler sur les unités légales. Cela permet d'éviter les ambiguïtés lors de la fusion des bases FICUS et FARE avec les autres bases (SIRENE, REE) car ces dernières contiennent exclusivement des unités légales. Selon la Loi de Modernisation de l'Économie (LME) de 2008, l'entreprise se définit comme «la plus petite combinaison d'unités légales constituant une unité organisationnelle de production et jouissant d'une certaine autonomie de décision». Notons que l'unité légale «est une entité juridique de droit public ou privé. Cette entité juridique peut être : une personne morale, dont l'existence est reconnue par la loi indépendamment des personnes ou des institutions qui la possèdent ou qui en sont membres ; une personne physique, qui, en tant qu'indépendant, peut exercer une activité économique» (INSEE). L'entreprise est confondue à l'unité légale lorsque cette dernière s'inscrit dans la définition de la LME. En revanche, pour les entreprises constituées de plusieurs unités légales, l'INSEE identifie celle.s pertinente.s pour l'analyse économique et reconstitue leurs comptes consolidés. Cette opération s'appelle proflage des groupes français et a débuté en 2000.

Le tableau 3.3.1 présente les codes et libellés des variables utilisées. Conformément à la littérature sur la survie des entreprises, et comme cela a été le cas dans les autres chapitres, trois catégories de variables ont été prises en compte. Nous avons dans un premier temps des variables spécifiques à l'entreprise, ensuite des variables spécifiques au secteur d'activité et enfin des variables décrivant l'environnement spatial du quartier général de

4. Système d'immatriculation au répertoire des unités statistiques

l'entreprise qui sont nos variables d'intérêt. Certaines des variables sont mesurées au début de l'étude (2005) tandis que d'autres sont mesurées chaque année (section 3.3.2).

Les variables : effectif de l'entreprise ; chiffre d'affaires total en milliers d'euros ; bénéfice attribué ou perte transférée en milliers d'euros ; perte supportée ou bénéfice transféré en milliers d'euros ; impôt sur les bénéfices en milliers d'euros ; et valeur ajoutée hors taxe en milliers d'euros ; sont celles pour lesquelles nous avons une variabilité sur toute la période 2005-2015. En ce qui concerne les autres, elles sont observées à l'instant t_0 (2005) en raison du nombre élevé d'informations manquantes pour les autres années.

L'intégralité des variables liées à la comptabilité de l'entreprise proviennent directement des bases FICUS-FARE et sont prises telles quelles. Nous proposons, dans le paragraphe suivant, de détailler un peu plus certaines d'entre elles. Toutes les précisions apportées ici proviennent des documents descriptifs des données.

Le bénéfice attribué ou la perte transférée (*benefat*) vient en augmentation du bénéfice d'exploitation, tandis que la perte supportée ou bénéfice transféré (*pertsup*) vient en diminution du bénéfice d'exploitation. Le total des charges financières (*tcharfi*) comprend les intérêts versés par l'entreprise à ses prêteurs, les escomptes accordés aux clients, les pertes sur créances, les pertes de changes et les pertes liées à la cession de valeurs immobilières de placement. Les emprunts et dettes assimilées (*empdett*) sont constitués d'emprunts obligataires convertibles⁵, d'emprunts et dettes auprès des établissements de crédits, d'emprunts et dettes financières divers⁶, et d'autres emprunts obligataires⁷. Les investissements corporels et incorporels avec apports (*invavap*) sont la somme des acquisitions en immobilisations corporelles et incorporelles, dont on soustrait les virements de poste à poste.

Le tableau 3.3.2 présente pour l'intégralité de l'échantillon, les statistiques descriptives des variables sur toute la période de l'étude. Le tableau 3.3.3 quant à lui reprend les mêmes statistiques en distinguant les entreprises censurées et celles ayant fermé.

5. Emprunts matérialisés par des obligations susceptibles d'être transformées en actions de la société émettrice sur demande de l'obligataire.

6. Fonds dus aux salariés au titre de la participation, avances remboursables consenties par l'État et emprunts participatifs.

7. Emprunts obligataires ordinaires, remboursables selon un plan d'amortissement rigide, pouvant être éventuellement assortis de clauses de participation.

TABLE 3.3.1 – Codes et libellés des variables

Codes	Libellés dans tableaux de résultat	Libellés
eff	Effectif	Effectif de l'entreprise
catotal	Chiffre d'affaires	Chiffre d'affaires total en milliers d'euros
benefat	Bénéfice attribué ou perte transférée	Bénéfice attribué ou perte transférée en milliers d'euros
pertsup	Perte supportée ou bénéfice transféré	Perte supportée ou bénéfice transféré en milliers d'euros
impoben	Impôt sur les bénéfices	Impôt sur les bénéfices en milliers d'euros
vaht	Valeur ajoutée hors taxe	Valeur ajoutée hors taxe en milliers d'euros
taux_dep_eco_09	Dépendance économique	Taux de dépendance économique de la zone d'emploi
tx_chom_t1_03	Chômage	Taux de chômage de la zone d'emploi
dpop16	Densité de population	Densité de la population dans la commune
actsoc_p_km2	Équipements en action sociale	Nombre d'équipements en action sociale, pour 100 personnes par km2, par commune
cmmrce_p_km2	Équipements dans le commerce	Nombre d'équipements dans le commerce pour 100 personnes par km2, par commune
ens1erdeg_p_km2	Équipements dans l'enseignement au 1er degré	Nombre d'équipements dans l'enseignement au 1er degré pour 100 personnes par km2, par commune
ens2ndeg_p_km2	Équipements dans l'enseignement au 2nd degré	Nombre d'équipements dans l'enseignement au 2nd degré pour 100 personnes par km2, par commune
ensupformautr_p_km2	Équipements dans l'enseignement supérieur et autres	Nombre d'équipements dans l'enseignement supérieur formation et services d'éducation, pour 100 personnes par km2, par commune
medparmed_p_km2	Équipements dans les services médicaux et paramédicaux	Nombre d'équipements dans les services médicaux et paramédicaux, pour 100 personnes par km2, par commune
sante_p_km2	Équipements dans les services de santé	Nombre d'équipements dans les services de santé, pour 100 personnes par km2, par commune
servparticul_p_km2	Équipements dans les services aux particuliers	Nombre d'équipements dans les services aux particuliers, pour 100 personnes par km2, par commune
sprtlsrucult_p_km2	Équipements dans le sport et la culture	Nombre d'équipements dans le sport et la culture, pour 100 personnes par km2, par commune
trsptrsm_p_km2	Équipements dans le transport et le tourisme	Nombre d'équipements dans le transport et le tourisme pour 100 personnes par km2, par commune
nbet_a	Nombre d'établissements actifs	Nombre d'établissements actifs en 2005
tcharfi	Total des charges financières	Total des charges financières en 2005 en milliers d'euros
empdett	Emprunts et dettes assimilées	Emprunts et dettes assimilées en 2005 en milliers d'euros
invavap	Investissement corporel et incorporel	Investissement corporel et incorporel, avec apports en 2005 en milliers d'euros
age_dbt	Âge de l'entreprise en 2005	Âge de l'entreprise en 2005
sect_be	Industrie manufacturière	Variante prenant la valeur 1 si l'entreprise opère dans l'industrie manufacturière
sect_fz	Secteur de la construction	Variante prenant la valeur 1 si l'entreprise opère dans le secteur de la construction
sect_gi	Secteur commerce, transport, hébergement et restauration	Variante prenant la valeur 1 si l'entreprise opère dans les commerces de gros et de détail, transport, hébergement et restauration
sect_mn	Secteur activités spécialisées, scientifiques et techniques, services administratifs et soutien	Variante prenant la valeur 1 si l'entreprise opère dans les Activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien

TABLE 3.3.2 – Statistiques descriptives 2005-2015 pour tout l'échantillon

	Moyenne	Écart type	min	max
Effectif de l'entreprise	13.86	343.06	0.00	183 196.00
Chiffre d'affaires	3525.28	104 189.41	-14 449.11	463 117 88.00
Bénéfice attribué ou perte transférée	3.06	332.42	-37 862.84	386 952.94
Perte supportée ou bénéfice transféré	1.78	184.91	-33 282.83	123 041.00
Impôt sur les bénéfices	34.23	2410.06	-1 528 725.00	1 176 369.00
Valeur ajoutée hors taxe	983.31	33 504.27	-2 463 507.00	15 560 995.00
Taux de dépendance économique	89.75	11.52	70.51	130.18
Taux de chômage	8.34	1.88	4.00	16.20
Densité de la population	3484.09	6952.95	0.44	40 168.58
Nombre d'équipements en action sociale	0.01	0.03	0.00	3.51
Nombre d'équipements dans le commerce	0.13	0.47	0.00	7.91
Nombre d'équipements dans l'enseignement au 1er degré	0.01	0.02	0.00	1.79
Nombre d'équipements dans l'enseignement au 2nd degré	0.00	0.01	0.00	0.30
Nombre d'équipements dans l'enseignement supérieur et autres	0.00	0.01	0.00	0.31
Nombre d'équipements dans les services médicaux et paramédicaux	0.08	0.18	0.00	10.71
Nombre d'équipements dans les services de santé	0.01	0.02	0.00	2.34
Nombre d'équipements dans les services aux particuliers	0.34	1.02	0.00	251.57
Nombre d'équipements dans le sport et la culture	0.02	0.10	0.00	62.89
Nombre d'équipements dans le transport et le tourisme	0.04	0.16	0.00	34.88
Nombre d'établissements actifs	1.36	13.75	1.00	7544.00
Total des charges financières	106.04	14 349.06	-2.00	59 422 233.00
Emprunts et dettes assimilées	914.45	91 454.25	0.00	39 313 686.00
Investissement corporel et incorporel avec apports	196.28	21 049.02	-1187.00	11 840 107.00
Âge de l'entreprise en 2005	12.94	10.58	0.00	104.97
Nombre d'observations : 6 434 521				

TABLE 3.3.3 – Statistiques descriptives 2005-2015 : Entreprises fermées versus censurées

	Entreprises censurées				Entreprises fermées			
	Moyenne	Écart type	min	max	Moyenne	Écart type	min	max
Effectif de l'entreprise	15.61	350.81	0.00	164308.00	3.25	28.04	0.00	9149.00
Chiffre d'affaires	4182.80	104382.46	-561.82	41552852.00	466.65	8912.76	-1182.18	1819808.50
Bénéfice attribué ou perte transférée	3.18	254.55	-1171.58	117176.92	0.23	24.29	-32.45	7835.00
Perte supportée ou bénéfice transféré	2.19	200.18	-1660.80	90959.65	0.15	20.09	-2643.51	6495.71
Impôt sur les bénéfices	37.89	2379.71	-540614.31	816725.63	5.42	349.34	-11347.00	163453.00
Valeur ajoutée hors taxe	1168.86	33405.67	-739739.63	14273798.00	123.01	1457.39	-25673.61	404475.78
Taux de dépendance économique	89.71	11.51	70.51	130.18	90.20	11.67	70.51	130.18
Taux de chômage	8.34	1.89	4.00	16.20	8.35	1.87	4.00	16.20
Densité de la population	3451.40	6931.68	0.44	40168.58	3422.06	6816.97	0.44	40168.58
Nombre d'équipements en action sociale	0.01	0.03	0.00	3.51	0.01	0.03	0.00	3.51
Nombre d'équipements dans le commerce	0.13	0.47	0.00	7.91	0.11	0.44	0.00	7.91
Nombre d'équipements dans l'enseignement au 1er degré	0.01	0.01	0.00	1.79	0.01	0.02	0.00	1.79
Nombre d'équipements dans l'enseignement au 2nd degré	0.00	0.01	0.00	0.30	0.00	0.01	0.00	0.30
Nombre d'équipements dans l'enseignement supérieur et autres	0.00	0.01	0.00	0.31	0.00	0.01	0.00	0.31
Nombre d'équipements dans les services médicaux et paramédicaux	0.08	0.18	0.00	10.71	0.08	0.17	0.00	10.71
Nombre d'équipements dans les services de santé	0.01	0.02	0.00	2.34	0.01	0.02	0.00	2.34
Nombre d'équipements dans les services aux particuliers	0.34	1.04	0.00	251.57	0.31	0.91	0.00	58.14
Nombre d'équipements dans le sport et la culture	0.02	0.11	0.00	62.89	0.02	0.06	0.00	6.21
Nombre d'équipements dans le transport et le tourisme	0.04	0.17	0.00	34.88	0.04	0.16	0.00	34.88
Nombre d'établissements actifs	1.41	14.97	1.00	7544.00	1.12	1.56	1.00	517.00
Total des charges financières	123.86	15654.44	-2.00	5942233.00	12.97	579.42	-2.00	135870.00
Emprunts et dettes assimilées	1062.75	99769.33	0.00	39313686.00	129.31	4247.83	0.00	948606.00
Investissement corporel et incorporel avec apports	225.69	22948.39	-1187.00	11840107.00	31.32	1645.50	-658.00	460039.00
Âge de l'entreprise en 2005	12.67	10.45	0.00	104.97	13.97	11.30	0.00	101.06
Nombre d'observations : 491 377 censurées et 269 731 fermées								

Sur les 761 108 entreprises étudiées, 269 731 ont fermé au cours de la période 2005-2015. On note 19,85% des fermetures en 2005; 21,32% en 2006; 17,17% en 2007; 5,98% en 2008; 6,56% en 2009; 6,71% en 2010; 6,25% en 2011; 5,88% en 2012; 5,31% en 2013; 4,93% en 2014 et 0,03% en 2015.

Le tableau 3.3.3 montre que les entreprises ayant fermé sur la période d'étude sont en moyenne les plus petites, que ce soit en termes d'effectifs, de chiffre d'affaires ou de valeur ajoutée. En effet, les entreprises ayant fermé ont un effectif de plus de quatre fois inférieur à celui des autres, un chiffre d'affaires et une valeur ajoutée de neuf fois inférieur à ceux des autres entreprises.

Concernant les communes d'implantation des quartiers généraux, les résultats du tableau 3.3.3 ne font pas état de différences notables entre les entreprises ayant fermé et les autres. Les localisations des différents quartiers généraux semble assez comparables en termes de populations et d'infrastructures. Des statistiques descriptives supplémentaires sont disponibles dans l'annexe 3.A.

3.3.2 Modèles économétriques

Dans cette partie, nous présentons la stratégie économétrique adoptée. Cette stratégie vise principalement à étudier la survie des entreprises en fonction des caractéristiques du lieu d'implantation de leur quartier général. Pour ce faire, nous avons estimé plusieurs modèles à hasards proportionnels de Cox (voir section 1.2.1). Ces modèles nous permettent de mesurer, pour les entreprises, la sensibilité du risque instantané de fermeture aux variations des caractéristiques considérées.

La variable dépendante est donc le taux instantané d'échec noté $h(t)$ encore appelé taux de hasard. Les variables explicatives, $X(t)$, considérées, sont celles que nous avons précédemment présentées dans la section 3.3.1. Conditionnellement aux variables $X(t)$, le taux de hasard $h(t | X(t))$ s'écrit :

$$h(t | X(t)) = h_0(t)\exp(X(t)\delta)$$

Le terme $h_0(t)$ représente le risque de base commun à toutes les entreprises, et δ correspond au vecteur des coefficients.

Le modèle est estimé par maximum de vraisemblance. La variable dépendante n'étant pas directement observée, deux indicateurs importants devront être calculés avant l'esti-

mation. Il s'agit de la durée de vie des entreprises, et une variable muette prenant la valeur 1 si l'entreprise a fermé sur la période d'observation, et 0 sinon (voir section 1.2.1). Le répertoire SIRENE fournit les informations nécessaires (dates de création, états administratifs) pour calculer ces indicateurs.

La matrice $X(t)$ contient les variables explicatives d'intérêts, qui sont en l'occurrence les caractéristiques de l'environnement spatial du quartier général; mais aussi plusieurs variables de contrôles relatives à l'entreprise et son secteur d'activité. De façon générale, les variables explicatives incluses dans un modèle de Cox sont considérées comme fixes dans le temps et sont souvent mesurées à la date de début de l'étude. Il est cependant possible d'incorporer dans le modèle, la variabilité temporelle de certaines variables explicatives. Certaines de nos variables explicatives sont ainsi mesurées au début de l'étude, et d'autres sont observées chaque année de la période d'observation. Pour un modèle contenant n_1 variables explicatives mesurées chaque année dans le temps et n_2 variables explicatives mesurées à l'instant $t_0(2005)$, nous avons :

$$h(t, X(t)) = h_0(t) \exp \left[\sum_{i=1}^{n_1} \beta_i X_i(t) + \sum_{j=1}^{n_2} \gamma_j X_j \right]$$

Nous avons également estimé un modèle de Cox avec stratification (section 1.2.4.2) afin de prendre en compte l'hétérogénéité des entreprises en termes de taille. D'abord, nous nous sommes référés à une classification de l'INSEE⁸ pour regrouper les entreprises en 4 groupes. On distingue les micros entreprises qui ont une taille inférieure à 10 employés; les petites et moyennes entreprises qui emploient entre 10 et 249 personnes; les entreprises de taille intermédiaire comptant entre 250 et 4 999 employés; et les grandes entreprises dont la taille excède les 4 999 employés. Ensuite, nous avons procédé à l'estimation en supposant un hasard de base commun aux entreprises de la même catégorie.

3.3.3 Résultats

Le tableau 3.3.4 présente les coefficients estimés de quatre différents modèles. Le premier modèle prend en compte les caractéristiques de l'environnement spatial du quartier général de l'entreprise, le second modèle inclut les caractéristiques de l'entreprise, le troisième modèle porte sur les caractéristiques de l'entreprise et celles de l'environnement

8. Institut National de la Statistique et des Études Économiques.

spatial du quartier général puis le dernier modèle rajoute les secteurs d'activités aux variables du troisième modèle. Les ratios de risques se trouvent dans l'annexe 3.B. Des tests de rapport de vraisemblance sont effectués pour les modèles trois et quatre, avec le modèle deux comme modèle contraint. Les résultats indiquent qu'il est statistiquement pertinent de considérer les caractéristiques de l'environnement spatial du quartier général de l'entreprise et les secteurs d'activités.

Avant de commenter plus en détails ces résultats, nous proposons de présenter sommairement les signes des coefficients associés aux variables et leur significativité.

3.3.3.1 Signes des coefficients et Significativité

Dans le premier modèle, 9 des 13 variables ont des coefficients significatifs. D'un côté, le taux de dépendance économique de la zone d'emploi; la densité de la population dans la commune; le nombre d'équipements dans l'enseignement du premier degré pour 100 personnes par km² dans la commune; le nombre d'équipements dans l'enseignement au second degré pour 100 personnes par km² dans la commune; le nombre d'équipements dans les services aux particuliers, pour 100 personnes par km² dans la commune, impactent négativement la survie des entreprises. De l'autre côté, le nombre d'équipements dans le commerce pour 100 personnes par km² dans la commune; le nombre d'équipements dans l'enseignement supérieur la formation et les services d'éducation pour 100 personnes par km² dans la commune; le nombre d'équipements dans les services médicaux et paramédicaux pour 100 personnes par km² dans la commune; le nombre d'équipements dans le sport et la culture pour 100 personnes par km² dans la commune; augmentent les chances de survie des entreprises.

Dans le second modèle, tous les coefficients sont significatifs à l'exception de ceux associés aux variables perte supportée ou bénéfice transféré et total des charges financières. Les résultats montrent que l'effectif de l'entreprise; le chiffre d'affaires total; le bénéfice attribué ou perte transférée; l'impôt sur les bénéfices; la valeur ajoutée hors taxe; le nombre d'établissements actifs; les emprunts et dettes assimilées; réduisent les risques de mortalité de l'entreprise alors que les investissements corporels et incorporels avec apports; et l'âge, augmentent les risques de mortalité.

TABLE 3.3.4 – Résultats

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Dépendance économique	0.00368*** (0.000202)		0.00113*** (0.000202)	0.0000701 (0.000203)
Chômage	0.000391 (0.00103)		-0.00189 (0.00104)	-0.000942 (0.00104)
Densité de population	0.00000377*** (0.000000372)		0.00000228*** (0.000000372)	0.00000324*** (0.000000371)
Équipements en action sociale	-0.0728 (0.108)		0.0195 (0.107)	0.0643 (0.107)
Équipements dans le commerce	-0.0471*** (0.0143)		-0.0607*** (0.0140)	-0.0874*** (0.0145)
Équipements dans l'enseignement au 1er degré	0.919*** (0.174)		0.745*** (0.168)	0.289 (0.177)
Équipements dans l'enseignement au 2nd degré	1.667*** (0.397)		1.092** (0.398)	1.091** (0.396)
Équipements dans l'enseignement supérieur et autres	-1.870*** (0.300)		-1.945*** (0.297)	-2.231*** (0.295)
Équipements dans les services médicaux et paramédicaux	-0.0653* (0.0289)		-0.105*** (0.0287)	-0.0275 (0.0286)
Équipements dans les services de santé	-0.102 (0.176)		0.0135 (0.170)	0.0515 (0.169)
Équipements dans les services aux particuliers	0.0177* (0.00877)		0.0452*** (0.00862)	0.0589*** (0.00902)
Équipements dans le sport et la culture	-0.108* (0.0448)		-0.209*** (0.0422)	-0.293*** (0.0488)
Équipements dans le transport et le tourisme	0.0137 (0.0177)		-0.0190 (0.0171)	-0.0455* (0.0177)
Effectif		-0.0329*** (0.000525)	-0.0329*** (0.000525)	-0.0321*** (0.000538)
Chiffre d'affaires		-0.000119*** (0.00000277)	-0.000119*** (0.00000277)	-0.000141*** (0.00000291)
Bénéfice attribué ou perte transférée		-0.000537*** (0.000131)	-0.000523*** (0.000137)	-0.000399** (0.000139)
Perte supportée ou bénéfice transféré		-0.0000122 (0.000136)	-0.00000863 (0.000140)	0.0000619 (0.000117)
Impôt sur les bénéfices		-0.0000209*** (0.00000157)	-0.0000209*** (0.00000158)	-0.0000222*** (0.00000207)
Valeur ajoutée hors taxe		-0.0000522*** (0.000000705)	-0.0000521*** (0.000000706)	-0.0000557*** (0.000000746)
Nombre d'établissements actifs		-0.0349*** (0.00423)	-0.0343*** (0.00423)	-0.0286*** (0.00422)
Total des charges financières		-0.000000669 (0.00000343)	-0.000000547 (0.00000349)	0.000000256 (0.00000340)
Emprunts et dettes assimilées		-0.00000245*** (0.000000505)	-0.00000248*** (0.000000514)	-0.00000257*** (0.000000501)
Investissement corporel et incorporel		0.0000136*** (0.00000150)	0.0000136*** (0.00000149)	0.0000141*** (0.00000135)
Âge de l'entreprise en 2005		0.0155*** (0.000177)	0.0156*** (0.000177)	0.0156*** (0.000177)
Industrie manufacturière				0.192*** (0.00751)
Secteur de la construction				0.283*** (0.00665)
Secteur commerce, transport, hébergement et restauration				0.395*** (0.00538)
Secteur activités spécialisées, scientifiques techniques, services administratifs et soutien				-0.0261** (0.00793)
N	6 434 521	6 434 521	6 434 521	6 434 521
Degré de liberté du test Ratio-Vraisemblance			13	17
Chi2 du test Ratio-Vraisemblance			236.1	7877.1
Probabilité critique du test Ratio-Vraisemblance			4.94e-43	0

Ecart types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Le sens des impacts observés dans les modèles 1 et 2 est confirmé par le modèle 3 tandis que de légers changements sont constatés dans le modèle 4 au niveau de certaines variables décrivant l'environnement spatial du quartier général de l'entreprise. En effet, lorsque les secteurs d'activité sont pris en compte, le taux de dépendance économique, le nombre d'équipements dans l'enseignement au 1er degré et le nombre d'équipements dans les services médicaux et paramédicaux n'ont plus d'effets significatifs alors que le coefficient estimé du nombre d'équipements dans le transport et le tourisme est devenu significatif.

3.3.3.2 Discussion

Environnement spatial du quartier général de l'entreprise

Les résultats du modèle 4 nous indiquent que le risque instantané de mortalité d'une entreprise est significativement sensible à plusieurs caractéristiques de l'environnement spatial de son quartier général.

D'après nos estimations, les entreprises qui implantent leur quartier général dans une commune plus densément peuplée ont un risque de fermeture plus élevé. Les grandes agglomérations sont un lieu de prédilection pour les quartiers généraux d'entreprises (A. W. Evans, 1973; Aarland et al., 2007; Zhang et al., 2018) en raison des multiples avantages (informations, sous-traitance, capital humain) qu'elles offrent. Cette attractivité des grandes métropoles, conduit à une concentration spatiale de quartiers généraux (Lovely et al., 2005; Ono, 2006; Davis & Henderson, 2008; Strauss-Kahn & Vives, 2009), et peut en conséquence engendrer une concurrence très agressive autour des ressources. Cela pourrait expliquer l'effet négatif de la densité de la commune d'implantation du quartier général sur la survie de l'entreprise.

Lorsque la commune d'implantation du quartier général de l'entreprise est davantage dotée d'équipements dans l'enseignement au second degré, et dans les services aux particuliers, les chances de survie de l'entreprise sont réduites. Les quartiers généraux pour mieux gérer les différentes unités de production, ont besoin d'avoir autour d'eux une diversité d'industrie (Davis & Henderson, 2008) et la possibilité d'externaliser certains services (Aarland et al., 2007). Les services offerts par des équipements dans l'enseignement au second degré, et dans les services aux particuliers, semblent assez peu utiles à des quartiers généraux d'entreprises du secteur marchand non agricole. Les communes abondamment dotées de ce genre d'équipements ne sauraient donc être des localisations optimales pour ces quartiers généraux.

Le nombre d'équipements dans le commerce, dans l'enseignement supérieur la formation et les services d'éducation, dans le sport et la culture, dans le transport et le tourisme, réduit les risques d'échec des entreprises. Un équipement additionnel dans le commerce pour 100 personnes par km² dans la commune d'implantation du quartier général de l'entreprise, réduit le risque instantané de mortalité de l'entreprise de 8,4%. La réduction du risque de mortalité est respectivement de 89,3%, 25,4% et 4,5% pour les équipements de l'enseignement supérieur la formation et les services d'éducation; les équipements dans le sport et la culture, et les équipements dans le transport et le tourisme. La majorité (40,43%) des entreprises de notre échantillon opèrent dans le secteur du commerce de gros et de détail, transport, hébergement et restauration. L'implantation du quartier général dans une commune comptant un nombre élevé d'équipements dans le commerce, peut permettre à l'entreprise d'avoir accès à des informations utiles. Cela pourrait justifier l'impact positif de la variable : nombre d'équipements dans le commerce.

Les équipements dans l'enseignement supérieur la formation et les services d'éducation, prennent en compte plusieurs institutions éducatives telles que les universités, les grandes écoles et divers autres centres de formations (voir annexe 1.A). Les communes ayant davantage ces équipements, peuvent permettre aux quartiers généraux d'avoir facilement accès au capital humain, qui constitue pour eux, l'une des ressources les plus importantes (Adler & Florida, 2020).

Les équipements en transport et tourisme incluent des infrastructures clés comme les gares et les aéroports. Les infrastructures aéroportuaires favorisent le développement des activités économiques et l'emploi (Neal, 2011; Florida et al., 2015), et peuvent également permettre aux quartiers généraux entreprises d'attirer plus facilement des employés qualifiés depuis l'extérieur. L'impact positif des équipements en transport et tourisme sur la survie peut donc se comprendre, et pourrait également s'inscrire dans la continuité des travaux de Strauss-Kahn & Vives (2009) et Adler & Florida (2020) qui ont, entre autres, montré que les villes dotées d'aéroports sont particulièrement convoitées par les quartiers généraux d'entreprises.

Variables de contrôle

Au niveau des caractéristiques spécifiques de l'entreprise, on note un effet positif de la taille de l'entreprise sur sa survie, et un effet négatif de l'âge. L'impact négatif de l'âge sur la survie, bien que contre-intuitif, n'est pas inexistant dans la littérature. Cela s'explique par la notion d'"aléa de la sénescence" (Hannan, 1998) qui stipule que les entreprises âgées voient leurs chances de survie réduites en raison des difficultés à s'adapter à un environnement

de plus en plus concurrentiel.

Un établissement actif en plus, réduit le risque de fermeture de 2,8%. Ce résultat est tout à fait logique d'autant plus qu'un nombre élevé d'établissements actifs traduit une certaine expansion économique de l'entreprise. En effet, l'implantation d'un nouvel établissement peut entraîner l'augmentation des parts de marché de l'entreprise, ce qui lui est bénéfique toutes choses égales par ailleurs.

Le chiffre d'affaires total, le bénéfice attribué ou perte transférée, l'impôt sur les bénéfices, la valeur ajoutée hors taxe, et les emprunts et dettes assimilées, ont un effet positif sur la survie des entreprises. Une valeur élevée de ses variables comptables reflète, d'un côté, la réussite de l'entreprise au niveau de ses activités de production, et de l'autre côté un capital financier conséquent. Toutefois, on note que les investissements corporels et incorporels avec apports affectent négativement la survie des entreprises.

Concernant les secteurs d'activités, les résultats d'estimation montrent que les entreprises qui opèrent dans l'industrie manufacturière ont plus de risque de fermer que les autres. Il en est de même pour celles opérant dans le secteur de la construction; et celui du commerce de gros et de détail, transport, hébergement et restauration. Quant aux entreprises menant des activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien; elles ont moins de risque de fermer que les autres.

Cette hétérogénéité inter-sectorielle du risque de fermeture entre les entreprises, a été déjà constatée dans la littérature empirique (Mata & Portugal, 1994; Audretsch, 1995; Baumöhl et al., 2020). Les secteurs d'activité où les entreprises vivent moins longtemps, sont généralement caractérisés par un fort degré de concurrence (Brown & Earle, 2000; Van Kranenburg et al., 2002) et une échelle minimale d'efficacité (niveau de production à atteindre pour minimiser le coût moyen) élevée (Audretsch, 1991; Fritsch et al., 2006).

3.3.3.3 Stratification par catégories d'entreprises

Comme mentionné dans la section 3.3.2, nous avons regroupé les entreprises en 4 catégories et estimé un modèle de Cox avec stratification. Nous distinguons les grandes entreprises (GRD), les entreprises de tailles intermédiaires (ETI), les petites et moyennes entreprises (PME), et les micro entreprises (MCR). Les tableaux 3.3.5, 3.3.6 et 3.3.7 présentent les fréquences relatives annuelles.

TABLE 3.3.5 – Description des catégories d'entreprises : 2005-2008

	2005		2006		2007		2008	
	effectif	proportion	effectif	proportion	effectif	proportion	effectif	proportion
Entreprises de tailles intermédiaires	3 300	0.43	3 376	0.48	3 510	0.55	3712	0.62
Grandes entreprises	70	0.01	68	0.01	72	0.01	75	0.01
Micro entreprises	663 860	87.22	605 369	86.11	544 628	84.57	494 662	83.12
Petites et moyennes entreprises	93 878	12.33	94 200	13.40	95 789	14.87	96 665	16.24
		100.00		100.00		100.00	595 114	100.00
Total	761 108		703 013		643 999		595 114	

TABLE 3.3.6 – Description des catégories d'entreprises : 2009-2012

	2009		2010		2011		2012	
	effectif	proportion	effectif	proportion	effectif	proportion	effectif	proportion
Entreprises de tailles intermédiaires	3 739	0.64	3 594	0.64	3 632	0.66	3 617	0.68
Grandes entreprises	82	0.01	79	0.01	78	0.01	79	0.01
Micro entreprises	482 408	83.05	468 547	83.07	451 596	82.60	437 007	82.35
Petites et moyennes entreprises	94 615	16.29	91 798	16.28	91 401	16.72	89 980	16.96
		100.00		100.00		100.00		100.00
Total	580 844		564 018		546 707		530 683	

TABLE 3.3.7 – Description des catégories d'entreprises : 2013-2015

	2013		2014		2015	
	effectif	proportion	effectif	proportion	effectif	proportion
Entreprises de tailles intermédiaires	3 612	0.70	3 618	0.72	3 604	0.73
Grandes entreprises	85	0.02	80	0.02	82	0.02
Micro entreprises	423 795	82.24	412 641	82.16	404 060	82.21
Petites et moyennes entreprises	87 832	17.04	85 904	17.10	83 722	17.04
		100.00		100.00		100.00
Total	515 324		502 243		491 468	

FIGURE 3.3.1 – Fonctions de survie de Kaplan Meier

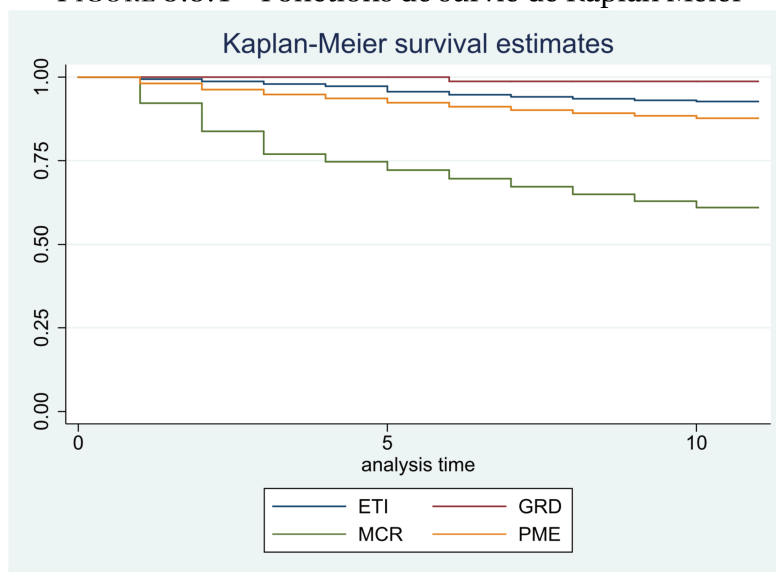


TABLE 3.3.8 – Test Log-rank

Catégories	Observés	Espérés
Entreprises de tailles intermédiaires	271	1497.38
Grandes entreprises	1	31.76
Micro entreprises	257 343	228 566.21
Petites et moyennes entreprises	12 116	39 635.66
Total	269 731	269 731.00
	$\chi^2(3)=25\ 176.42$	
	$\text{Pr}>\chi^2=0.0000$	

La figure 3.3.1 et le tableau 3.3.8 présentent respectivement les fonctions de survie de Kaplan-Meier et le log-rank test. Les résultats montrent clairement que les entreprises de grandes tailles vivent plus longtemps.

Le tableau 3.3.9 contient les résultats du modèle avec stratification. Comparativement au modèle précédemment interprété, il n'y a eu aucun changement d'impact au niveau des variables décrivant l'environnement spatial et les secteurs d'activités. En revanche, on note des changements au niveau de certaines variables comptables.

Dans le modèle sans stratification, l'impôt sur les bénéfices avait un effet positif sur la survie des entreprises, mais dans le modèle avec stratification l'effet devient négatif. Les

impôts sont proportionnels aux bénéfices réalisés, et les entreprises les plus profitables ont de meilleures perspectives de survie (Fotopoulos & Louri, 2000). L'impact positif des impôts sur la survie, est donc beaucoup plus intuitif. Par ailleurs, si les grandes entreprises, grâce à des profits colossaux, ne ressentent pas toujours la pression des charges fiscales; les petites, moyennes et micros entreprises peuvent avoir du mal à s'en sortir en raison des impôts. Cevik & Miryugin (2022) travaillent sur un échantillon de plus de 4 millions d'entreprises de 21 pays, et trouvent que le taux d'imposition augmente significativement le risque d'échec des entreprises. Nous pensons ainsi que la prise en compte de l'hétérogénéité des entreprises en termes de taille, a permis d'apporter plus de nuances quant à l'effet de l'impôt sur bénéfices. Cette stratification par la taille a également entraîné des changements de significativité au niveau des variables *total des charges financières*, et *emprunts et dettes assimilées*. En effet, le *total des charges financières* est devenu significatif avec un effet positif sur la survie alors que la variable *emprunts et dettes assimilées* a perdu sa significativité.

TABLE 3.3.9 – Modèle avec stratification

Modèle	Coefficients
Dépendance économique	0.000158 (0.000203)
Chômage	-0.000611 (0.00104)
Densité de population	0.00000313** (0.000000371)
Équipements en action sociale	0.0503 (0.108)
Équipements dans le commerce	-0.0875** (0.0144)
Équipements dans l'enseignement au 1er degré	0.314 (0.175)
Équipements dans l'enseignement au 2nd degré	1.050* (0.398)
Équipements dans l'enseignement supérieur et autres	-2.216** (0.294)
Équipements dans les services médicaux et paramédicaux	-0.0223 (0.0286)
Équipements dans les services de santé	0.00288 (0.171)
Équipements dans les services aux particuliers	0.0600** (0.00894)
Équipements dans le sport et la culture	-0.288** (0.0470)
Équipements dans le transport et le tourisme	-0.0475* (0.0176)
Chiffre d'affaires	-0.000164** (0.00000257)
Bénéfice attribué ou perte transférée	-0.000379** (0.000105)
Perte supportée ou bénéfice transféré	0.0000555 (0.0000947)
Impôt sur les bénéfices	0.0000202** (0.00000371)
Valeur ajoutée hors taxe	-0.0000423** (0.00000152)
Nombre d'établissements actifs	-0.0444** (0.00408)
Total des charges financières	-0.0000129** (0.00000295)
Emprunts et dettes assimilées	0.000000375 (0.000000367)
Investissement corporel et incorporel	0.00000766** (0.00000142)
Âge de l'entreprise en 2005	0.0157** (0.000176)
Industrie manufacturière	0.183** (0.00747)
Secteur de la construction	0.267** (0.00664)
Secteur commerce, transport, hébergement et restauration	0.386** (0.00538)
Secteur activités spécialisées, scientifiques techniques, services administratifs et soutien	-0.0369** (0.00792)
<i>N</i>	6 434 521

Écarts types entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conclusion

Le quartier général de l'entreprise est le centre de commande où se décide la coordination des différentes unités de production. Sa mission est tout aussi importante que celle des unités de production car une mauvaise gestion peut aussi bien conduire à la faillite qu'un défaut de production. Plusieurs travaux dans la littérature (Lovely et al., 2005; Ono, 2006; Aarland et al., 2007; Davis & Henderson, 2008; Strauss-Kahn & Vives, 2009) nous apprennent que le lieu d'implantation du quartier général n'est pas choisi de façon hasardeuse et ce choix peut avoir des conséquences sur les performances des différentes unités de production (Audia & Rider, 2010; Landier et al., 2009; Kalnins & Lafontaine, 2013). Toutefois, ces conséquences ne sont que partiellement explorées dans la littérature, en ce sens que les travaux existants se limitent à analyser l'impact de la distance entre quartier général et établissements, sur la survie de ces derniers.

Dans ce chapitre, nous avons étudié dans un cadre plus général, les conséquences du choix de localisation du quartier général de l'entreprise. Nous nous sommes notamment intéressés au risque de fermeture de l'entreprise en fonction de l'environnement spatial de son quartier général.

L'estimation de modèles à hasards proportionnels de Cox sur un échantillon de 761 108 entreprises françaises, nous a permis de voir que les caractéristiques du lieu d'implantation du quartier général d'une entreprise ont un impact significatif sur la survie de cette dernière, toutes choses égales par ailleurs. D'après nos résultats, les entreprises qui implantent leur quartier général dans une commune avec beaucoup d'équipements dans le commerce, dans l'enseignement supérieur la formation et les services d'éducation, dans le sport et la culture, dans le transport et le tourisme, ont plus de chances de survie. Ces équipements semblent fournir aux quartiers généraux les ressources nécessaires pour mieux coordonner les activités de l'entreprise.

En revanche, le fait que la commune de localisation du quartier général soit plus densément peuplée, ou le fait qu'elle soit davantage dotée d'équipements dans l'enseignement au second degré, et dans les services aux particuliers, réduit les chances de survie de l'entreprise. L'impact négatif de la densité de la population peut s'expliquer par la forte concurrence que les entreprises se livrent dans les grandes agglomérations. Quant aux équipements dans l'enseignement au second degré, et dans les services aux particuliers, ils proposent des services beaucoup plus destinés aux ménages qu'aux entreprises. Les communes abondamment dotées de ces équipements ne sauraient donc être considé-

rées comme des localisations optimales pour des quartiers généraux d'entreprises.

Ce chapitre et les deux précédents, montrent que l'environnement spatial joue un rôle important dans la survie de l'entreprise. Toutefois, la définition de l'environnement spatial est plutôt large dans la mesure où l'entreprise est observée dans sa commune ou dans sa zone d'emploi. Cette conception de l'espace est l'une des limites de ces trois premiers chapitres car les réalités mises en exergue lorsque l'on observe les caractéristiques de la commune ou la zone d'emploi d'implantation de l'entreprise, ne sont pas forcément informatives à propos de ce qui se passe dans le voisinage très immédiat de l'entreprise.

Une entreprise est inévitablement impactée par sa relation avec les entreprises voisines, mais les variables mesurées à l'échelle de la commune ou de la zone d'emploi, peuvent par exemple masquer certaines réalités microéconomiques découlant de ces interactions. L'un des résultats les plus récurrents le long des trois chapitres est l'effet négatif de la densité de population sur la survie des entreprises. Ce résultat que nous avons expliqué par la forte concurrence présente dans les milieux urbains, reflète probablement une facette des interactions entre entreprises voisines dans les grandes agglomérations. La variable densité de population en elle-même n'a pas de fondements théoriques microéconomiques, et donc ne permet pas de décrire convenablement les mécanismes derrière la concurrence entre entreprises.

Le chapitre 4, toujours dans la perspective d'explicitier la relation entre la survie de l'entreprise et son environnement spatial, propose d'étudier cet effet de concurrence à partir d'une modélisation économétrique sur micro-données.

Annexes

3.A Statistiques descriptives

TABLE 3.A.1 – Secteurs d'activités

	BE ¹		FZ ²		GI ²		MN ⁴	
	effectif	pourcentage	effectif	pourcentage	effectif	pourcentage	effectif	pourcentage
0	666473	87.57	643754	84.58	453405	59.57	674779	88.66
1	94635	12.43	117354	15.42	307703	40.43	86329	11.34
Total	761 108	100.00	761 108	100.00	761 108	100.00	761 108	100.00

¹ Industrie manufacturière, industries extractives et autres

² Construction

³ Commerce de gros et de détail, transports, hébergement et restauration

⁴ Activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien

TABLE 3.A.2 – Variables quantitatives 2005-2007

	2005					2006					2007				
	mo	écart	min	max	type	mo	écart	min	max	type	mo	écart	min	max	type
Effectif de l'entreprise	10,78576	338.44	0.00	183196.00		11,52728	338.40	0.00	177614.00		12,82381	340.73	0.00	162391.00	
Chiffre d'affaires	2470,92681	79875.89	-85.00	31339256.00		2828,02715	88916.56	-89.00	33718100.00		3256,62892	97936.38	-34.00	36717648.00	
Bénéfice attribué ou perte transférée	2,47490	357.26	-132.00	248134.00		2,33751	212.63	-71.00	113300.00		2,98243	235.35	-7000.00	78126.00	
Perte supportée ou bénéfice transféré	1,19499	116.36	-3.00	68779.00		1,29796	132.99	-830.00	82906.00		1,65033	190.20	-27.00	102331.00	
Impôt sur les bénéfices	27,32885	2236.82	-1528725.00	567787.00		34,72876	2437.27	-1103772.00	1176369.00		41,16481	2884.71	-1516872.00	834587.00	
Valeur ajoutée hors taxe économique	717,54796	33034.07	-1066328.00	14755224.00		813,54228	35127.33	-2358761.00	15560995.00		916,99176	33310.13	-2463507.00	14076333.00	
Taux de dépendance	89,81551	11.57	70.51	130.18		89,81107	11.55	70.51	130.18		89,79640	11.52	70.51	130.18	
Taux de chômage	8,34399	1.88	4.00	16.20		8,34397	1.88	4.00	16.20		8,33991	1.88	4.00	16.20	
Densité de population	3520,60037	6973.76	0.44	40168.58		3498,61607	6954.49	0.44	40168.58		3480,48600	6939.15	0.44	40168.58	
Nombre d'équipements en action sociale	0,00964	0.03	0.00	3.51		0,00962	0.03	0.00	3.51		0,00964	0.03	0.00	3.51	
Nombre d'équipements dans le commerce	0,12498	0.47	0.00	7.91		0,12472	0.47	0.00	7.91		0,12545	0.47	0.00	7.91	
Nombre d'équipements dans l'enseignement au 1er degré	0,00712	0.02	0.00	1.79		0,00712	0.02	0.00	1.79		0,00711	0.02	0.00	1.79	
Nombre d'équipements dans l'enseignement au 2nd degré	0,00239	0.01	0.00	0.30		0,00238	0.01	0.00	0.30		0,00238	0.01	0.00	0.30	
Nombre d'équipements dans l'enseignement supérieur et autres	0,00318	0.01	0.00	0.31		0,00317	0.01	0.00	0.31		0,00319	0.01	0.00	0.31	
Nombre d'équipements dans les services médicaux et paramédicaux	0,08106	0.18	0.00	10.71		0,08089	0.18	0.00	10.71		0,08106	0.18	0.00	10.71	
Nombre d'équipements dans les services de santé	0,00983	0.02	0.00	2.34		0,00982	0.02	0.00	2.34		0,00983	0.02	0.00	2.34	
Nombre d'équipements dans les services aux particuliers	0,33279	1.00	0.00	251.57		0,33202	1.00	0.00	251.57		0,33312	1.01	0.00	251.57	
Nombre d'équipements dans le sport et la culture	0,02415	0.09	0.00	62.89		0,02420	0.09	0.00	62.89		0,02420	0.10	0.00	62.89	
Nombre d'équipements dans le transport et le tourisme	0,04021	0.16	0.00	34.88		0,03995	0.15	0.00	34.88		0,04003	0.15	0.00	34.88	
Nombre d'établissements actifs	1,30486	12.07	1.00	7544.00		1,32226	12.55	1.00	7544.00		1,34490	13.11	1.00	7544.00	
Total des charges financières	84,56084	12583.13	-2.00	5942233.00		90,74819	13091.50	-2.00	5942233.00		97,43756	13675.91	-2.00	5942233.00	
Emprunts et dettes assimilées	731,94336	80205.45	0.00	39313686.00		785,54657	83450.55	0.00	39313686.00		844,10944	87171.31	0.00	39313686.00	
Investissement corporel et incorporel avec apports	156,80791	18465.20	-1187.00	11840107.00		168,32400	19212.45	-1187.00	11840107.00		182,00634	20069.83	-1187.00	11840107.00	
Âge de l'entreprise en 2005	13,13183	10.78	0.00	104.97		13,09164	10.71	0.00	104.97		13,05840	10.65	0.00	104.97	
N	761108					703013					643999				

TABLE 3.A.3 – Variables quantitatives 2008-2010

	2008					2009					2010				
	moyenne	écart type	min	max	max	moyenne	écart type	min	max	max	moyenne	écart type	min	max	max
Effectif de l'entreprise	14.29362	351.16	0.00	159262.00	171922.00	14.49627	338.17	0.00	171922.00	171922.00	14.41949	336.40	0.00	170663.00	170663.00
Chiffre d'affaires	3625.83162	110353.93	-14449.11	44582916.00	38895280.00	3376.60525	99453.30	-1182.18	38895280.00	38895280.00	3618.93565	107164.86	-2616.47	40906332.00	40906332.00
Bénéfice attribué ou perte transférée	3.89667	549.03	-2045.58	366409.00	386952.94	3.62591	557.53	-4120.00	386952.94	386952.94	2.85187	259.20	-37862.84	95707.38	95707.38
Perte supportée ou bénéfice transféré	2.08139	239.18	-1255.05	123041.00	120184.00	2.09465	235.96	-3767.02	120184.00	120184.00	1.75792	181.34	-33282.83	69281.00	69281.00
Impôt sur les bénéfices	32.14395	2574.74	-1517028.00	336806.63	401597.66	27.28707	2215.33	-1274391.50	401597.66	401597.66	33.19420	2415.17	-1214191.13	659593.75	659593.75
Valeur ajoutée hors taxe	983.61255	31591.17	-195498.53	13806726.00	13554915.00	956.97008	31445.21	-262557.66	13554915.00	13554915.00	1018.51610	32516.13	-178848.22	12777733.00	12777733.00
Taux de dépendance économique	89.75158	11.51	70.51	130.18	130.18	89.73271	11.51	70.51	130.18	130.18	89.72595	11.52	70.51	130.18	130.18
Taux de chômage	8.33833	1.88	4.00	16.20	16.20	8.33935	1.88	4.00	16.20	16.20	8.33906	1.88	4.00	16.20	16.20
Densité de la population	3486.72118	6951.63	0.44	40168.58	40168.58	3491.27341	6961.87	0.44	40168.58	40168.58	3485.75147	6958.92	0.44	40168.58	40168.58
Nombre d'équipements en action sociale	0.00965	0.03	0.00	3.51	3.51	0.00965	0.03	0.00	3.51	3.51	0.00965	0.03	0.00	3.51	3.51
Nombre d'équipements dans le commerce	0.12614	0.47	0.00	7.91	7.91	0.12650	0.47	0.00	7.91	7.91	0.12705	0.47	0.00	7.91	7.91
Nombre d'équipements dans l'enseignement au 1er degré	0.00709	0.02	0.00	1.79	1.79	0.00710	0.02	0.00	1.79	1.79	0.00711	0.02	0.00	1.79	1.79
Nombre d'équipements dans l'enseignement au 2nd degré	0.00239	0.01	0.00	0.30	0.30	0.00238	0.01	0.00	0.30	0.30	0.00238	0.01	0.00	0.30	0.30
Nombre d'équipements dans l'enseignement supérieur et autres	0.00321	0.01	0.00	0.31	0.31	0.00323	0.01	0.00	0.31	0.31	0.00324	0.01	0.00	0.31	0.31
Nombre d'équipements dans les services médicaux et paramédicaux	0.08135	0.18	0.00	10.71	10.71	0.08150	0.18	0.00	10.71	10.71	0.08171	0.18	0.00	10.71	10.71
Nombre d'équipements dans les services de santé	0.00985	0.02	0.00	2.34	2.34	0.00985	0.02	0.00	2.34	2.34	0.00987	0.02	0.00	2.34	2.34
Nombre d'équipements dans les services aux particuliers	0.33465	1.02	0.00	251.57	251.57	0.33550	1.02	0.00	251.57	251.57	0.33660	1.02	0.00	251.57	251.57
Nombre d'équipements dans le sport et la culture	0.02419	0.10	0.00	62.89	62.89	0.02419	0.10	0.00	62.89	62.89	0.02424	0.10	0.00	62.89	62.89
Nombre d'équipements dans le transport et le tourisme	0.04033	0.17	0.00	34.88	34.88	0.04039	0.16	0.00	34.88	34.88	0.04042	0.15	0.00	34.88	34.88
Nombre d'établissements actifs	1.36586	13.64	1.00	7544.00	7544.00	1.37148	13.80	1.00	7544.00	7544.00	1.37593	14.00	1.00	7544.00	7544.00
Total des charges financières	104.49864	14226.13	-2.00	5942233.00	5942233.00	106.68180	14399.21	-2.00	5942233.00	5942233.00	109.39990	14612.30	-2.00	5942233.00	5942233.00
Emprunts et dettes assimilées	902.62545	90670.90	0.00	39313686.00	39313686.00	920.81402	91771.09	0.00	39313686.00	39313686.00	942.75767	93127.18	0.00	39313686.00	39313686.00
Investissement corporel et incorporel avec apports	195.48588	20877.31	-1187.00	11840107.00	11840107.00	198.89966	21127.94	-1187.00	11840107.00	11840107.00	203.47214	21440.01	-1187.00	11840107.00	11840107.00
Âge de l'entreprise en 2005	13.02458	10.60	0.00	104.58	104.58	12.95795	10.56	0.00	104.57	104.57	12.92137	10.53	0.00	104.97	104.97
N	595 114			580 844		580 844			564 018		564 018				

TABLE 3.A.4 – Variables quantitatives 2011-2013

	2011					2012					2013					
	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max
Effectif de l'entreprise	14.82485	339.64	0.00	168012.00	15.13964	344.02	0.00	168171.00	15.38081	348.31	0.00	167917.00				
Chiffre d'affaires	3928.87255	115707.80	-1088.89	43321536.00	4067.82294	123411.75	-1738.72	46311788.00	4101.53695	114871.66	-4386.78	43422984.00				
Bénéfice attribué ou perte transférée	3.27309	261.50	-295.38	114788.76	3.01220	200.23	-757.68	54366.22	3.04955	217.24	-5446.94	73656.53				
Perte supportée ou bénéfice transféré	1.79807	173.06	-1049.69	76639.00	1.83299	163.78	-688.95	61873.00	1.94711	167.55	-5477.32	74250.53				
Impôt sur les bénéfices	38.38344	2099.55	-656471.38	594121.56	35.43682	2793.05	-1284142.75	486487.50	36.61217	2288.78	-767744.88	747352.00				
Valeur ajoutée hors taxe	1071.23387	33173.21	-292912.50	12305871.00	1095.50801	32814.35	-112938.29	11796454.00	1131.95808	36171.98	-768883.38	14485726.00				
Taux de dépendance économique	89.71933	11.51	70.51	130.18	89.71451	11.51	70.51	130.18	89.70339	11.51	70.51	130.18				
Taux de chômage	8.34030	1.89	4.00	16.20	8.34116	1.89	4.00	16.20	8.34077	1.89	4.00	16.20				
Densité de la population	3480.85334	6957.25	0.44	40168.58	3473.27454	6952.09	0.44	40168.58	3468.27914	6947.28	0.44	40168.58				
Nombre d'équipements en action sociale	0.00965	0.03	0.00	3.51	0.00962	0.03	0.00	3.51	0.00962	0.03	0.00	3.51				
Nombre d'équipements dans le commerce	0.12739	0.47	0.00	7.91	0.12767	0.48	0.00	7.91	0.12782	0.48	0.00	7.91				
Nombre d'équipements dans l'enseignement au 1er degré	0.00711	0.02	0.00	1.79	0.00710	0.01	0.00	1.79	0.00709	0.01	0.00	1.79				
Nombre d'équipements dans l'enseignement au 2nd degré	0.00238	0.01	0.00	0.30	0.00237	0.01	0.00	0.30	0.00237	0.01	0.00	0.30				
Nombre d'équipements dans l'enseignement supérieur et autres	0.00325	0.01	0.00	0.31	0.00325	0.01	0.00	0.31	0.00326	0.01	0.00	0.31				
Nombre d'équipements dans les services médicaux et paramédicaux	0.08179	0.18	0.00	10.71	0.08182	0.18	0.00	10.71	0.08194	0.18	0.00	10.71				
Nombre d'équipements dans les services de santé	0.00986	0.02	0.00	2.34	0.00984	0.02	0.00	2.34	0.00984	0.02	0.00	2.34				
Nombre d'équipements dans les services aux particuliers	0.33727	1.02	0.00	251.57	0.33803	1.03	0.00	251.57	0.33881	1.03	0.00	251.57				
Nombre d'équipements dans le sport et la culture	0.02425	0.10	0.00	62.89	0.02422	0.10	0.00	62.89	0.02426	0.10	0.00	62.89				
Nombre d'équipements dans le transport et le tourisme	0.04047	0.15	0.00	34.88	0.04066	0.16	0.00	34.88	0.04082	0.17	0.00	34.88				
Nombre d'établissements actifs	1.38248	14.22	1.00	7544.00	1.38898	14.43	1.00	7544.00	1.39598	14.64	1.00	7544.00				
Total des charges financières	112.35512	14841.42	-2.00	5942233.00	115.43711	15063.70	-2.00	5942233.00	118.50472	15286.43	-2.00	5942233.00				
Emprunts et dettes assimilées	968.17335	94588.74	0.00	39313686.00	993.39177	96004.11	0.00	39313686.00	1018.74129	97424.25	0.00	39313686.00				
Investissement corporel et incorporel avec apports	208.01271	21767.65	-1187.00	11840107.00	213.45789	22093.75	-1187.00	11840107.00	217.80390	22413.45	-1187.00	11840107.00				
Âge de l'entreprise en 2005	12.88008	10.51	0.00	104.97	12.84276	10.49	0.00	104.97	12.79803	10.47	0.00	104.97				
N	546 707				530 683				515 324							

TABLE 3.A.5 – Variables quantitatives 2014-2015

	2014				2015			
	moyenne	écart type	min	max	moyenne	écart type	min	max
Effectif de l'entreprise	15.54871	352.26	0.00	167 048.00	15.61029	350.78	0.00	164 308.00
Chiffre d'affaires	4144.64914	111 382.52	-2225.29	41 716 660.00	4182.04963	104 372.81	-561.82	41 552 852.00
Bénéfice attribué ou perte transférée	3.34824	259.31	-1238.03	93 391.05	3.18129	254.53	-1171.58	117 176.92
Perte supportée ou bénéfice transféré	2.16986	219.85	-2643.51	94 481.77	2.18575	200.16	-1660.80	90 959.65
Impôt sur les bénéfices	34.93861	1951.64	-377 677.75	576 524.25	37.87903	2379.49	-540 614.31	816 725.63
Valeur ajoutée hors taxe	1151.09754	35975.25	-745 801.19	14 196 946.00	1168.65820	33 402.58	-739 739.63	14 273 798.00
Taux de dépendance économique	89.70286	11.51	70.51	130.18	89.71025	11.51	70.51	130.18
Taux de chômage	8.34154	1.89	4.00	16.20	8.34247	1.89	4.00	16.20
Densité de la population	3462.83754	6944.03	0.44	40 168.58	3451.43031	6931.65	0.44	40 168.58
Nombre d'équipements en action sociale	0.00961	0.03	0.00	3.51	0.00959	0.03	0.00	3.51
Nombre d'équipements dans le commerce	0.12806	0.48	0.00	7.91	0.12789	0.48	0.00	7.91
Nombre d'équipements dans l'enseignement au 1er degré	0.00709	0.01	0.00	1.79	0.00708	0.01	0.00	1.79
Nombre d'équipements dans l'enseignement au 2nd degré	0.00237	0.01	0.00	0.30	0.00236	0.01	0.00	0.30
Nombre d'équipements dans l'enseignement supérieur et autres	0.00327	0.01	0.00	0.31	0.00326	0.01	0.00	0.31
Nombre d'équipements dans les services médicaux et paramédicaux	0.08203	0.18	0.00	10.71	0.08198	0.18	0.00	10.71
Nombre d'équipements dans les services de santé	0.00984	0.02	0.00	2.34	0.00982	0.02	0.00	2.34
Nombre d'équipements dans les services aux particuliers	0.33946	1.04	0.00	251.57	0.33919	1.04	0.00	251.57
Nombre d'équipements dans le sport et la culture	0.02426	0.11	0.00	62.89	0.02424	0.11	0.00	62.89
Nombre d'équipements dans le transport et le tourisme	0.04085	0.17	0.00	34.88	0.04078	0.17	0.00	34.88
Nombre d'établissements actifs	1.40070	14.81	1.00	7544.00	1.40593	14.97	1.00	7544.00
Total des charges financières	121.35646	15 484.19	-2.00	59 422 233.00	123.83675	15 652.99	-2.00	59 422 233.00
Emprunts et dettes assimilées	1042.67799	98 684.62	0.00	39 313 686.00	1062.55941	99 760.09	0.00	39 313 686.00
Investissement corporel et incorporel avec apports	221.84862	22 700.06	-1187.00	11 840 107.00	225.65497	22 946.27	-1187.00	11 840 107.00
Âge de l'entreprise en 2005	12.73848	10.46	0.00	104.97	12.67348	10.45	0.00	104.97
N	502 243				491 468			

TABLE 3.A.6 – Variables quantitatives 2005-2007 (entreprises fermées)

	2005					2006					2007					
	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max
Effectif de l'entreprise	2.19359	14.99	0.00	2041.00	1.97068	12.04	0.00	965.00	2.16514	14.80	0.00	1350.00				
Chiffre d'affaires	281.82059	3095.13	-40.00	286802.00	319.67002	5455.16	-89.00	691.411.00	231.80484	2386.80	-7.00	213556.00				
Bénéfice attribué ou perte transférée	0.31513	36.35	-1.00	7835.00	0.23564	16.86	-7.00	2282.00	0.13662	20.35	0.00	4319.00				
Perte supportée ou bénéfice transféré	0.06146	5.23	0.00	921.00	0.08613	5.83	-3.00	1032.00	0.07350	6.63	0.00	1017.00				
Impôt sur les bénéfices	3.48748	87.70	-2247.00	10494.00	4.05885	201.05	-11347.00	30059.00	6.37148	761.58	-1057.00	163453.00				
Valeur ajoutée hors taxe	78.30790	649.07	-2600.00	67764.00	85.73815	1043.60	-9879.00	144406.00	69.87598	667.76	-10849.00	53992.00				
Taux de dépendance économique	89.95855	11.76	70.51	130.18	90.05338	11.82	70.51	130.18	90.44700	11.66	70.51	130.18				
Taux de chômage	8.3521	1.85	4.00	16.20	8.39432	1.86	4.00	16.20	8.35698	1.87	4.00	16.20				
Densité de la population	3680.94426	7154.54	0.94	40168.58	3614.78954	7045.81	1.34	40168.58	3271.58602	6626.39	0.94	40168.58				
Nombre d'équipements en action sociale	0.00966	0.03	0.00	3.51	0.00941	0.03	0.00	3.51	0.00943	0.03	0.00	1.65				
Nombre d'équipements dans le commerce	0.12246	0.46	0.00	7.91	0.11784	0.44	0.00	7.91	0.10910	0.42	0.00	7.91				
Nombre d'équipements dans l'enseignement au 1er degré	0.00715	0.01	0.00	1.79	0.00715	0.01	0.00	0.60	0.00728	0.02	0.00	1.79				
Nombre d'équipements dans l'enseignement au 2nd degré	0.00239	0.01	0.00	0.30	0.00237	0.01	0.00	0.30	0.00225	0.01	0.00	0.30				
Nombre d'équipements dans l'enseignement supérieur et autres	0.00312	0.01	0.00	0.14	0.00300	0.01	0.00	0.31	0.00271	0.01	0.00	0.14				
Nombre d'équipements dans les services médicaux et paramédicaux	0.08100	0.18	0.00	10.71	0.07894	0.17	0.00	5.00	0.07549	0.18	0.00	10.71				
Nombre d'équipements dans les services de santé	0.00980	0.02	0.00	2.34	0.00966	0.02	0.00	2.34	0.00935	0.02	0.00	1.79				
Nombre d'équipements dans les services aux particuliers	0.33419	0.97	0.00	58.14	0.32406	0.91	0.00	16.12	0.30372	0.92	0.00	58.14				
Nombre d'équipements dans le sport et la culture	0.02376	0.05	0.00	5.33	0.02409	0.05	0.00	4.69	0.02407	0.05	0.00	6.21				
Nombre d'équipements dans le transport et le tourisme	0.04156	0.20	0.00	34.88	0.03935	0.13	0.00	10.03	0.03639	0.20	0.00	34.88				
Nombre d'établissements actifs	1.07850	0.67	1.00	67.00	1.08761	0.89	1.00	164.00	1.09107	0.90	1.00	106.00				
Total des charges financières	7.66568	612.00	0.00	135870.00	19.49639	851.80	0.00	105851.00	10.72399	319.57	0.00	35170.00				
Emprunts et dettes assimilées	63.71560	2289.74	0.00	373888.00	152.12903	5895.02	0.00	948606.00	132.80635	4788.16	0.00	684392.00				
Investissement corporel et incorporel avec apports	11.66890	383.19	-658.00	50833.00	21.84536	1268.86	-160.00	229368.00	16.35994	322.43	-227.00	37870.00				
Âge de l'entreprise en 2005	13.83077	11.59	0.00	101.06	13.64108	11.40	0.00	97.07	13.67847	11.41	0.00	95.98				
N	53530				57507				46325							

TABLE 3.A.7 – Variables quantitatives 2008-2010 (entreprises fermées)

	2008					2009					2010				
	moyenne	écart type	min	max		moyenne	écart type	min	max		moyenne	écart type	min	max	
Effectif de l'entreprise	5.32401	28.08	0.00	1413.00		6.51565	41.81	0.00	1623.00		5.38873	73.79	0.00	9149.00	
Chiffre d'affaires	726.50451	6803.07	-58.04	605917.50		792.23382	7523.98	-1182.18	542473.38		655.16574	7578.95	-932.16	584903.13	
Bénéfice attribué ou perte transférée	0.39433	25.26	-2.01	2846.15		0.17443	9.13	-0.01	742.67		0.05963	3.29	-7.22	362.78	
Perte supportée ou bénéfice transféré	0.53200	51.88	-68.37	6495.71		0.06859	3.17	0.00	219.61		0.22941	20.25	0.00	2598.47	
Impôt sur les bénéfices	8.27155	140.04	-747.65	11754.08		8.84735	180.89	-2214.03	12663.32		5.14917	91.10	-4058.59	6117.48	
Valeur ajoutée hors taxe	205.58992	1481.14	-2112.59	116898.08		219.45322	1812.06	-8745.86	90354.58		181.74286	1712.08	-21792.90	126546.54	
Taux de dépendance économique	90.76549	11.42	70.51	130.18		90.14626	11.51	70.51	130.18		90.23683	11.62	70.51	130.18	
Taux de chômage	8.31452	1.89	4.00	16.20		8.34599	1.86	4.00	16.20		8.31603	1.85	4.00	16.20	
Densité de la population	3019.35962	6257.95	1.46	40168.58		3336.92305	6676.55	0.94	40168.58		3344.85130	6703.75	0.44	40168.58	
Nombre d'équipements en action sociale	0.00932	0.02	0.00	1.65		0.00921	0.02	0.00	0.90		0.00949	0.02	0.00	1.55	
Nombre d'équipements dans le commerce	0.10376	0.41	0.00	7.91		0.10567	0.39	0.00	5.59		0.11418	0.43	0.00	7.91	
Nombre d'équipements dans l'enseignement au 1er degré	0.00692	0.01	0.00	0.47		0.00694	0.01	0.00	0.45		0.00703	0.02	0.00	1.55	
Nombre d'équipements dans l'enseignement au 2nd degré	0.00236	0.01	0.00	0.30		0.00220	0.01	0.00	0.22		0.00234	0.01	0.00	0.30	
Nombre d'équipements dans l'enseignement supérieur et autres	0.00252	0.01	0.00	0.13		0.00267	0.01	0.00	0.13		0.00278	0.01	0.00	0.13	
Nombre d'équipements dans les services médicaux et paramédicaux	0.07306	0.18	0.00	7.72		0.07315	0.15	0.00	5.00		0.07729	0.17	0.00	5.00	
Nombre d'équipements dans les services de santé	0.00936	0.03	0.00	1.34		0.00914	0.02	0.00	0.60		0.00966	0.02	0.00	1.55	
Nombre d'équipements dans les services aux particuliers	0.28464	0.83	0.00	16.12		0.29402	0.79	0.00	10.81		0.30829	0.88	0.00	35.71	
Nombre d'équipements dans le sport et la culture	0.02416	0.05	0.00	2.43		0.02312	0.04	0.00	1.19		0.02367	0.06	0.00	6.21	
Nombre d'équipements dans le transport et le tourisme	0.03284	0.11	0.00	1.82		0.03555	0.11	0.00	4.40		0.03684	0.12	0.00	3.11	
Nombre d'établissements actifs	1.13501	0.80	1.00	46.00		1.21453	2.73	1.00	163.00		1.16697	1.60	1.00	153.00	
Total des charges financières	19.01994	853.84	0.00	90973.00		14.89313	301.03	0.00	24462.00		15.83833	610.78	0.00	78333.00	
Emprunts et dettes assimilées	192.18032	7357.15	0.00	872507.00		174.87494	3808.09	0.00	347670.00		142.75371	2688.91	0.00	257356.00	
Investissement corporel et incorporel avec apports	54.34726	2557.37	-162.00	314108.00		45.29683	1015.09	-196.00	107526.00		64.41237	3552.17	-78.00	460039.00	
Âge de l'entreprise en 2005	15.07435	11.54	0.00	73.62		13.99228	11.35	0.00	77.05		14.06614	11.12	0.00	99.90	
N	16 132			17 704		18 105									

TABLE 3.A.8 – Variables quantitatives 2011-2013 (entreprises fermées)

	2011					2012					2013					
	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max	moyenne	écart type	min	max
Effectif de l'entreprise	4.31858	24.17	0.00	1221.00	4.23766	28.37	0.00	1729.00	4.14572	26.09	0.00	1732.00	4.14572	26.09	0.00	1732.00
Chiffre d'affaires	542.92865	5950.04	-284.64	613923.50	1275.10287	30202.56	-117.60	1819808.50	626.57986	8086.05	-208.55	466115.25	626.57986	8086.05	-208.55	466115.25
Bénéfice attribué ou perte transférée	0.14442	11.56	-0.01	1363.75	0.48778	42.00	0.00	5123.76	0.11871	6.06	-32.45	624.22	0.11871	6.06	-32.45	624.22
Perte supportée ou bénéfice transféré	0.03568	1.79	-0.06	176.04	0.09110	3.85	0.00	275.23	0.79647	47.46	-1.42	5078.91	0.79647	47.46	-1.42	5078.91
Impôt sur les bénéfices	4.45394	94.40	-2997.96	9835.16	10.92270	321.59	-1763.04	23446.09	5.57894	173.36	-2241.66	19226.02	5.57894	173.36	-2241.66	19226.02
Valeur ajoutée hors taxe	158.23127	1190.36	-16816.56	70298.90	200.15122	2143.19	-25673.61	114179.87	191.86695	3866.14	-6852.04	404475.78	191.86695	3866.14	-6852.04	404475.78
Taux de dépendance économique	90.19915	11.56	70.51	130.18	90.28588	11.53	70.51	130.18	90.13546	11.51	70.51	130.18	90.13546	11.51	70.51	130.18
Taux de chômage	8.32711	1.88	4.00	16.20	8.36385	1.90	4.00	16.20	8.33176	1.88	4.00	16.20	8.33176	1.88	4.00	16.20
Densité de la population	3339.38744	6681.87	1.05	40168.58	3251.89482	6615.60	1.65	40168.58	3259.45089	6608.30	1.10	40168.58	3259.45089	6608.30	1.10	40168.58
Nombre d'équipements en action sociale	0.00966	0.03	0.00	3.51	0.00965	0.03	0.00	1.65	0.00941	0.02	0.00	1.55	0.00941	0.02	0.00	1.55
Nombre d'équipements dans le commerce	0.11377	0.44	0.00	5.59	0.11717	0.46	0.00	7.91	0.11383	0.44	0.00	7.91	0.11383	0.44	0.00	7.91
Nombre d'équipements dans l'enseignement au 1er degré	0.00703	0.01	0.00	0.36	0.00728	0.02	0.00	1.79	0.00741	0.02	0.00	1.55	0.00741	0.02	0.00	1.55
Nombre d'équipements dans l'enseignement au 2nd degré	0.00238	0.01	0.00	0.27	0.00237	0.01	0.00	0.30	0.00233	0.01	0.00	0.30	0.00233	0.01	0.00	0.30
Nombre d'équipements dans l'enseignement supérieur et autres	0.00288	0.01	0.00	0.13	0.00281	0.01	0.00	0.21	0.00283	0.01	0.00	0.13	0.00283	0.01	0.00	0.13
Nombre d'équipements dans les services médicaux et paramédicaux	0.07643	0.16	0.00	2.32	0.07798	0.19	0.00	10.71	0.07716	0.17	0.00	5.00	0.07716	0.17	0.00	5.00
Nombre d'équipements dans les services de santé	0.00965	0.03	0.00	2.34	0.00978	0.03	0.00	1.79	0.00957	0.02	0.00	1.55	0.00957	0.02	0.00	1.55
Nombre d'équipements dans les services aux particuliers	0.30664	0.87	0.00	10.81	0.31441	0.95	0.00	35.71	0.31126	0.91	0.00	35.71	0.31126	0.91	0.00	35.71
Nombre d'équipements dans le sport et la culture	0.02450	0.06	0.00	4.69	0.02394	0.07	0.00	6.21	0.02542	0.08	0.00	6.21	0.02542	0.08	0.00	6.21
Nombre d'équipements dans le transport et le tourisme	0.03638	0.12	0.00	1.45	0.03750	0.12	0.00	3.11	0.03777	0.15	0.00	10.03	0.03777	0.15	0.00	10.03
Nombre d'établissements actifs	1.16249	1.54	1.00	137.00	1.15228	0.85	1.00	47.00	1.20552	4.50	1.00	517.00	1.20552	4.50	1.00	517.00
Total des charges financières	10.22935	326.69	-2.00	40548.00	12.96273	319.71	-1.00	27203.00	8.81453	153.48	0.00	11534.00	8.81453	153.48	0.00	11534.00
Emprunts et dettes assimilées	118.94394	1751.99	0.00	151750.00	161.11911	3176.43	0.00	228146.00	105.74384	963.34	0.00	45229.00	105.74384	963.34	0.00	45229.00
Investissement corporel et incorporel avec apports	26.15288	183.16	-15.00	11058.00	73.54903	3229.79	-8.00	331623.00	52.89214	2328.42	-20.00	272682.00	52.89214	2328.42	-20.00	272682.00
Âge de l'entreprise en 2005	13.92250	10.95	0.00	98.99	14.16592	10.91	0.00	71.62	14.53543	10.65	0.00	51.04	14.53543	10.65	0.00	51.04
N	16 856				15 859				14 315				14 315			

TABLE 3.A.9 – Variables quantitatives 2014-2015 (entreprises fermées)

	2014					2015				
	moyenne	écart type	min	max		moyenne	écart type	min	max	
Effectif de l'entreprise	3.47802	21.78	0.00	1710.00		1.34066	5.69	0.00	51.00	
Chiffre d'affaires	428.25975	3145.36	-497.92	192776.47		150.79240	240.27	0.00	1364.58	
Bénéfice attribué ou perte transférée	0.30468	27.72	-1.72	3174.86		0.01218	0.12	0.00	1.11	
Perte supportée ou bénéfice transféré	0.09989	38.62	-2643.51	3581.58		0.00003	0.00	0.00	0.00	
Impôt sur les bénéfices	2.68102	58.48	-1678.49	3779.27		0.80341	5.00	-0.74	44.57	
Valeur ajoutée hors taxe	130.05898	820.96	-3443.99	41682.51		66.80928	122.66	-60.52	611.86	
Taux de dépendance économique	90.26623	11.58	70.51	130.18		89.12209	11.50	72.18	128.43	
Taux de chômage	8.34619	1.89	4.00	16.20		8.36154	1.86	5.20	13.70	
Densité de la population	3259.21336	6628.21	0.80	40168.58		3614.94339	6803.14	20.97	29757.98	
Nombre d'équipements en action sociale	0.00921	0.02	0.00	0.80		0.00900	0.01	0.00	0.07	
Nombre d'équipements dans le commerce	0.10705	0.40	0.00	6.97		0.17565	0.69	0.00	4.64	
Nombre d'équipements dans l'enseignement au 1er degré	0.00688	0.01	0.00	0.39		0.00783	0.01	0.00	0.06	
Nombre d'équipements dans l'enseignement au 2nd degré	0.00227	0.01	0.00	0.27		0.00242	0.01	0.00	0.04	
Nombre d'équipements dans l'enseignement supérieur et autres	0.00271	0.01	0.00	0.13		0.00290	0.01	0.00	0.05	
Nombre d'équipements dans les services médicaux et paramédicaux	0.07464	0.16	0.00	5.00		0.09493	0.18	0.00	0.98	
Nombre d'équipements dans les services de santé	0.00921	0.02	0.00	0.61		0.01098	0.02	0.00	0.13	
Nombre d'équipements dans les services aux particuliers	0.29755	0.80	0.00	13.17		0.39868	1.08	0.01	6.89	
Nombre d'équipements dans le sport et la culture	0.02398	0.04	0.00	1.82		0.03038	0.06	0.00	0.54	
Nombre d'équipements dans le transport et le tourisme	0.03640	0.14	0.00	10.03		0.05356	0.18	0.00	1.14	
Nombre d'établissements actifs	1.14864	1.24	1.00	118.00		1.10989	0.59	1.00	6.00	
Total des charges financières	8.14579	112.86	0.00	8754.00		4.79121	22.50	0.00	214.00	
Emprunts et dettes assimilées	128.36026	1728.34	0.00	120057.00		57.30769	112.39	0.00	651.00	
Investissement corporel et incorporel avec apports	44.89968	1481.55	-198.00	166356.00		15.47253	52.62	0.00	481.00	
Âge de l'entreprise en 2005	14.63017	10.55	0.00	76.14		10.55905	9.19	0.00	41.03	
N	13 307					91				

3.B Ratios de risques

TABLE 3.B.1 – Ratios de risques modèle sans stratification

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Dépendance économique	1.004*** (0.000202)		1.001*** (0.000202)	1.000 (0.000203)
Chômage	1.000 (0.00103)		0.998 (0.00103)	0.999 (0.00103)
Densité de population	1.000*** (0.000000372)		1.000*** (0.000000372)	1.000*** (0.000000371)
Équipements en action sociale	0.930 (0.101)		1.020 (0.110)	1.066 (0.114)
Équipements dans le commerce	0.954*** (0.0136)		0.941*** (0.0132)	0.916*** (0.0133)
Équipements dans l'enseignement au 1er degré	2.506*** (0.437)		2.106*** (0.354)	1.335 (0.237)
Équipements dans l'enseignement au 2nd degré	5.294*** (2.104)		2.981** (1.186)	2.977** (1.180)
Équipements dans l'enseignement supérieur et autres	0.154*** (0.0463)		0.143*** (0.0425)	0.107*** (0.0316)
Équipements dans les services médicaux et paramédicaux	0.937* (0.0271)		0.900*** (0.0258)	0.973 (0.0278)
Équipements dans les services de santé	0.903 (0.159)		1.014 (0.172)	1.053 (0.178)
Équipements dans les services aux particuliers	1.018* (0.00893)		1.046*** (0.00902)	1.061*** (0.00957)
Équipements dans le sport et la culture	0.897* (0.0402)		0.811*** (0.0342)	0.746*** (0.0364)
Équipements dans le transport et le tourisme	1.014 (0.0180)		0.981 (0.0168)	0.955* (0.0169)
Effectif		0.968*** (0.000508)	0.968*** (0.000508)	0.968*** (0.000521)
Chiffre d'affaires		1.000*** (0.00000277)	1.000*** (0.00000277)	1.000*** (0.00000291)
Bénéfice attribué ou perte transférée		0.999*** (0.000131)	0.999*** (0.000137)	1.000** (0.000139)
Perte supportée ou bénéfice transféré		1.000 (0.000136)	1.000 (0.000140)	1.000 (0.000117)
Impôt sur les bénéfices		1.000*** (0.00000157)	1.000*** (0.00000158)	1.000*** (0.00000207)
Valeur ajoutée hors taxe		1.000*** (0.000000705)	1.000*** (0.000000706)	1.000*** (0.000000746)
Nombre d'établissements actifs		0.966*** (0.00409)	0.966*** (0.00409)	0.972*** (0.00410)
Total des charges financières		1.000 (0.00000343)	1.000 (0.00000349)	1.000 (0.00000340)
Emprunts et dettes assimilées		1.000*** (0.000000505)	1.000*** (0.000000514)	1.000*** (0.000000501)
Investissement corporel et incorporel		1.000*** (0.00000150)	1.000*** (0.00000149)	1.000*** (0.00000135)
Âge de l'entreprise en 2005		1.016*** (0.000179)	1.016*** (0.000180)	1.016*** (0.000179)
Industrie manufacturière				1.211*** (0.00910)
Secteur de la construction				1.327*** (0.00883)
Secteur commerce, transport, hébergement et restauration				1.484*** (0.00799)
Secteur activités spécialisées, scientifiques techniques, services administratifs et soutien				0.974** (0.00773)
N	6 434 521	6 434 521	6 434 521	6 434 521

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 4

Competition effect in spatial concentration of firms: A spatial survival model approach

Introduction

Space is not a neutral entity for firms, and it has multilevel impacts on their activities. Characteristics of spatial environment determine firms' location choice and affect their life cycle. In order to benefit from favorable conditions, firms are sometimes strikingly concentrated in certain locations, while other areas seem to be less fertile for business creation because they appear starved of environmental amenities. Spatial concentration of economic activities spawns some positive externalities that attract new businesses but could also be a source of tough competition whose effects dominate agglomeration economies, leading then to higher hazard rates for market failure. [Dumais et al. \(2002\)](#) rightly pointed out the dynamic aspect of geographic concentration, stating: "Geographic concentration is the outcome of a life cycle process in which new plants are constantly being born, existing plants are expanding and contracting at different rates, and a substantial number of businesses are failing".

Although the question of firm survival chance in densely populated areas has not often received primary attention in the literature, some papers still shed a little light on the subject. [Brixy & Grotz \(2007\)](#) and [Huiban \(2011\)](#) found that firms are more likely to fail in densely populated areas. This pattern, interpreted as the consequence of competition, seems to prevail in our data since we have consistently found in the first three chapters that firm survival is negatively affected by the population density. However, one limitation of the previous chapters is that firms are observed in a quite large areas (municipalities and employment areas). This does not allow to accurately investigate the competition effect for at least two reasons. Firstly, with the population density calculated at municipalities level, we cannot fully rely on microeconomic foundations when discussing the competition. Secondly, the competition that firms face in certain municipalities, probably emanates from their close proximity. Thus, observing things at an aggregate level might not be sufficient to capture the mechanism. We then think that working on microdata can be a good alternative to get a better view and tackle the problem.

Competition is a type of interaction, and according to [Tobler \(1970\)](#): «everything is re-

lated to everything else, but near things are more related than distant things». Close firms are then expected to interact more, whether is for commercial relationship or for trying to drop competitor out of the market. A firm will tend to be positively affected by its business partner's survival and negatively affected by its competitor's survival (Lang & Stulz, 1992). In either case, it makes sense to assume and model spatial dependence when analyzing the survival of neighboring firms.

In survival analysis, data, especially those involving human subjects, sometimes contain unobserved heterogeneities that might produce bias to parameter estimates. Moreover, these unobserved heterogeneities can be spatially correlated. To handle this, spatial frailty models have been introduced (see Zhou & Hanson, 2015; Banerjee, 2016). Nevertheless, the "frailty" term, that was primarily designed for medical study, may not be suitable to account for spatial dependence in certain types of study such as firm survival analysis, and what is more, it can be quite hard to find theoretical underpinnings for that in economics.

Spatial correlation in medical data does not necessarily mean that subjects (patients) are interacting with neighbors, it could come from the fact they share a specificity associated with a common spatial environment. In firm survival analysis case, spatial dependence may, for instance, stem from the fact that two neighboring firms have a customer/supplier relationship, or conversely, two neighboring firms are competing against each other, and trying through different strategies evict each other from the market. To account for this aspect, spatial probit will be frequently used in the literature (Craioveanu & Terrell, 2016; Maté-Sánchez-Val et al., 2018; Vivel-Búa & Lado-Sestayo, 2021). However, the problem with using probit or logit models on survival data, is that these models do not account for the time to event and censoring (Kleinbaum & Klein, 2010; Liu, 2012). Indeed, even though the outcome variable described in survival analysis seem similar to the one studied when using probit or logit model, there are significant differences that should not be overlooked. In a survival analysis, one is interested in the: time to an event. The duration of the process, the occurrence of the event and the censorship are all considered in the analysis. But on the flip side, probit or logit models are designed to carry out analysis of situations with dichotomous outcome without taking neither the timing of the occurrence, nor the censorship into account. These methods would then treat equally 2 individuals who experience the studied event, even though there is a huge difference in the timing. Thus, applying these statistical approaches on survival data can generate misleading conclusions. It is therefore important to think about a new way to account for spatial dependencies in firm survival analysis using survival models. In this train of thought, this chapter proposes a spatial Accelerated Failure Time (AFT) model using bayesian approach. This leads us

to a posterior distribution from which one can sample using a Random Walk Metropolis-Hastings Within Gibbs algorithm.

The contribution of this chapter is twofold. First, it extends the other chapters by analyzing the interfirm competition using microdata, thus providing a framework that allows reference to microeconomic foundations of competition. Secondly, the chapter proposes a new way to account for spatial dependencies in survival analysis. We then apply the model on 300 manufacturing plants located in Île-de-France region and find that firm survival time is negatively affected by neighbors' survival time.

The rest of the chapter is structured as follows. Section 4.1 reviews the literature on the relationship between competition and firm survival. Section 4.2 discusses spatial dependence in survival analysis. Section 4.3 presents the model and the likelihood. Section 4.4 briefly describes the Bayesian approach. Sections 4.5 and 4.6 present data, summary statistics and results.

4.1 Competition and firm survival

Profit maximization is one of the first goals of every established firm. Reaching this goal depends not only on the firm internal resources, but also imply for the firms to deal with market conditions. Except for a case of monopoly where there is only one producer in the market, or a perfectly competitive paradigm where the firm's decisions have no impact on the market (Belleflamme & Peitz, 2010); a firm usually has to anticipate competitors' behavior and take that into account when elaborating its strategies which can range from setting products price or output level, to the decision to innovate. This situation results in permanent interaction between firms which is not without consequences on their life duration in the market.

In industrial organization, the competition game is sometimes considered as a kind of sorting process that drop non-efficient firms out of the market. In this regard, some theoretical works on imperfect markets have established interesting results that remain robust to the model of competition whether it is a price-competition (Bertrand model), a output level – competition (Cournot model), or a monopolistic competition where each firms produce a single variety of a differentiated product (Chamberlin, 1933). Those results are well described in Belleflamme & Peitz (2010) and overall point out the fact that only efficient firms would be able to survive in a competitive market. Indeed, in a price-competition setting, it has been demonstrated that when two firms set the same price, the most efficient

one ends up cornering the market. Likewise, in a Cournot competition, a firm profit grows as its marginal cost decreases or when its rival's marginal cost increases. It was also established that in monopolistically competitive growing markets, only efficient firms would be able to stay. On the empirical literature side, findings are quite mixed. While some studies show that competition increases the hazard rate for business failure, others demonstrate the opposite.

[Agarwal & Gort \(1996\)](#) in their study of entry, exit and firm survival, considered different stages in the evolution of the product market. Those stages are defined based on the change in the number of firms in the market. Based on their results, they concluded that the intensity of competition counts for a lot in exiting the market. This finding is corroborated in another paper ([Agarwal & Gort, 2002](#)) from same authors. Indeed, they found that market maturity and increased competitiveness are associated with higher hazard rate for firm failure. This positive impact of the competition on firms' exit probability is quite often evidenced in the literature.

[Buehler et al. \(2004\)](#) use a natural experiment approach on Swiss data and find that the tightening of antitrust law increases the exit hazard rates. [Mengistae \(2006\)](#) worked on ethiopian manufacturing firms and found that when the competition pressure get stronger in the industry, firms face a higher hazard for failure. [Brown & Earle \(2000\)](#) use a panel of Russian industrial firms to study the effect of market competition on firm efficiency. They find that firm survival is negatively affected by competition proxied by market concentration indices. Their results also show that there are lower survival chances for firms located in region with good quality transportation infrastructure due to the high intensity competition there are exposed to. [Van Kranenburg et al. \(2002\)](#) studied firm survival in daily newspaper industry in the Netherlands. Their results show that exit chances are positively impacted by the intensity of the competition proxied by the number of newspapers. Regarding this finding, authors conclude that as competition intensifies, more efficient firms survive, and inefficient ones are driven out of the industry. However, some studies come up with different conclusions about the link between competition and firm survival.

[Audretsch & Mahmood \(1995\)](#) found that price-cost margin has a negative effect on firm survival. Price-cost margin is a proxy of the Lerner index (market power) used to measure competition, and it is as high as the competition is low. [Audretsch & Mahmood \(1995\)](#)'s result means that less intense competition leads to an increase in the hazard rate. [Resende et al. \(2016\)](#) found that Herfindahl index reduces firm survival chances. The Herfindahl index measures the concentration and can be used to capture the market power, it is also pro-

portional to Lerner index (Belleflamme & Peitz, 2010). Thus, Resende et al. (2016)'s results are in line with Audretsch & Mahmood (1995)'s work. Nonetheless, measuring competition using the price-cost margin or the Herfindahl index raises some questions that could create confusion in interpreting results of Audretsch & Mahmood (1995) and Resende et al. (2016). Higher values of the price-cost margin and the Herfindahl index are empirically interpreted as being less competitive, but this is not always the case from theoretical perspective. Boone (2008) has demonstrated that more intense competition can sometimes lead to higher price-cost margin and/or Herfindahl index. Thus, conclusions from Audretsch & Mahmood (1995) and Resende et al. (2016), may not be in opposition to studies establishing a negative relationship between the intensity of competition and firm survival.

These mixed empirical results do not necessarily deviate from the theoretical conclusions on competition. Indeed, empirical studies on samples composed mainly of inefficient firms will tend to find a negative effect of competition on survival, whereas studies on samples of efficient firms will capture more of the positive effect of competition. In any case, competition does not affect all firms equally since efficient firms will survive, gain more market share, and thrive, while inefficient ones will exit the market. This is therefore a type of interaction where the success of some is equivalent to the failure of others, and this deserves to be modeled in certain contexts (agglomerations) where spatial proximity exacerbates competition (Baum & Mezas, 1992).

4.2 Spatial dependence in survival analysis

In survival analysis, "frailty" is used to represent individual proneness to experience the studied event (Vaupel et al., 1979). The frailty term is set to quantify unobserved heterogeneities such as genetic predisposition, physiological senescence parameters, economic capability, and so forth (Liu, 2012). The main idea consists in introducing a random component into the hazard function to account for heterogeneities among subjects (see Hosmer et al., 2008; Liu, 2012, for further details). Survival data can sometimes be spatially correlated due to non-measurable location-specific similarities. In those cases, spatial dependencies are handled through particular forms of frailties called "spatial frailties".

Spatial frailty models extend traditional frailty models by adding a random effect that incorporates spatial dependence. The modelling of spatially correlated frailties is customized to data, whether it is point referenced or areally referenced data (Zhou & Hanson, 2015; Banerjee, 2016). Let's consider the following survival model formulated using the propor-

tional hazards specification: $h(t_i, X_i) = h_0(t_i) \exp(X_i' \beta)$, the accelerated failure time specification would give: $S(t_i, X_i) = S_0(t_i \exp(-X_i' \beta))$. S_0 and h_0 are baseline survival function and baseline hazard function respectively, t_i refers to the survival time (or censoring time) of the i th subject. The matrix X contains covariates while β is the vector of parameters. To account for spatial dependencies through frailties, the model is extended by replacing the term $X_i' \beta$ with $X_i' \beta + \gamma_i$ where γ_i is a random term corresponding to the location specific frailty. When data are point referenced, spatial frailties $\Gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ can be modeled to follow a multivariate Gaussian distribution $N_n(0, \Sigma_\Gamma)$. Σ_Γ is constructed using an isotropic correlation function where the correlation between γ_i and γ_j only depends on the distance between their location (Banerjee et al., 2003; Zhou & Hanson, 2015; Bastos & Gamerman, 2006). In the case of areal data, spatial frailties can still be modeled using a multivariate gaussian distribution with a covariance matrix constructed using distance between centroids. Another approach is to use conditionally autoregressive structure for frailties (Banerjee et al., 2003; Zhou & Hanson, 2015). These models are estimated through bayesian approach and are widely relied upon in biostatistics literature. By way of example, we might cite the works of Henderson et al. (2002); Li & Ryan (2002); Bastos & Gamerman (2006) and Thamrin et al. (2021). Henderson et al. (2002) worked on survival of adult acute myeloid leukemia patients in northwest England. To investigate whether there is some spatial between-patient heterogeneities, they developed a model that allows spatial dependence which reduces to standard gamma frailty model if there is no spatial variation. Their results show that there are concentrated regions with high and low risk. Li & Ryan (2002) studied risk factors associated with childhood asthma using a semiparametric frailty model they proposed. To account for potential spatial correlation, they introduced spatial dependencies in ordinary frailty models. More precisely, they allow for random effect (mean-zero stationary Gaussian process) accommodating general correlation structure to enter into the baseline hazard. Bastos & Gamerman (2006) considered spatial frailty while allowed coefficients to be time-varying. Their model includes a spatially structured frailty term following Gaussian distribution. Thamrin et al. (2021) also used spatial survival (conditional autoregressive frailty) models on medical data.

When considering event like firm survival, spatial dependencies do not necessarily arise from similarities in social or ecological environments (of course, this may be the case). Neighboring firms interact with each other whether is for business relation or competition rivalry. In either case, a firm can experience a contagion effect when its business partner or its competitor failed. Lang & Stulz (1992) studied the contagion effect of bankruptcy and found positive and negative effect of bankruptcy announcements on the bankrupt firm's

competitors. Their results show that in highly leveraged industries, firms are negatively affected by their competitors' bankruptcy while in more concentrated industries (Herfindal index), firms benefit from their competitor's bankruptcy by experiencing a wealth gain. Those contagion effects could sometimes extend beyond industry competitors along the supply chain; [Hertzel et al. \(2008\)](#)'s results shed light on that.

More recently, spatial econometrics methods will be used to handle spatial interaction in firm survival analysis. [Cainelli et al. \(2014\)](#) assessed firm exit rate sensibility to spatial agglomeration in Italy. They worked on aggregated data (Industry-province level) and controlled for spatial spillovers effect through the estimation of a spatial dynamic panel model via a Generalized Method of Moments estimator. Their results showed an evidence (even though it was weak) of spatial spillovers effects. They found that firm exit rates in a province are negatively affected by lag exit rates in neighbor provinces. [Arcuri et al. \(2019\)](#) worked on the same kind of data (departmental level) and used the same econometric strategy. They also find a significant spatial spillovers effects. However, contrary to [Cainelli et al. \(2014\)](#), they found that firm exit rates in a certain department are positively influenced by exit rates in neighbour departments. They explained this contagion effect as the consequence of inter-company relationships (trade, debt, outsourcing). [Craioveanu & Terrell \(2016\)](#) account for spatial interaction in firm survival by using a [Smith & LeSage \(2004\)](#) version of spatial probit model (spatial dependencies in the model disturbances). They find evidence for the presence of spatial component. [Maté-Sánchez-Val et al. \(2018\)](#) use a spatial probit ([LeSage & Pace, 2009](#)) to evaluate how business failure in Madrid metropolitan area is affected by the geographical location of external economic agent. Results show evidence for spatial correlation in probabilities of business failure (failed companies are likely to be surrounded by other failed companies). [Vivel-Búa & Lado-Sestayo \(2021\)](#) also use spatial probit on Spanish hotel sector, and find significant contagion effect on business failure. Their results show that the probability of failure of a hotel is significantly influenced by the failure of neighboring hotels.

Although spatial probit may, to some extent, be suitable to model spatial interactions in survival analysis, survival models, due to their capability to account for the time to event and censoring, are preferred over discrete choice models ([Kleinbaum & Klein, 2010](#); [Liu, 2012](#)). In what follows, we propose to use a survival model (Exponential AFT) and account for spatial dependencies in a different way than using classical spatial frailty modelling. This model will be useful to capture the competition effect that we have found evidence of in previous chapters' results.

4.3 Model and Likelihood

Consider that we are interested in studying the survival time of n individuals, each having a precise location. Let T_0 be the baseline survival time for the n individuals.

Assumption A_1 : T_0 follows a standard exponential distribution. The density (let f_0), the cumulative distribution (F_0) function and the survival function (S_0) are respectively ($\forall t > 0$)

$$\begin{cases} f_0(t) = \exp(-t) \\ F_0(t) = 1 - \exp(-t) \\ S_0 = \text{Prob}(T_0 > t) = 1 - F_0(t) = \exp(-t) \end{cases}$$

Let $\gamma = (\gamma_1 \ \gamma_2 \ \dots \ \gamma_k)'$ be a $K \times 1$ vector of parameters and $X = (X_{1s} \ X_{2s} \ \dots \ X_{ns})'$ a $n \times K$ matrix where $X_{is} = (X_{i1} \ X_{i2} \ \dots \ X_{ik})$ are $1 \times K$ vectors.

The AFT model assumes the survival function conditional on X to be $S(t/X_i) = S_0(t \exp(X_i \gamma))$ (Klein et al., 2016).

Let's consider a $n \times 1$ random vector $\varepsilon = (\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n)'$

Assumption A_2 : Each ε_i has the same distribution with $\log(T_0)$, and ε is independent of $X\gamma$

Given A_2 , the distribution function and the cumulative distribution of ε_i can easily be derived:

$$\begin{cases} f_\varepsilon(t) = \exp(t) f_0(\exp(t)) \\ F_\varepsilon(t) = F_0(\exp(t)) \end{cases}$$

Another interesting results are:

$$\begin{aligned} S_\varepsilon(t) &= 1 - F_\varepsilon(t) \\ &= 1 - F_0(\exp(t)) \end{aligned}$$

$$\begin{aligned}
S_\varepsilon(\log(t)) &= 1 - F_0(\exp(\log(t))) \\
&= S_0(t), \text{ then} \\
S_0(t \exp(X_i \gamma)) &= S_\varepsilon(\log(t \exp(X_i \gamma))) \\
&= S_\varepsilon(\log(t) - X_i \beta) \text{ with } \beta = -\gamma
\end{aligned}$$

$$\Rightarrow S(t/X_i) = S_\varepsilon(\log(t) - X_i \beta) \quad (\text{E1})$$

What (E1) tells us, is that :

$$\text{Prob}(T_i > t/X_i) = \text{Prob}(\varepsilon_i > \log(t) - X_i \beta)$$

$$\Rightarrow \text{Prob}\{\log(T_i) > \log(t)/X_i\} = \text{Prob}\{\varepsilon_i > \log(t) - X_i \beta\} \quad (\text{E2})$$

Where T_i is the sequence of random variable that corresponds to the survival time of each individual. Setting $\log(T_i) = X_i \beta + \varepsilon_i$, (E2) can be established. Using matrices, we can write for all of n individuals

$$Y = X\beta + \varepsilon$$

$$\text{Where } Y = \left(\log(T_1) \quad \log(T_2) \quad . \quad . \quad . \quad \log(T_n) \right)'$$

As aforementioned, the goal of this chapter, is to investigate the competition effect we found evidence of in previous chapters. We assume this competition to take place among neighboring firms and expect the life duration of those neighboring firms to be spatially correlated. To model this reality, we depart from the traditional "frailty" approach because of the lack of economic theories in industrial organization supporting this idea. Moreover, we think that relying on a random component might not properly describe the situation since the spatial dependence in our case is supposed to arise from strategies consciously

set up by firms when interacting with competitors.

In a competitive market, when a firm elaborates its profit maximizing strategies, it takes into account possible reactions of its competitors, and vice versa. Thus, those same strategies that determine a firm's success potentially affect the success of the firm's competitors. This should result in spatial spillover effects in firm survival and we try to capture this through a spatial autoregressive structure.

The spatial autoregressive structure has already been used in the literature on firm survival, whether on aggregate data (Cainelli et al., 2014; Arcuri et al., 2019), or individual data (Maté-Sánchez-Val et al., 2018; Vivel-Búa & Lado-Sestayo, 2021). All these studies evinced the existence of spatial spillover effects, showing that inter-firm relationships are source of spatial pattern that should be considered when modeling firm survival.

To continue our modeling, we include a spatial lag in the previous equation :

Assumption A_3 : W is a row stochastic matrix whose eigenvalues are $\lambda_{i\{i=1,\dots,n\}}$ and $(I - \rho W)$ is a positive definite matrix. It implies that $\forall i \quad |\lambda_i| \leq 1$ (Gantmakher, 1959) and $\rho \in (-1, 1)$ (see Anselin, 1982; Kazar & Celik, 2012; LeSage & Pace, 2009)

$$\begin{aligned} Y &= \rho W Y + X\beta + \varepsilon \\ \Rightarrow Y &= (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \end{aligned} \quad (4.1)$$

I is the $n \times n$ identity matrix, W is a $n \times n$ weight matrix and ρ is a scalar.

Given assumption A_3 and equation (E2) , the survival function for the whole sample is :

$$\text{Prob}\{(I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon > \tau\} = \text{Prob}\{\varepsilon > (I - \rho W)\tau - X\beta\}$$

Where $\tau = (\log t_1 \quad \log t_2 \quad . \quad . \quad . \quad \log t_n)'$ is a $n \times 1$ vector

When it comes to write this for each individual, we have :

$$\text{Prob}\{\varepsilon_i > (I_i - \rho W_i)\tau - X_i\beta\} = S_\varepsilon\{(I_i - \rho W_i)\tau - X_i\beta\}$$

where I_i and W_i are respectively line vectors corresponding to the i^{th} line of I and W .

Based on the assumption A_1 , $\varepsilon_{i\{\forall i\}}$ follows a standard extreme value distribution for minimum, that is :

$$\Rightarrow \begin{cases} f_\varepsilon(t) = \exp(t) \exp[-\exp(t)] \\ F_\varepsilon(t) = 1 - \exp[-\exp(t)] \\ S_\varepsilon(t) = \exp[-\exp(t)] \end{cases}$$

$$\Rightarrow \begin{cases} f_\varepsilon\{(I_i - \rho W_i)\tau - X_i\beta\} = \exp\{(I_i - \rho W_i)\tau - X_i\beta\} \exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}] \\ F_\varepsilon\{(I_i - \rho W_i)\tau - X_i\beta\} = 1 - \exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}] \\ S_\varepsilon\{(I_i - \rho W_i)\tau - X_i\beta\} = \exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}] \end{cases}$$

We can now write the likelihood of our model. In survival model, the likelihood is specified as follows $L = \prod_{i=1}^n f^{\delta_i} S^{1-\delta_i}$ where $\delta_i = 1$ for individuals who have experienced the studied event and $\delta_i = 0$ for those who have not (see Liu, 2012). Let $L(\rho, \beta / Z)$ ¹ be the likelihood.

$$\begin{aligned} L(\rho, \beta / Z) &= \prod_{i=1}^n f_\varepsilon^{\delta_i}\{(I_i - \rho W_i)\tau - X_i\beta\} S_\varepsilon^{1-\delta_i}\{(I_i - \rho W_i)\tau - X_i\beta\} \\ &= \prod_{i=1}^n \{\exp\{(I_i - \rho W_i)\tau - X_i\beta\} \exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}]\}^{\delta_i} \\ &\quad \times \{\exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}]\}^{1-\delta_i} \\ &= \prod_{i=1}^n \{\exp[\delta_i \{(I_i - \rho W_i)\tau - X_i\beta\}]\} \{\exp[-\exp\{(I_i - \rho W_i)\tau - X_i\beta\}]\} \\ L(\rho, \beta / Z) &= \exp\left\{\sum_{i=1}^n \delta_i [(I_i - \rho W_i)\tau - X_i\beta]\right\} \exp\left\{-\sum_{i=1}^n \exp[(I_i - \rho W_i)\tau - X_i\beta]\right\} \end{aligned}$$

Likelihood being clearly specified, we can now move to the Bayesian approach.

1. Z is defined as $(W\tau, X)$

4.4 Bayesian Approach

Posterior

Unlike the frequentist Maximum Likelihood approach which is interested in estimating the mode of the likelihood and then relies on Central Limit Theorem to find its asymptotic distribution, Bayesian derives a posterior distribution for the parameters. That posterior which is derived using the likelihood and prior distribution, can then be used to compute some quantities like mean, variance and quantiles.

Let $\pi(\rho, \beta / Z)$, $\pi(\rho, \beta)$ be respectively joint posterior and prior:

$$\pi(\rho, \beta / Z) \propto L(\rho, \beta / Z) \pi(\rho, \beta) \quad (4.2)$$

Assumption A_4 : ρ and β are independent, so that $\pi(\rho, \beta) = \pi(\rho)\pi(\beta)$

The posterior can then be rewritten: $\pi(\rho, \beta / Z) \propto L(\rho, \beta / Z) \pi(\rho)\pi(\beta)$

Specifying Prior

The prior does not depend on data, it is supposed to summarize all we know about parameters (G. M. Koop, 2003). When nothing is known about parameters, or when our background does not allow us to be such as precise, we use non-informative prior, so that the inference is minimally affected (Syversveen, 1998; G. M. Koop, 2003; G. Koop et al., 2007; Greenberg, 2012).

The parameter ρ has been of great interest in spatial econometrics, and in the same way, $\pi(\rho)$ has been widely considered specially when analysing the SAR model using Bayesian approach (see Anselin, 2013; LeSage & Pace, 2009). Given the assumption A_3 , ρ is set to be uniformly distributed in $(-1, 1)$ ², thus

$$\pi(\rho) \propto C, \text{ where } C \text{ is a constant.} \quad (E3)$$

As for β , we assume our prior knowledge is not such sufficient to allow us to be precise when specifying $\pi(\beta)$. Then we are going to choose a non-informative prior for β .

2. $\rho \sim U(-1, 1)$

$\beta \sim N_k(0, V)$ where $V = 10^6 I_{k \times k}$

$$\Rightarrow \pi(\beta) \propto \exp \left[-\frac{1}{2} (\beta' V^{-1} \beta) \right] \quad (\text{E4})$$

Given (E3) and (E4), (4.2) becomes

$$\begin{aligned} \pi(\rho, \beta / Z) \propto & \exp \left\{ \sum_{i=1}^n \delta_i [(I_i - \rho W_i) \tau - X_i \beta] \right\} \exp \left\{ -\sum_{i=1}^n \exp [(I_i - \rho W_i) \tau - X_i \beta] \right\} \\ & \times \pi(\rho) \exp \left[-\frac{1}{2} (\beta' V^{-1} \beta) \right] \end{aligned} \quad (4.3)$$

Knowing the posterior, we can then be interested in computing quantities $E(g(\theta) / Z) = \int g(\theta) \pi(\theta / Z)$ with $\theta = (\rho, \beta)$, $g(\theta) = \theta$ for mean, and $g(\theta) = (\theta - E(\theta))^2$ for variance.

Bayesian computation

The posterior, as described, in the equation (4.3) does not have a usual form. Thus, it will be quite complicated to integrate it to produce summaries of parameters. To fix this issue, bayesian resort to simulation methods. Most frequently used methods are Markov chain Monte Carlo (MCMC), specially Gibbs sampling and the Metropolis-hastings algorithm. The idea is to take random draws from the posterior and use Monte Carlo integration to produce estimates of $E(g(\theta) / Z)$.

Monte Carlo Integration: Let $\theta^{r\{r=1, \dots, R\}}$ be a random sample from $\pi(\theta / Z)$ and $h(\cdot)$ be any function :

$$\lim_{R \rightarrow \infty} \left(\frac{1}{R} \sum_{r=1}^R h(\theta^r) \right) = E[h(\theta) / Z]$$

Gibbs Sampling

The Gibbs sampler is a basic MCMC method used for posterior simulation. It is useful when sampling from a multivariate posterior is not feasible but conditional distributions are available and sampling for them are feasible (Lynch, 2007).

Let $\pi(\theta_1, \dots, \theta_k)$ be nonstandard joint distribution that we want to sample from, and suppose we know full conditional distributions for each parameter, that is $\pi(\theta_i/\theta_{-i})$. The Gibbs sampler can be summarized in these three main steps:

1. Choose a starting values $\theta_2^{(0)}, \dots, \theta_k^{(0)}$
2. Draw
 - $\theta_1^{(1)}$ from $\pi(\theta_1/\theta_2^{(0)}, \dots, \theta_k^{(0)})$
 - $\theta_2^{(1)}$ from $\pi(\theta_2/\theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$
 - .
 - .
 - .
 - $\theta_k^{(1)}$ from $\pi(\theta_k/\theta_1^{(1)}, \dots, \theta_{k-1}^{(1)})$
3. At the R^{th} iteration, draw
 - $\theta_1^{(R)}$ from $\pi(\theta_1/\theta_2^{(R-1)}, \dots, \theta_k^{(R-1)})$
 - .
 - .
 - .
 - $\theta_k^{(R)}$ from $\pi(\theta_k/\theta_1^{(R)}, \dots, \theta_{k-1}^{(R)})$

Metropolis-Hastings algorithm

In contrast to Gibbs sampler, the Metropolis-Hastings Algorithm does not require conditional distributions to be available. Suppose we want to sample from $\pi(\theta)$ (θ is a vector). The Metropolis-Hastings Algorithm consist in generating a proposed value Θ^{s+1} given θ^s , from pre-specified proposal distribution let $q(\theta, \Theta)$. The proposed value Θ^{s+1} is then accepted as θ^{s+1} with the probability

$$p = \min \left\{ \frac{\pi(\Theta) q(\Theta, \theta)}{\pi(\theta) q(\theta, \Theta)}, 1 \right\}$$

If $q(\theta, \Theta) = q(\Theta, \theta)$ ³, then $p = \min \left\{ \frac{\pi(\Theta)}{\pi(\theta)}, 1 \right\}$. If Θ^{s+1} is rejected, the algorithm return to θ^s . The Metropolis-Hastings algorithm can then be summarized in these following steps:

1. Choose a starting value $\theta^{(0)}$
2. At the iteration $s + 1$, draw $\Theta^{(s+1)}$ from $q(\theta^{(s)}, \Theta)$
3. Generate u from $U(0, 1)$ and accept $\Theta^{(s+1)}$ as $\theta^{(s+1)}$ if $u \leq \min \left\{ \frac{\pi(\Theta^{(s+1)})}{\pi(\theta^s)}, 1 \right\}$, otherwise return θ^s

Sampling from $\pi(\rho, \beta / Z)$

$$\begin{aligned} \pi(\rho, \beta / Z) \propto & \exp \left\{ \sum_{i=1}^n \delta_i [(I_i - \rho W_i)\tau - X_i \beta] \right\} \exp \left\{ - \sum_{i=1}^n \exp [(I_i - \rho W_i)\tau - X_i \beta] \right\} \\ & \times \pi(\rho) \exp \left[-\frac{1}{2} (\beta' V^{-1} \beta) \right] \end{aligned}$$

Let's first try to derive conditional posteriors $\pi(\rho / \beta, Z)$ and $\pi(\beta / \rho, Z)$. Conditional posteriors are obtained by picking out from joint posterior terms involving the parameters whose conditional posterior we are interested in. The remaining terms are relegated to proportionality constant. Thus :

$$\pi(\rho / \beta, Z) \propto \exp \left\{ \left(-\rho \sum_{i=1}^n \delta_i W_i \tau \right) + \left(- \sum_{i=1}^n \exp [(I_i - \rho W_i)\tau - X_i \beta] \right) \right\} \quad (4.4)$$

$$\pi(\beta / \rho, Z) \propto \exp \left\{ - \left[\left(\sum_{i=1}^n X_i \beta \right) + \frac{1}{2} \beta' V^{-1} \beta \right] + \left[- \sum_{i=1}^n \exp [(I_i - \rho W_i)\tau - X_i \beta] \right] \right\} \quad (4.5)$$

Equations (4.4) and (4.5) show us conditional posterior distribution of our parameters. It may then seem straightforward to use Gibbs Sampler to draw from $\pi(\rho, \beta / Z)$, but it is not that easy. The fact of the matter is that, generating ρ^{s+1} from $\pi(\rho / \beta^s, Z)$ or β^{s+1} from $\pi(\beta / \rho^s, Z)$ in Gibbs sampler steps, may be pretty complicated given the form of these conditional posteriors. Then, we are going to use Metropolis-Hastings Algorithm within Gibbs Sampler.

3. Symmetric proposal

Metropolis-within-Gibbs algorithm is commonly used when conditional posteriors are not easy to draw from, but can easily be derived from the joint posterior . In our case, the algorithm can be summarized in these following steps:

1. Choose a starting values $\rho^{(0)}, \beta^{(0)}$
2. Use a **Random Walk Metropolis-Hastings** to draw
 - $\rho^{(1)}$ from $\pi(\rho, \beta^{(0)} / Z)$
 - $\beta^{(1)}$ from $\pi(\beta / \rho^{(1)}, Z)$
3. At the R^{th} iteration, draw
 - $\rho^{(R)}$ from $\pi(\rho, \beta^{(R-1)} / Z)$, using a **Random Walk Metropolis-Hastings**

Random Walk Metropolis-Hastings (RWMH) The RWMH is the one of the most frequently used Markov Chain Monte Carlo algorithms in practice. Suppose $\pi(\theta)$ is the target distribution we wish to sample from, the RWMH proceeds by generating proposed values in random directions from the current point:

$$\Theta^{(s)} = \theta^{(s-1)} + \omega \quad (\text{E5})$$

where ω is a random variable whose probability density is f_ω . In our case, we choose a centered normal distribution for f_ω . This ensures the proposal distribution , let $q(\theta, \Theta)$, to be symmetric and reduce the acceptance probability to $p = \min \left\{ \frac{\pi(\Theta)}{\pi(\theta)}, 1 \right\}$ ⁴. We can see through p that proposed values with higher posterior probability ($p = 1$ if $\pi(\Theta) > \pi(\theta)$) will always be accepted.

In the RWMH algorithm, one of the most important element is the choice of the dispersion of f_ω . This determines the way the chain moves. Let Σ be the variance of f_ω . When Σ is too small, the proposed values are too close to current value and this leads to a high acceptance probability. This is not desirable because the chain could take more time than one can afford to explore the entire posterior. Conversely, when Σ is too large, candidates draws are generated far away from current values at the cost of very small acceptance rate (see [Roberts et al., 2001](#); [G. M. Koop, 2003](#); [Gamerman & Lopes, 2006](#); [J. S. Rosenthal et al., 2011](#)). This critical issue has been widely considered by number of studies ([Roberts et al., 1997, 2001](#); [Bédard et al., 2007](#); [Bédard & Rosenthal, 2008](#)) in Mathematical Statistics litterature. Thus,

4. (E5) means that proposed value $\Theta^{(s)}$ is generated given the current value $\theta^{(s-1)}$. The proposal distribution is $q(\theta^{(s-1)}, \Theta^{(s)}) = f_\omega(\Theta^{(s)} - \theta^{(s-1)}) \Rightarrow q(\Theta^{(s)}, \theta^{(s-1)}) = f_\omega(\theta^{(s-1)} - \Theta^{(s)})$, and $q(\theta^{(s-1)}, \Theta^{(s)}) = q(\Theta^{(s)}, \theta^{(s-1)})$ when f_ω is a normal probability density.

it mainly appears that, instead of manually searching for better proposal distribution (what can be a very cumbersome process in high dimension), one can use adaptive chains where Σ is progressively adjusted to get good acceptance rate. Suppose θ is a k -dimensional vector and $\Sigma = \sigma^2 I$ where I is the $k \times k$ identity matrix. The adaptive RWMH here, will consist in randomly updating σ to get an optimal rate of acceptance. When $k \rightarrow \infty$, the optimal rate of acceptance is 0.234 (Roberts et al., 2001). In practice, this result still holds, even when $k = 5$. Nevertheless, when $k = 1$, the optimal acceptance is approximately 0.44 (J. S. Rosenthal et al., 2011).

4.5 Data and summary statistics

To apply our model, we randomly sample, from the dataset that we described in the section 1.2.2 of first chapter, 300⁵ manufacturing plants located in Île-de-France. Then, covariates presented in the Table 4.5.1 are the same detailed in chapter 1. The failure rate is 43.67%. The sampling is conducted keeping the same spatial structure as the original dataset⁶. The Île-de-France region is chosen because it is a highly densely populated area and the phenomenon we are trying to capture is more likely to take place there.

Île-de-France

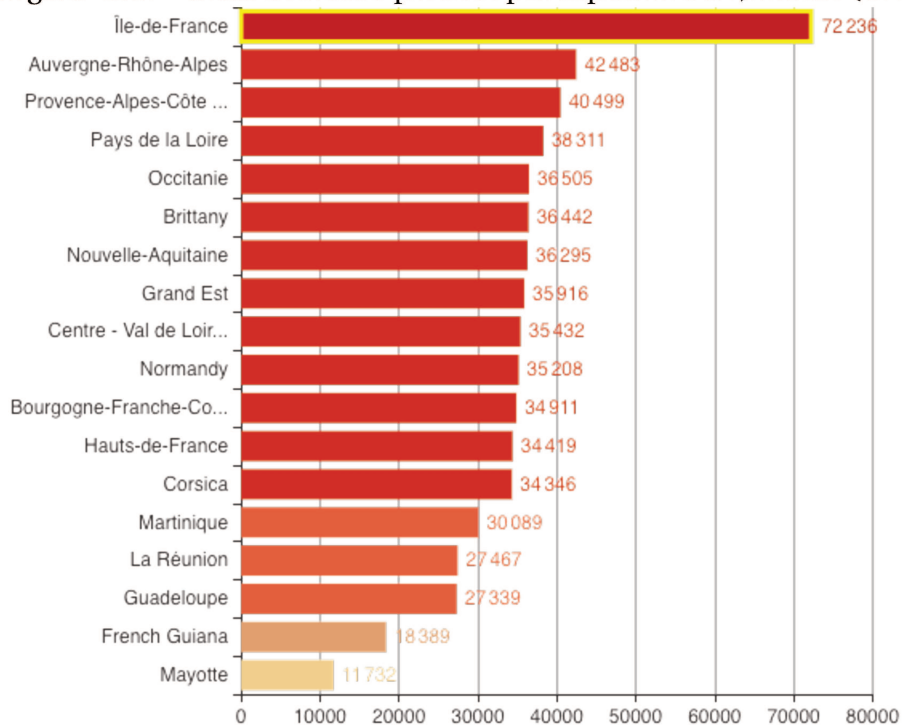
Île-de-France is the richest (OECD (2019)), the most innovative (OECD (2015)) and the most densely populated region in France (OECD (2020)). With a Gross Domestic Product per capita of 72 2236 USD in 2019, 240.5 patent PCT⁷ applications per million inhabitants in 2015 and 1028.41 inhabitants per km² in 2020; Île-de-France outperforms all other regions in France.

5. We have also tried the model on 3001 observations but it was such time-consuming that we have only drawn 8002 observations from the posterior (see appendix 4.B).

6. We sample by zone d'emploi (see section 1.2.2).

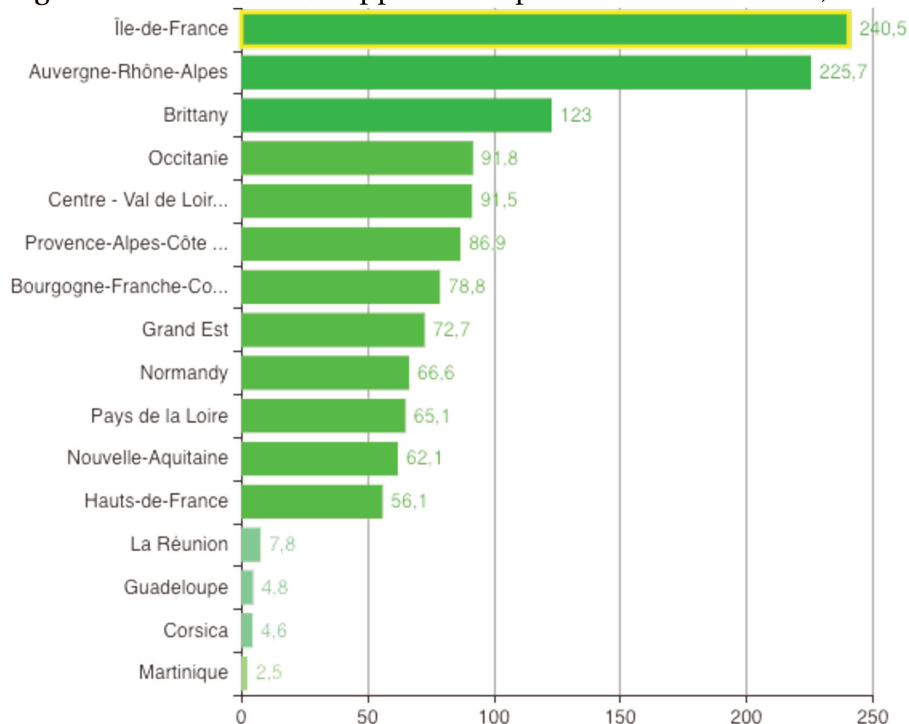
7. Patent Cooperation Treaty.

Figure 4.5.1 – Gross domestic product per capita in USD, France (OECD 2019)

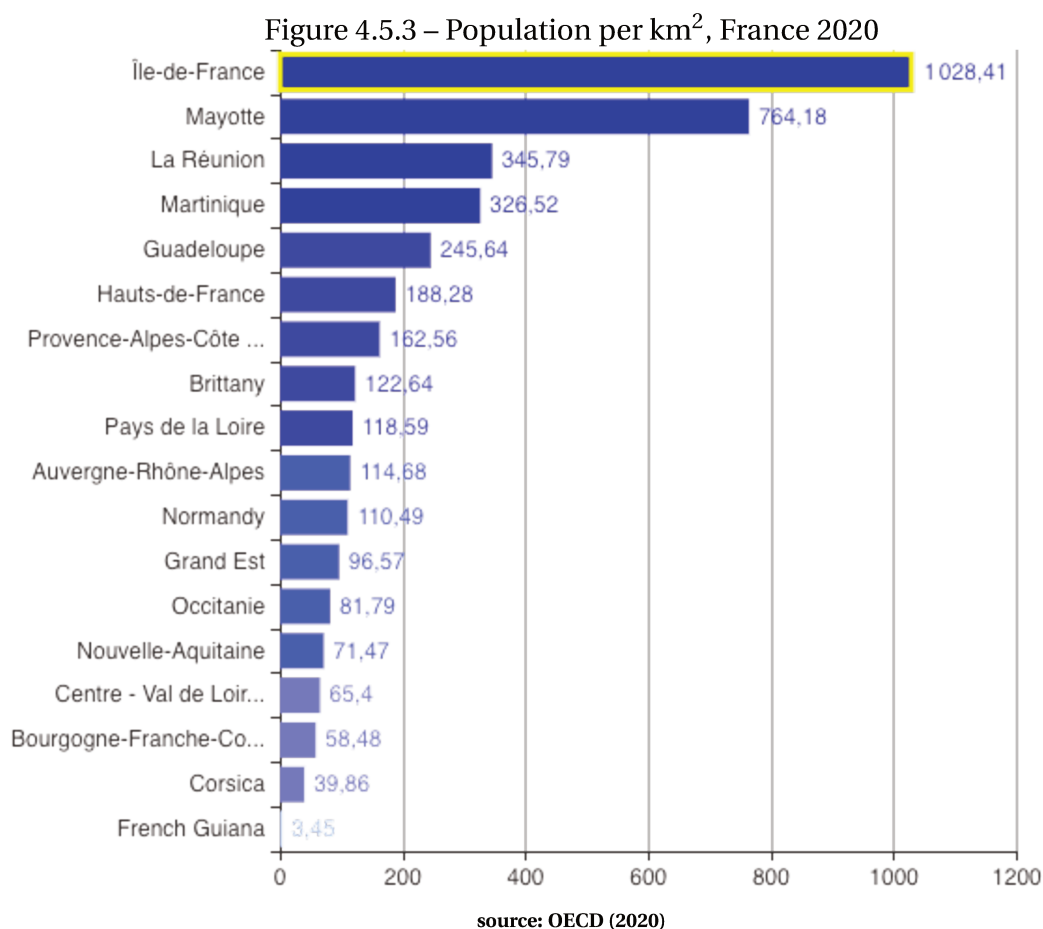


source: OECD (2019)

Figure 4.5.2 – Patent PCT applications per million inhabitants, France 2015



source: OECD (2015)



Such regions are very attractive for business creation because of resources they offer, but at the same time, it could be complicated for firm to survive there because of the competition (Brixy & Grotz, 2007; Huiban, 2011). Focusing on Île-de-France could help to capture the competition effect found in previous chapters. Thus, we expect here that firm survival will be negatively affected by neighboring firms survival.

Figures 4.5.4, 4.5.5 and 4.5.6 show micro-data on Île-de-France map. The weight matrix W is a row-standardized matrix of inverse distance.

Table 4.5.1 – Variables

Codes	Labels
taux_dep_eco_09	Economic dependency ratio
tx_chom_t1_03	Unemployment rate at employment area level
dpop16	Population density at city level
atsoc_p_km2	Number of facilities in social action per 100 people per square kilometer at city level
cmmrce_p_km2	Number of trade facilities per 100 people per square kilometer at city level
ens1erdeg_p_km2	Number of elementary school facilities per 100 people per square kilometer at city level
ens2nddeg_p_km2	Number of secondary school facilities per 100 people per square kilometer at city level
ensupformautr_p_km2	Number of facilities in higher education, training and education services per 100 people per square kilometer at city level
medparmed_p_km2	Number of facilities in medical and paramedical services per 100 people per square kilometer at city level
sante_p_km2	Number of facilities in health services per 100 people per square kilometer at city level
servparticul_p_km2	Number of facilities in personal services per 100 people per square kilometer at city level
sprtdrscult_p_km2	Number of facilities in sport and culture per 100 people per square kilometer at city level
trsptrsm_p_km2	Number of facilities in transport and tourism per 100 people per square kilometer at city level
sect_be	Dummy taking 1 for plants operating in Manufacturing, mining and quarrying and other industrial activities
sect_fz	Dummy taking 1 for plant operating in Construction
sect_gi	Dummy taking 1 for plant operating in Wholesale and retail trade, transportation and storage, accommodation and food service activities
sect_mn	Dummy taking 1 for plant operating in Professional, scientific, technical, administrative and support service activities
sect_oq	Dummy taking 1 for plant operating in Public administration and defence, education, human health and social work activities
distcaptl_km	Distance to the regional capital
d_km2_nbredetabs	Number of plants per square kilometer at city level
h_index	Herfindahl index
et_siege	Dummy taking 1 for headquarters
size_95	Number of employees in 1995

Figure 4.5.4 – Failed

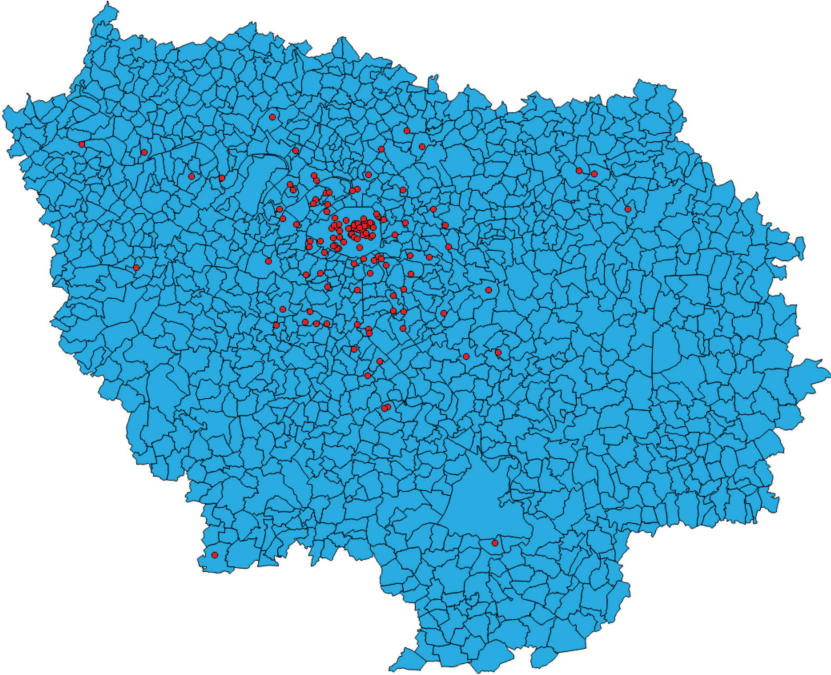


Figure 4.5.5 – censored

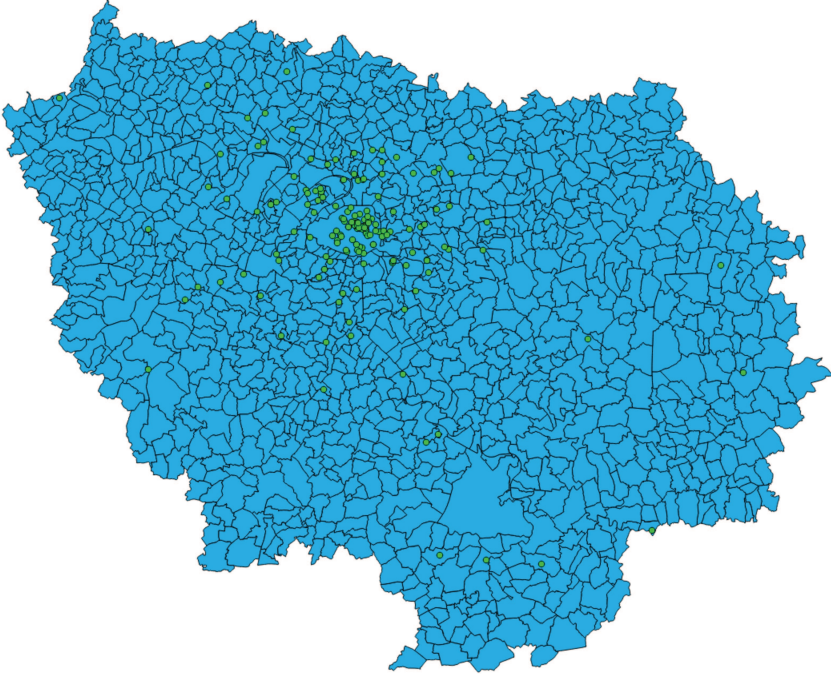
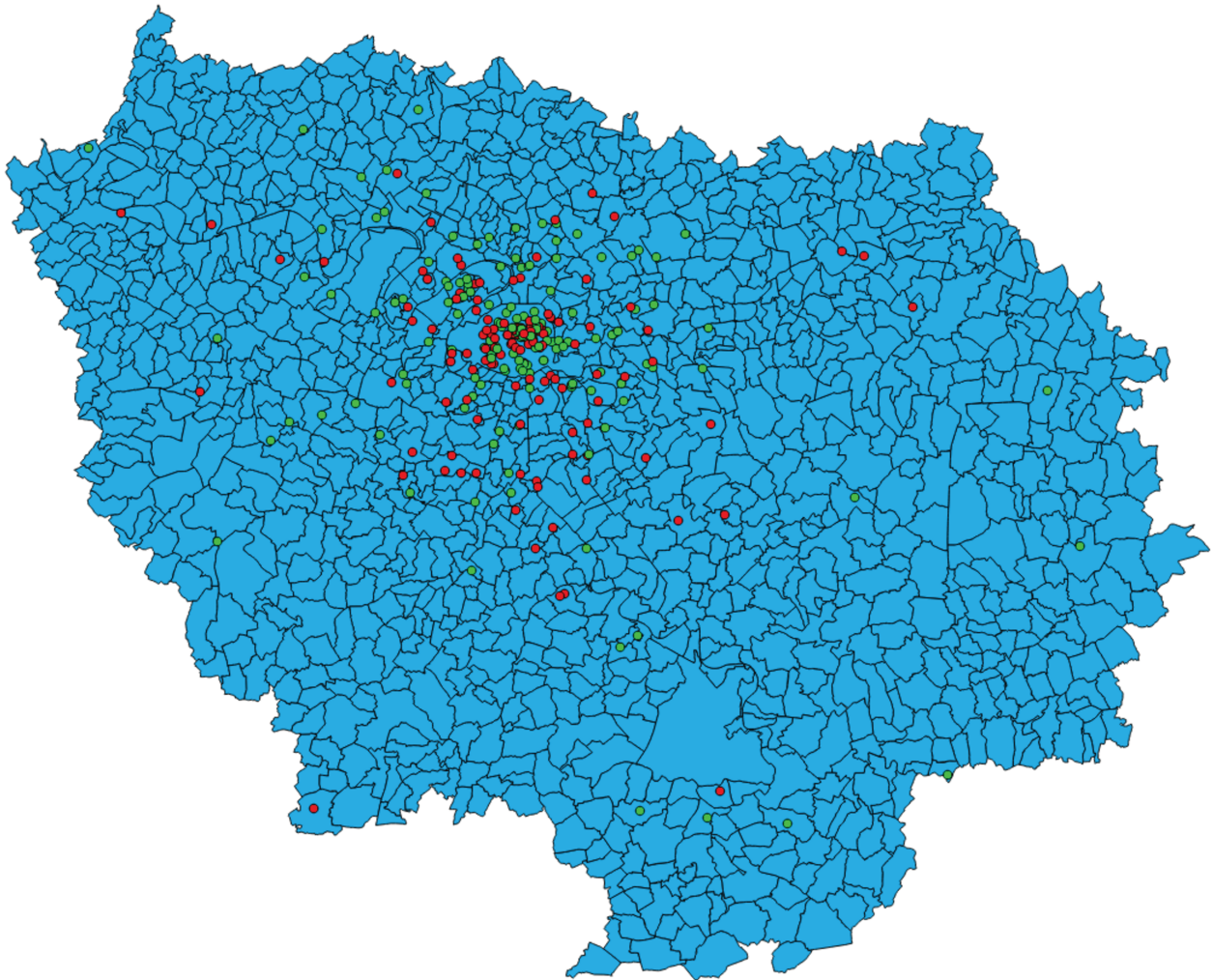


Figure 4.5.6 – Failed and Censored



Based on Figures 4.5.4, 4.5.5 and 4.5.6, establishments seem to be spatially concentrated. Figure 4.5.6 shows an agglomeration of failed and censored observations. This tend to suggest that successful firms are surrounded by non-successful ones and vice versa.

Table 4.5.2 – Statistics

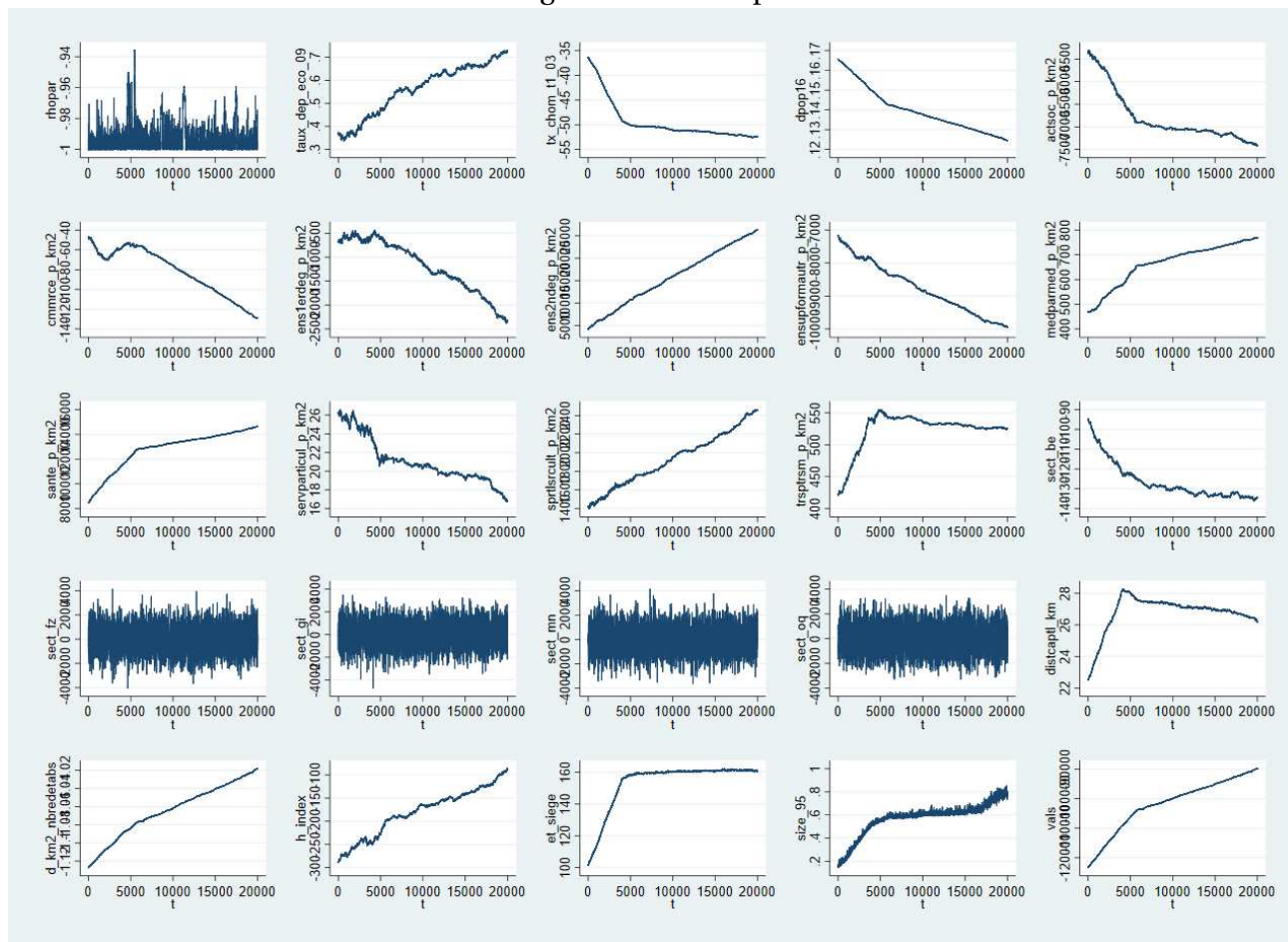
	Summary statistics			
	mean	standard deviation	min	max
taux_dep_eco_09	75.49	5.30	70.51	98.99
tx_chom_t1_03	8.28	1.12	4.90	9.80
dpop16	13640.65	12269.31	12.06	40168.58
actsoc_p_km2	0.02	0.02	0.00	0.11
cmmrce_p_km2	0.43	1.02	0.00	4.93
ens1erdeg_p_km2	0.01	0.01	0.00	0.07
ens2ndeg_p_km2	0.01	0.01	0.00	0.03
ensupformautr_p_km2	0.01	0.03	0.00	0.13
medparmed_p_km2	0.19	0.29	0.00	1.29
sante_p_km2	0.02	0.03	0.00	0.13
servparticul_p_km2	1.04	2.04	0.04	10.81
sprtlsrscult_p_km2	0.03	0.03	0.00	0.14
trsptrsm_p_km2	0.14	0.24	0.00	1.17
distcaptl_km	8.84	10.38	0.25	61.59
d_km2_nbredetabs	1138.49	1723.58	0.08	7189.90
pro_fer_com	42.89	4.97	0.00	71.43
pro_fer_ze	42.98	0.87	37.69	46.87
h_index	0.21	0.02	0.19	0.27
age_95	8.97	8.77	1.07	46.29
size_95	20.27	229.84	0.00	3976.00
Obs: 300				

Descriptive statistics are presented in Table 4.5.2.

4.6 Results

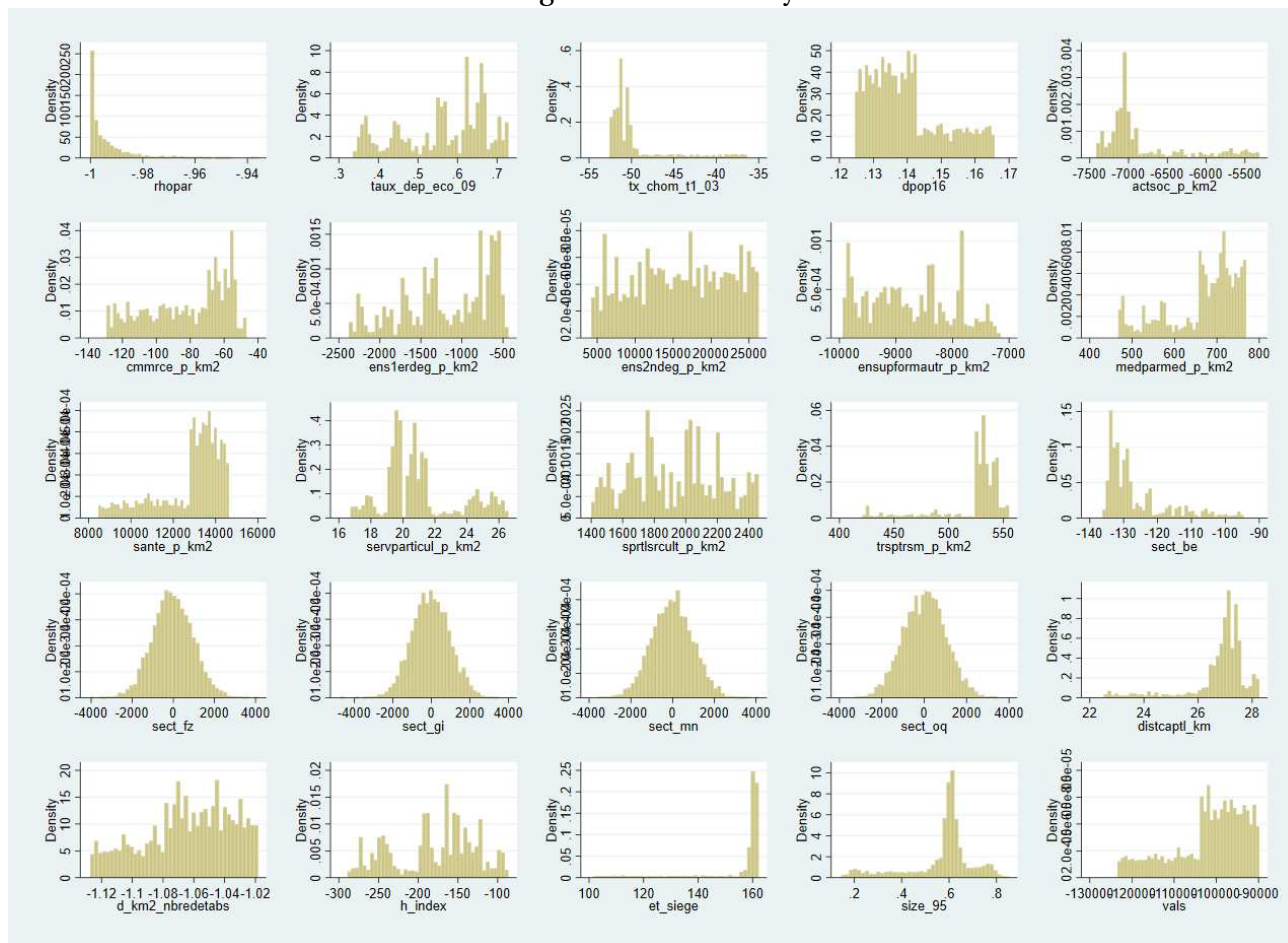
Figure 4.6.1 gives an overview of the estimates while Figure 4.6.2 shows the distribution of the parameters after MCMC run. Summary statistics of parameters are presented in Table 4.6.1.

Figure 4.6.1 – Sample



We draw a sample of 30 000 observations from our posterior distribution and discard 10 000 so we end up with 20 000 observations. Indeed, the starting value of the algorithm is not drawn from the posterior distribution, that is why some portion of the sample (called *burn-in*) must be discarded.

Figure 4.6.2 – Density



In our discussion, we will mainly focus on the autoregressive parameter ρ (*rhopar* in Table 4.6.1), because we are interested in the spatial spillover effects. The mean value of the parameter *rhopar* is -0.99 with a relatively low standard deviation (0.00904). This value kind of corroborates what Figure 4.5.6 foreshadows, that is, a firm survival's time is negatively affected by neighboring firms' survival time. This result differs from the contagion effect usually found in the literature (Maté-Sánchez-Val et al., 2018; Vivel-Búa & Lado-Sestayo, 2021), but sheds additional light on the competition effect consistently observed in the previous chapters.

In general, locations like Île-de-France are characterized by intense competition between firms. It is also well acknowledged that firm survival chances are reduced by the intensity of the competition (Brown & Earle, 2000; Buehler et al., 2004; Van Kranenburg et al., 2002). But the fact of the matter is that, when the competition heats up, the firms that exit the market are the less efficient ones. Instead, the most efficient firms survive, gain

more power, and the market eventually becomes concentrated in the hands of few producers. This means that sometimes, the competition is such fierce that for one firm, staying in the market means leaving it for another. That could be the reason why neighboring firms' survival times are negatively correlated. This reality is, to some extent, illustrated by some studies international economics, where [Coucke \(2005\)](#) has for example found that domestic firms have higher exit chances as foreign firms' market share raises.

Table 4.6.1 – Parameters

	Coefficients			
	mean	standard deviation	min	max
rhopar	-0.99	0.01	-1.00	-0.94
taux_dep_eco_09	0.56	0.11	0.34	0.73
tx_chom_t1_03	-49.47	3.87	-52.57	-36.37
dpop16	0.14	0.01	0.12	0.17
actsoc_p_km2	-6803.93	546.50	-7414.50	-5314.34
cmmrce_p_km2	-81.94	23.39	-129.26	-46.86
ens1erdeg_p_km2	-1186.24	538.49	-2375.55	-435.04
ens2ndeg_p_km2	15769.38	6293.56	4203.18	26262.62
ensupformautr_p_km2	-8758.61	759.23	-9942.87	-7154.85
medparmed_p_km2	666.35	82.85	467.39	768.79
sante_p_km2	12802.82	1533.46	8459.73	14630.74
servparticul_p_km2	21.04	2.33	16.73	26.59
sprtlsrcult_p_km2	1930.87	283.78	1399.35	2463.34
trsptrsm_p_km2	523.85	28.79	420.98	555.18
sect_be	-126.36	9.34	-136.20	-94.62
sect_fz	-28.38	987.49	-4036.31	4124.51
sect_gi	-28.54	986.47	-4719.98	3497.74
sect_mn	-10.52	988.65	-3621.45	4126.40
sect_oq	21.93	1010.96	-3288.46	3472.90
distcaptl_km	26.80	1.09	22.52	28.26
d_km2_nbredetabs	-1.06	0.03	-1.13	-1.02
h_index	-179.42	50.88	-289.02	-86.59
et_siege	153.87	14.67	101.62	162.21
size_95	0.57	0.14	0.15	0.85
Obs: 20 000				

Conclusion

As has already been reported in the literature ([Brixy & Grotz, 2007](#); [Huiban, 2011](#)), we found in previous chapters that there is higher hazard rates for business failure in more

densely populated areas. This can be understood as the consequence of competition among firms in urban areas, since prior research (Brown & Earle, 2000; Buehler et al., 2004; Van Kranenburg et al., 2002) suggested that competition has a positive effect on firms' exit probability. In this chapter we investigated this competition effect by modeling spatial dependencies in survival analysis.

Originally used for epidemiological and biomedical studies, survival models have since been democratized and are now used in several other fields such as economics. When survival data are spatially correlated, spatial frailty models are used (Zhou & Hanson, 2015; Banerjee, 2016), but because of the idea behind "frailty" (Vaupel et al., 1979), these models may not be suitable for certain types of data. When considering firms level data, spatial correlation does not necessarily stem from environment specific similarities, but it could emanate from business relation or competition rivalry firms have with their neighbors. Even though spatial probit is frequently used to account for spatial dependencies in firm survival analysis, probit model has some limitations in fitting time-to-events analysis (Kleinbaum & Klein, 2010; Liu, 2012). It's therefore important to think about a new way to account for spatial dependencies in firm survival analysis.

Our modeling strategy consists in using a spatial autoregressive structure in an AFT exponential model. Several studies have used the spatial autoregressive structure, and we think it is the most suitable approach to capture the spatial spillover effects generated by inter-firm competition. We use bayesian approach, and then derive a posterior distribution that can be drawn from using a Random Walk Metropolis-Hastings within Gibbs. We apply the model to 300 manufacturing plants located in Île-de-France. Our results suggest that firm survival time is negatively affected by neighbors survival time. This provides further clarification on competition effect observed in the first three chapters. Île-de-France is a rich urban area which can be very attractive for business creation and at the same time hostile. Our findings suggest that the level of competition is such that for one firm staying in the market means leaving it for another. Regarding other covariates, the algorithm does not draw from proper range so more iterations are needed. We also think that this chapter could be extended by using a distribution other than standard exponential distribution to model the baseline survival time. Some monte carlo experiments might also be used to assess how good the model is, compare to spatial frailty and spatial probit models.

Annexes

4.A Implementing the Model using Stata's Matrix programming

```
import delimited C:\Users\Public\Documents\Bases\zstata\IDF.csv
drop if sect_be==0
sample .775, by(ze2010)
mata
function PST(real rowvector T){
Xid=st_data(., ("siret", "taux_dep_eco_09", "tx_chom_t1_03", "dpop16", "actsoc_p_km2",
"cmrce_p_km2", "ens1erdeg_p_km2", "ens2ndeg_p_km2", "ensupformautr_p_km2", "medparmed_p_km2",
"sante_p_km2", "servparticul_p_km2", "sprtlsrucult_p_km2", "trsptrsm_p_km2", "sect_be",
"sect_fz", "sect_gi", "sect_mn", "sect_oq", "distcaptl_km", "d_km2_nbredetabs", "h_index", "et_siege", "size_95"))
X=sort(Xid,1)[.,(2\3\4\5\6\7\8\9\10\11\12\13\14\15\16\17\18\19\20\21\22\23\24)]
Yid=st_data(.,("siret", "survt"))
Y=sort(Yid,1)[.,2]
deltaid=st_data(.,("siret", "delta"))
delta=sort(deltaid,1)[.,2]
Did=st_data(.,("siret", "lat_y", "long_x"))
D=sort(Did,1)[.,(2\3)]
D=D*(pi()/180)
WW = J(rows(D),rows(D), 0)
for (i=1; i<=rows(D); i++){
for (j=1; j<i; j++){
WW[i,j]=WW[j,i] = 1/(6372.795*(2*asin(sqrt(sin((D[i,1]-D[j,1])/2)^2 +
cos(D[i,1])*cos(D[j,1])*sin((D[i,2]- D[j,2])/2)^2 ))))
}
}
W=WW:/(WW*J(rows(D),1,1))
I=I(rows(D))
V=1000000*I(23)
PSTA=sum(delta:*((I-((-1/(1+((T[1])^2)+ ((T[1])^2)/(1+((T[1])^2))))*W))*Y-X*(T[2] \T[3]
\T[4]\T[5]\T[6]\T[7]\T[8]\T[9]\T[10]\T[11]\T[12]\T[13]\T[14]\T[15]\T[16]
\T[17]\T[18]\T[19]\T[20]\T[21]\T[22]\T[23]\T[24]))))
PSTB=(-sum(exp((I-((-1/(1+((T[1])^2)+ ((T[1])^2)/(1+((T[1])^2))))*W))*Y-X*(T[2] \T[3]
\T[4]\T[5]\T[6]\T[7]\T[8]\T[9]\T[10]\T[11]\T[12]\T[13]\T[14]\T[15]\T[16]\T[17]
\T[18]\T[19]\T[20]\T[21]\T[22]\T[23]\T[24]))))
PSTC=(-0.5*((T[2] \T[3]\T[4]\T[5]\T[6]\T[7]\T[8]\T[9]\T[10]\T[11]\T[12]\T[13]\T[14]\T[15]
\T[16]\T[17]\T[18]\T[19]\T[20]\T[21]\T[22]\T[23]\T[24])' *luinv(V)*(T[2] \T[3]
\T[4]\T[5]\T[6]\T[7]\T[8]\T[9]\T[10]\T[11]\T[12]\T[13]\T[14]\T[15]\T[16]
\T[17]\T[18]\T[19]\T[20]\T[21]\T[22]\T[23]\T[24]))))
return (PSTA+PSTB+PSTC)
}
end
```

```

mata
alginfo="mwg"
MC=amcmc(alginfo,&PST(),J(1,24,0.5),I(24),30000,10000,1/2,0.234,arate=.,vals=.,lam=.,...)
end
mata
rho=J(rows(MC),1,-1)/(J(rows(MC),1,1)+(MC[.,1]:*MC[.,1])) + (MC[.,1]:*MC[.,1])/(J(rows(MC),1,1)+(MC[.,1]:*MC[.,1]))
RWMG=rho,MC[.,(2\3\4\5\6\7\8\9\10\11\12\13\14\15\16\17\18\19\20\21\22\23\24)]
rhoes=mean(rho)
allparm=mean(RWMG)
rhoes
allparm
varcovar=variance(RWMG)
vars=diagonal(varcovar)
sderr=sqrt(vars)
end
local varnames rhopar taux_dep_eco_09 tx_chom_t1_03 dpop16 actsoc_p_km2 cmmrce_p_km2 ens1erdeg_p_km2 ens2ndeg_p_km2 ///
ensupformautr_p_km2 medparmed_p_km2 sante_p_km2 servparticul_p_km2 sprtlsrcult_p_km2 trsptrsm_p_km2 ///
sect_be sect_fz sect_gi sect_mn sect_oq distcapt1_km d_km2 nbredetabs h_index et_siege size_95
getmata ('varnames')=RWMG
getmata vals=vals
generate t=_n
local graphs
local tgraphs
foreach var of local varnames {
quietly {
histogram `var', saving(`var', replace)nodraw
tway line `var' t, saving(t`var', replace) nodraw
}
local graphs "`graphs' `var'.gph"
local tgraphs "`tgraphs' t`var'.gph"
}
histogram vals, saving(vals,replace) nodraw
tway line vals t, saving(vals_t,replace) nodraw
graph combine `graphs' vals.gph
graph export vals_mwg.eps, replace
graph combine `tgraphs' vals_t.gph
graph export valst_mwg.eps, replace

```

4.B Using 3001 randomly selected plants from Île-de-France

Figure 4.B.1 – Sample

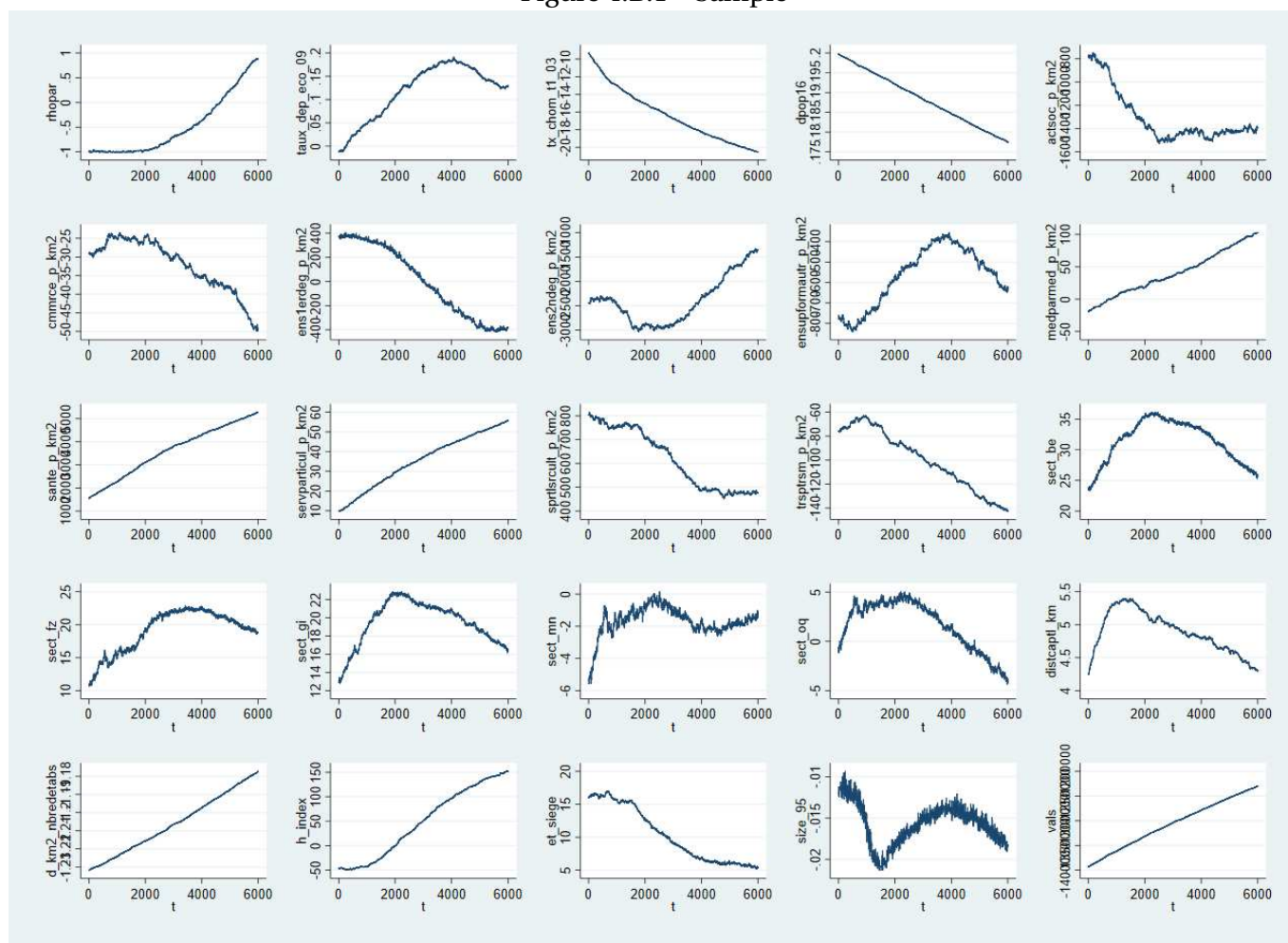


Figure 4.B.2 – Density

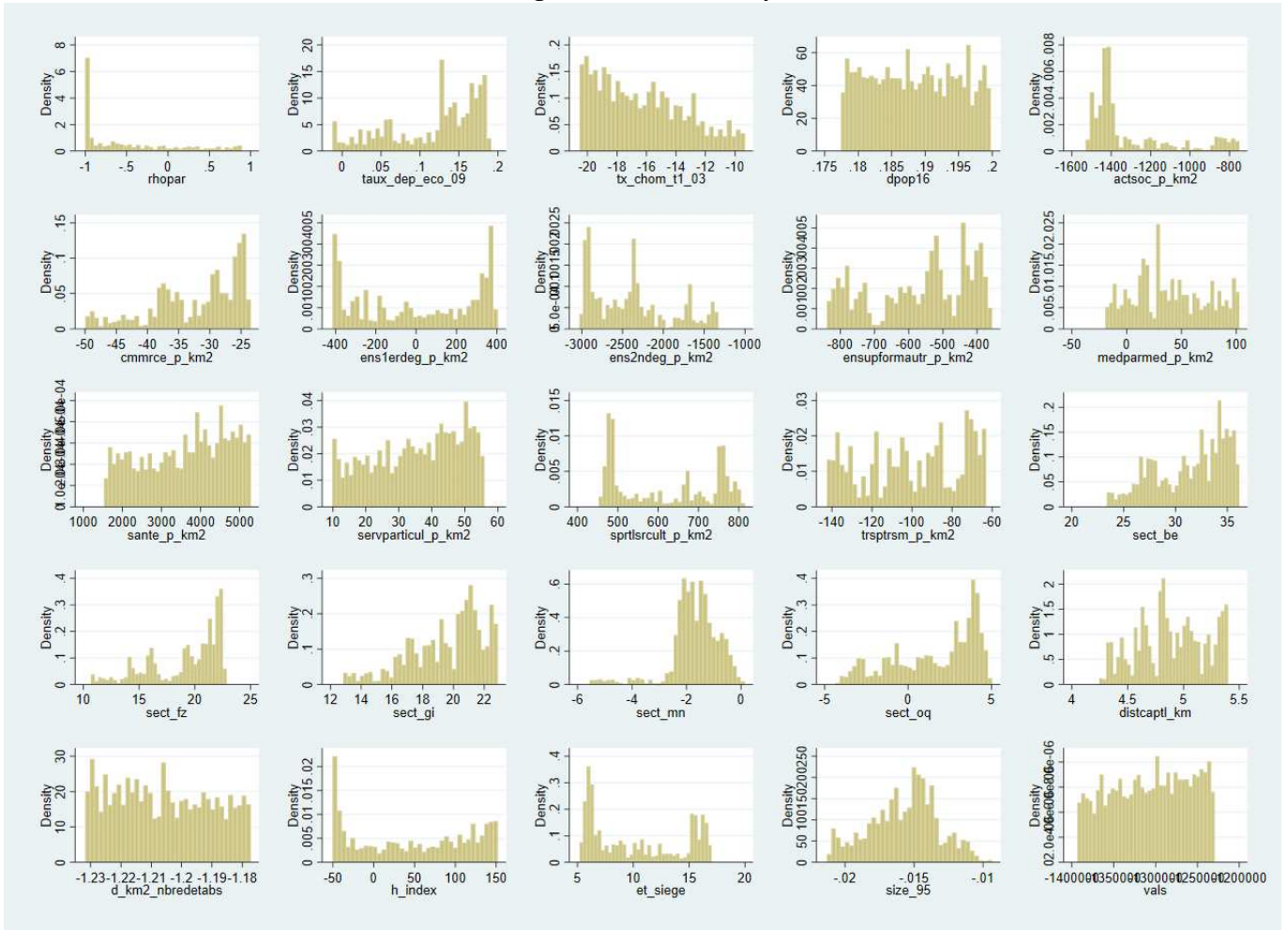


Table 4.B.1 – Sampled parameters

	Coefficients			
	mean	standard deviation	min	max
rhopar	-0.48	0.59	-1.00	0.89
taux_dep_eco_09	0.12	0.06	-0.01	0.19
tx_chom_t1_03	-16.31	2.97	-20.52	-9.32
dpop16	0.19	0.01	0.18	0.20
actsoc_p_km2	-1308.49	218.34	-1529.29	-750.07
cmmrce_p_km2	-32.19	6.94	-49.93	-23.47
ens1erdeg_p_km2	2.11	289.90	-414.64	405.17
ens2ndeg_p_km2	-2397.40	483.33	-3028.10	-1328.63
ensupformautr_p_km2	-566.39	142.28	-839.67	-355.94
medparmed_p_km2	40.44	33.67	-19.24	102.79
sante_p_km2	3614.96	1080.12	1537.29	5279.39
servparticul_p_km2	35.53	13.25	9.85	55.96
sprtlsrcult_p_km2	617.36	123.65	452.90	815.36
trsptrsm_p_km2	-99.68	24.43	-142.54	-62.95
sect_be	31.43	3.38	23.39	36.15
sect_fz	19.16	3.16	10.69	22.84
sect_gi	19.54	2.37	12.84	22.88
sect_mn	-1.68	0.89	-5.56	0.15
sect_oq	1.66	2.48	-4.36	5.04
distcaptl_km	4.89	0.31	4.24	5.40
d_km2_nbredetabs	-1.21	0.02	-1.23	-1.18
h_index	48.18	69.72	-50.44	152.06
et_siege	10.30	4.12	5.18	17.02
size_95	-0.02	0.00	-0.02	-0.01
Obs: 6002				

Conclusion générale

Les entreprises, à travers l'innovation et la promotion de l'emploi, apportent une contribution majeure au développement économique régional et même national. Cependant, il ne suffit pas pour une entreprise de naître pour avoir de l'impact, mais il lui faut survivre et prospérer, d'où l'importance de s'interroger sur les déterminants de la survie des entreprises. La littérature associée à cette thématique pointe généralement trois groupes de facteurs à savoir : les caractéristiques de l'entreprise, les caractéristiques de l'industrie dans laquelle l'entreprise opère et les caractéristiques de l'environnement spatial de l'entreprise. Les deux premières catégories de facteurs ayant déjà fait l'objet d'une abondante littérature, nous nous sommes, dans les quatre chapitres de cette thèse, focalisés sur l'importance de l'environnement spatial de l'entreprise. La mise en œuvre d'une variété de techniques économétriques sur une diversité d'échantillon a permis de mettre en exergue l'impact significatif qu'exerce l'environnement spatial de l'entreprise sur sa survie.

Nous avons débuté par une analyse à l'échelle des établissements qui sont définis comme des unités de production géographiquement individualisées (INSEE). L'estimation de modèles à hasards proportionnels de Cox sur 1 651 798 établissements français nous montre que l'environnement spatial affecte significativement la survie des établissements français. Cette étude empirique a été complétée par une analyse théorique qui a permis d'établir que l'âge auquel l'entreprise ferme, est lié aux caractéristiques de l'environnement spatial lorsque ces dernières sont prises en compte.

Nous nous sommes ensuite intéressés à la survie des jeunes entreprises françaises dans le chapitre 2. Nous en avons observé 45 268 sur une période de cinq ans à partir de leur naissance. Dans ce chapitre, nous avons estimé la propension de l'entreprise à choisir sa localisation observée et nous avons pris cette information en compte dans l'analyse de la survie. Nos résultats suggèrent dans un premier temps que les caractéristiques de l'entreprise, de l'entrepreneur, du secteur d'activité, et de l'environnement spatial, impactent significativement la survie des jeunes entreprises. Dans un second temps, la prise en compte du choix de localisation des entreprises nous révèle que plus il est probable pour une entreprise de choisir sa localisation actuelle, moins il est risqué pour elle de fermer.

Dans le troisième chapitre, l'accent est mis sur l'environnement géographique du quartier général de l'entreprise, qui constitue une entité de premier plan dont les fonctions diffèrent de celles des unités de production. La stratégie empirique du chapitre est faite d'estimation de modèles à hasards proportionnels de Cox sur un échantillon de 761 108

entreprises françaises observées de 2005 à 2015. Les résultats indiquent que le lieu d'implantation du quartier général d'une entreprise joue un rôle significatif dans sa survie.

Le dernier chapitre étudie un peu plus en détail l'une des régularités spatiales mises en évidence dans les chapitres précédents. Il s'agit de l'effet de concurrence qui se manifeste par des risques de fermeture élevés auxquels sont confrontées les entreprises dans les communes densément peuplées. Pour mieux modéliser cette problématique, nous avons proposé un modèle de survie spatial applicable sur micro-données. En effet, nous avons inclus un décalage spatial dans un modèle à durée de vie accélérée exponentielle afin de prendre en compte les effets d'entraînement spatiaux découlant des interactions stratégiques entre les entreprises voisines. Le modèle est traité par approche bayésienne. En l'appliquant sur 300 établissements franciliens du secteur manufacturier, nous notons que la survie d'une entreprise est négativement impactée par la survie de ses voisines.

L'une des forces de cette thèse est qu'elle utilise diverses techniques économétriques sur des données microéconomiques de grande qualité, et une variété d'échantillons de tailles conséquentes. Nous avons également fait montre d'une certaine rigueur statistique en n'interprétant que les résultats significatifs à 5%, 1% et 1%. De plus les caractéristiques territoriales que nous avons considérées, sont principalement des infrastructures découlant de politiques publiques. Ainsi, la compréhension de l'impact de ces caractéristiques permet d'orienter les politiques publiques en faveur de la création d'environnements favorables aux entreprises.

Toutefois, certains éléments de ce travail peuvent faire l'objet d'une considération davantage approfondie. Dans notre thèse, les variables décrivant l'environnement géographique de l'entreprise sont calculées à l'échelle communale et de la zone d'emploi, ce qui fait que nous n'avons pas beaucoup de variabilité au niveau de ces variables. Un travail plus minutieux peut être effectué en calculant le nombre d'équipements dans un voisinage plus immédiat de l'entreprise. L'on pourrait également aller plus loin en utilisant une classification plus raffinée des équipements. Cela permettrait de connaître l'impact de chaque type spécifique d'équipement. Un autre point possible d'approfondissement de ce travail pourrait être la redéfinition de la notion de réussite pour l'entreprise. Dans notre thèse, la notion de réussite se résume au fait que l'entreprise soit active ou non. Pourtant, une entreprise active n'est pas systématiquement profitable. Une recherche future pourrait combiner les notions de survie et de performances économique de l'entreprise (profitabilité) pour construire un nouvel indicateur de réussite.

Bibliographie

- Aarland, K., Davis, J. C., Henderson, J. V., & Ono, Y. (2007). Spatial organization of firms : The decision to split production and administration. *The RAND Journal of Economics*, 38(2), 480–494.
- Acs, Z. J., Armington, C., & Zhang, T. (2007). The determinants of new-firm survival across regional economies : The role of human capital stock and knowledge spillover. *Papers in Regional Science*, 86(3), 367–391.
- Adler, P., & Florida, R. (2020). Geography as strategy : the changing geography of corporate headquarters in post-industrial capitalism. *Regional Studies*, 54(5), 610–620.
- Agarwal, R., & Gort, M. (1996). The evolution of markets and entry, exit and survival of firms. *The review of Economics and Statistics*, 489–498.
- Agarwal, R., & Gort, M. (2002). Firm and product life cycles and firm survival. *The American Economic Review*, 92(2), 184–190.
- Andersson, P., & Wadensjö, E. (2007). Do the unemployed become successful entrepreneurs? *International Journal of Manpower*.
- Anselin, L. (1982). A note on small sample properties of estimators in a first-order spatial autoregressive model. *Environment and Planning A*, 14(8), 1023–1030.
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93–115.
- Anselin, L. (2013). *Spatial econometrics : methods and models* (Vol. 4). Springer Science & Business Media.
- Arcuri, G., Brunetto, M., & Levratto, N. (2019). Spatial patterns and determinants of firm exit : an empirical analysis on france. *The Annals of Regional Science*, 62(1), 99–118.
- Arrow, K. (1962). The economic implications of learning by doing. review of economic studies, june.
- Asheim, B. T., & Isaksen, A. (2002). Regional innovation systems : the integration of local 'sticky' and global 'ubiquitous' knowledge. *The Journal of Technology Transfer*, 27(1), 77–86.

- Audia, P. G., & Rider, C. I. (2010). Close, but not the same : locally headquartered organizations and agglomeration economies in a declining industry. *Research Policy*, 39(3), 360–374.
- Audretsch, D. B. (1991). New-firm survival and the technological regime. *The Review of Economics and Statistics*, 441–450.
- Audretsch, D. B. (1995). Innovation, growth and survival. *International journal of industrial organization*, 13(4), 441–457.
- Audretsch, D. B., & Fritsch, M. (2002). Growth regimes over time and space. *Regional Studies*, 36(2), 113–124.
- Audretsch, D. B., Lehmann, E. E., & Warning, S. (2005). University spillovers and new firm location. *Research Policy*, 7(34), 1113–1122.
- Audretsch, D. B., & Mahmood, T. (1995). New firm survival : new results using a hazard function. *The Review of Economics and Statistics*, 97–103.
- Banerjee, S. (2016). Spatial survival models. In *Handbook of spatial epidemiology* (pp. 321–332). Chapman and Hall/CRC.
- Banerjee, S., Wall, M. M., & Carlin, B. P. (2003). Frailty modeling for spatially correlated survival data, with application to infant mortality in minnesota. *Biostatistics*, 4(1), 123–142.
- Baptista, R., Escária, V., & Madruga, P. (2008). Entrepreneurship, regional development and job creation : the case of portugal. *Small Business Economics*, 30(1), 49–58.
- Baptista, R., Lima, F., & Mendonça, J. (2011). Establishment of higher education institutions and new firm entry. *Research Policy*, 40(5), 751–760.
- Basile, R., Pittiglio, R., & Reganati, F. (2017). Do agglomeration externalities affect firm survival? *Regional Studies*, 51(4), 548–562.
- Bastos, L. S., & Gamerman, D. (2006). Dynamic survival models with spatial frailty. *Lifetime data analysis*, 12(4), 441–460.
- Baum, J. A., & Mezias, S. J. (1992). Localized competition and organizational failure in the manhattan hotel industry, 1898-1990. *Administrative science quarterly*, 580–604.

- Baumöhl, E., Iwasaki, I., & Kočenda, E. (2019). Institutions and determinants of firm survival in european emerging markets. *Journal of Corporate Finance*, 58, 431–453.
- Baumöhl, E., Iwasaki, I., & Kočenda, E. (2020). Firm survival in new EU member states. *Economic Systems*, 44(1), 100743.
- Baumol, W. J. (2002). *The free-market innovation machine : Analyzing the growth miracle of capitalism*. Princeton University Press.
- Bédard, M., et al. (2007). Weak convergence of metropolis algorithms for non-iid target distributions. *The Annals of Applied Probability*, 17(4), 1222–1244.
- Bédard, M., & Rosenthal, J. S. (2008). Optimal scaling of metropolis algorithms : Heading toward general target distributions. *Canadian Journal of Statistics*, 36(4), 483–503.
- Beguín, H. (1988). La région et les lieux centraux. In *Analyse économique spatiale* (pp. 231–276). Presses Universitaires de France.
- Belleflamme, P., & Peitz, M. (2010). *Industrial organization* (Reprint. éd.). Cambridge [u.a.] : Cambridge Univ. Press. (Literaturangaben)
- Bellman, R. (1965). An application of dynamic programming to location—allocation problems. *Siam Review*, 7(1), 126–128.
- Bernard, A. B., & Jensen, J. B. (1999). Exceptional exporter performance : cause, effect, or both? *Journal of international economics*, 47(1), 1–25.
- Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic perspectives*, 21(3), 105–130.
- Birch, D. L. (1979). *The job generation process* (Vol. 302). MIT program on neighborhood and regional change Cambridge, MA.
- Boone, J. (2008). Competition : Theoretical parameterizations and empirical measures. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 587–611.
- Børing, P. (2015). The effects of firms' r&d and innovation activities on their survival : a competing risks analysis. *Empirical Economics*, 49(3), 1045–1069.
- Boschma, R., & Martin, R. (2007). *Constructing an evolutionary economic geography*. Oxford University Press.

- Boyer, T., & Blazy, R. (2014). Born to be alive? the survival of innovative and non-innovative french micro-start-ups. *Small Business Economics*, 42(4), 669–683.
- Brickley, J. A., & Dark, F. H. (1987). The choice of organizational form the case of franchising. *Journal of financial economics*, 18(2), 401–420.
- Brixy, U., & Grotz, R. (2007). Regional patterns and determinants of birth and survival of new firms in western germany. *Entrepreneurship and Regional Development*, 19(4), 293–312.
- Brock, W. A. (1972). On models of expectations that arise from maximizing behavior of economic agents over time. *Journal of Economic Theory*, 5(3), 348–376.
- Brown, D., & Earle, J. (2000). Market competition and firm performance in russia. *Russian Economic Trends*, 9(1), 13–18.
- Buddelmeyer, H., Jensen, P. H., & Webster, E. (2010). Innovation and the determinants of company survival. *Oxford Economic Papers*, 62(2), 261–285.
- Buehler, S., Kaiser, C., & Jaeger, F. (2004). Competition and exit : Evidence from switzerland. *University of Zurich SOI Working Paper*(0405).
- Cainelli, G., Montresor, S., & Marzetti, G. V. (2014). Spatial agglomeration and firm exit : a spatial dynamic analysis for italian provinces. *Small Business Economics*, 43(1), 213–228.
- Cefis, E., & Marsili, O. (2005). A matter of life and death : innovation and firm survival. *Industrial and Corporate change*, 14(6), 1167–1192.
- Cefis, E., & Marsili, O. (2011). Born to flip. exit decisions of entrepreneurial firms in high-tech and low-tech industries. *Journal of Evolutionary Economics*, 21(3), 473–498.
- Cevik, S., & Miryugin, F. (2022). Death and taxes : Does taxation matter for firm survival? *Economics & Politics*, 34(1), 92–112.
- Chamberlin, E. (1933). *The theory of monopolistic competition*. Harvard University Press.
- Chang, Z., & Zheng, L. (2022). High-speed rail and the spatial pattern of new firm births : Evidence from china. *Transportation Research Part A : Policy and Practice*, 155, 373–386.
- Chiang, A. (1992). Elements of dynamic optimization. McGaw-Hill. *International Editions, Singapore*.

- Claval, P. (1966). Chronique de géographie économique. la théorie des lieux centraux. *Revue géographique de l'Est*, 6(1), 131–152.
- Colombelli, A., Krafft, J., & Vivarelli, M. (2016). To be born is not enough : the key role of innovative start-ups. *Small Business Economics*, 47(2), 277–291.
- Cooper, A. C., Gimeno-Gascon, F. J., & Woo, C. Y. (1994). Initial human and financial capital as predictors of new venture performance. *Journal of business venturing*, 9(5), 371–395.
- Corporation, S. W. P., Mills, C. W., & Ulmer, M. J. (1946). *Small business and civic welfare : Report of the smaller war plants corporation to the special committee to study problems of american small business, united states senate*. US Government Printing Office.
- Coucke, K. (2005). Import competition and the exit of firms in belgian manufacturing. Available at SSRN 870395.
- Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34(2).
- Craioveanu, M., & Terrell, D. (2016). The impact of storms on firm survival : a bayesian spatial econometric model for firm survival. In *Spatial econometrics : Qualitative and limited dependent variables*. Emerald Group Publishing Limited.
- Cressy, R. (1996). Are business startups debt-rationed? *The economic journal*, 106(438), 1253–1270.
- Danson, M. (2002). *Small firm formation and regional economic development*. Routledge.
- Dasgupta, P., & Stiglitz, J. (1981). Entry, innovation, exit : Towards a dynamic theory of oligopolistic industrial structure. *European Economic Review*, 15(2), 137–158.
- Davidsson, P., Lindmark, L., & Olofsson, C. (1994). New firm formation and regional development in sweden. *Regional Studies*, 28(4), 395–410.
- Davis, J. C., & Henderson, J. V. (2008). The agglomeration of headquarters. *Regional Science and Urban Economics*, 38(5), 445–460.
- Derbyshire, J. (2013). The survival half-life of firms and its effect on economic development. *Local Economy*, 28(1), 114–122.
- Dicharry, B., Nguyen-Van, P., & Pham, T. K. C. (2019). La politique de cohésion de l'ue et la convergence économique. *Bulletin de l'Observatoire des politiques économiques en Europe*, 41(1), 5–11.

- Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, 67(3), 297–308.
- Dubé, J., Brunelle, C., & Legros, D. (2016). Location theories and business location decision : A micro-spatial investigation of a nonmetropolitan area in canada. *Review of Regional Studies*, 46(2), 143–170.
- Dubé, J., & Legros, D. (2014). *Économétrie spatiale appliquée des microdonnées*. ISTE Group.
- Dumais, G., Ellison, G., & Glaeser, E. L. (2002). Geographic concentration as a dynamic process. *Review of economics and Statistics*, 84(2), 193–204.
- Dunne, T., Roberts, M. J., & Samuelson, L. (1989). The growth and failure of us manufacturing plants. *The Quarterly Journal of Economics*, 104(4), 671–698.
- Ebert, T., Brenner, T., & Brixy, U. (2019). New firm survival : the interdependence between regional externalities and innovativeness. *Small Business Economics*, 53(1), 287–309.
- Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries : a dartboard approach. *Journal of political economy*, 105(5), 889–927.
- Essletzbichler, J., & Rigby, D. L. (2010). Generalized darwinism and evolutionary economic geography. In *The handbook of evolutionary economic geography*. Edward Elgar Publishing.
- Esteve-Pérez, S., & Mañez-Castillejo, J. A. (2008). The resource-based theory of the firm and firm survival. *Small Business Economics*, 30(3), 231–249.
- Evans, A. W. (1973). The location of the headquarters of industrial companies. *Urban Studies*, 10(3), 387–395.
- Evans, D. S. (1987). The relationship between firm growth, size, and age : Estimates for 100 manufacturing industries. *The journal of industrial economics*, 567–581.
- Ezcurra, R., & Del Villar, A. (2021). Globalization and spatial inequality : Does economic integration affect regional disparities? *The Annals of Regional Science*, 1–24.
- Ezcurra, R., & Rodríguez-Pose, A. (2013). Does economic globalization affect regional inequality? a cross-country analysis. *World Development*, 52, 92–103.

- Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities : : Science-based diversity, specialization and localized competition. *European economic review*, 43(2), 409–429.
- Florida, R., Mellander, C., & Holgersson, T. (2015). Up in the air : the role of airports for regional economic development. *The annals of regional science*, 54(1), 197–214.
- Fotopoulos, G., & Louri, H. (2000). Location and survival of new entry. *Small business economics*, 14(4), 311–321.
- Friedland, R., & Palmer, D. (1984). Park place and main street : Business and the urban power structure. *Annual Review of Sociology*, 10(1), 393–416.
- Fritsch, M., Brixy, U., & Falck, O. (2006). The effect of industry, region, and time on new business survival—a multi-dimensional analysis. *Review of industrial organization*, 28(3), 285–306.
- Fudenberg, D., & Tirole, J. (1986). A theory of exit in duopoly. *Econometrica : Journal of the Econometric Society*, 943–960.
- Gamerman, D., & Lopes, H. F. (2006). *Markov chain monte carlo : stochastic simulation for bayesian inference*. Chapman and Hall/CRC.
- Gantmakher, F. R. (1959). *The theory of matrices* (Vol. 131). American Mathematical Soc.
- Gehan, E. A. (1965). A generalized wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, 52(1-2), 203–224.
- Glaeser, E. L. (1999). Learning in cities. *Journal of urban Economics*, 46(2), 254–277.
- Goldstein, G. S., & Gronberg, T. J. (1984). Economies of scope and economies of agglomeration. *Journal of Urban Economics*, 16(1), 91–104.
- Greenberg, E. (2012). *Introduction to bayesian econometrics*. Cambridge University Press.
- Haltiwanger, J. C., Lane, J. I., & Spletzer, J. (1999). Productivity differences across employers : The roles of employer size, age, and human capital. *American Economic Review*, 89(2), 94–98.
- Hannan, M. T. (1998). Rethinking age dependence in organizational mortality : Logical formalizations. *American Journal of Sociology*, 104(1), AJSv104p126–164.
- Hannan, M. T., & Freeman, J. (1977). The population ecology of organizations. *American journal of sociology*, 82(5), 929–964.

- Helmers, C., & Rogers, M. (2010). Innovation and the survival of new firms in the uk. *Review of Industrial Organization*, 36(3), 227–248.
- Helsley, R. W., & Strange, W. C. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics*, 20(2), 189-212. Consulté sur <https://www.sciencedirect.com/science/article/pii/016604629090004M> doi: [https://doi.org/10.1016/0166-0462\(90\)90004-M](https://doi.org/10.1016/0166-0462(90)90004-M)
- Henderson, R., Shimakura, S., & Gorst, D. (2002). Modeling spatial variation in leukemia survival data. *Journal of the American Statistical Association*, 97(460), 965–972.
- Hertzel, M. G., Li, Z., Officer, M. S., & Rodgers, K. J. (2008). Inter-firm linkages and the wealth effects of financial distress along the supply chain. *Journal of Financial Economics*, 87(2), 374–387.
- Holmes, P., Hunt, A., & Stone, I. (2010). An analysis of new firm survival using a hazard function. *Applied Economics*, 42(2), 185–195.
- Holmes, T. J. (1999). Localization of industry and vertical disintegration. *Review of Economics and Statistics*, 81(2), 314–325.
- Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica : Journal of the Econometric Society*, 1127–1150.
- Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis*. doi: 10.1002/9780470258019
- Hotelling, H. (1929). Stability in competition. *The economic journal*, 39(153), 41–57.
- Houdebine, M. (1999). Concentration géographique des activités et spécialisation des départements français. *Economie et statistique*, 326(1), 189–204.
- Howell, A., He, C., Yang, R., & Fan, C. C. (2018). Agglomeration,(un)-related variety and new firm survival in china : Do local subsidies matter? *Papers in Regional Science*, 97(3), 485–500.
- Huggins, R., Prokop, D., & Thompson, P. (2017). Entrepreneurship and the determinants of firm survival within regions : human capital, growth motivation and locational conditions. *Entrepreneurship & Regional Development*, 29(3-4), 357–389.
- Huiban, J.-P. (2011). The spatial demography of new plants : urban creation and rural survival. *Small Business Economics*, 37(1), 73–86.

- Hyytinen, A., Pajarinen, M., & Rouvinen, P. (2015). Does innovativeness reduce startup survival rates? *Journal of business venturing*, 30(4), 564–581.
- INSEE. (2019). *Les entreprises en france-édition 2019*. Institut national de la statistique et des études économiques.
- Iwasaki, I., & Kočenda, E. (2020). Survival of service firms in european emerging economies. *Applied Economics Letters*, 27(4), 340–348.
- Jacobs, J. (1969). *Economy of cities*. Penguin books Limited.
- Jofre-Monseny, J., Marín-López, R., & Viladecans-Marsal, E. (2011). The mechanisms of agglomeration : Evidence from the effect of inter-industry relations on the location of new firms. *Journal of Urban Economics*, 70(2-3), 61–74.
- Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica : Journal of the Econometric Society*, 649–670.
- Kalnins, A., & Lafontaine, F. (2013). Too far away? the effect of distance to headquarters on business establishment performance. *American Economic Journal : Microeconomics*, 5(3), 157–79.
- Kazar, B. M., & Celik, M. (2012). *Spatial autoregression (sar) model : Parameter estimation techniques*. Springer Science & Business Media.
- Kimura, F., & Fujii, T. (2003). Globalizing activities and the rate of survival : : Panel data analysis on japanese firms. *Journal of the Japanese and International Economies*, 17(4), 538–560.
- Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G., & Scheike, T. H. (2016). *Handbook of survival analysis*. CRC Press.
- Kleinbaum, D. G., & Klein, M. (2010). *Survival analysis* (Vol. 3). Springer.
- Klepper, S. (2002). Firm survival and the evolution of oligopoly. *RAND journal of Economics*, 37–61.
- Koop, G., Poirier, D. J., & Tobias, J. L. (2007). *Bayesian econometric methods*. Cambridge University Press.
- Koop, G. M. (2003). *Bayesian econometrics*. John Wiley & Sons Inc.

- Kuenne, R. E., & Soland, R. M. (1972). Exact and approximate solutions to the multisource weber problem. *Mathematical Programming*, 3(1), 193–209.
- Kulin, H. W., & Kuenne, R. E. (1962). An efficient algorithm for the numerical solution of the generalized weber problem in spatial economics. *Journal of Regional Science*, 4(2), 21–33.
- Lafontaine, F. (1992). Agency theory and franchising : some empirical results. *The rand journal of economics*, 263–283.
- Landier, A., Nair, V. B., & Wulf, J. (2009). Trade-offs in staying close : Corporate decision making and geographic dispersion. *The Review of Financial Studies*, 22(3), 1119–1148.
- Landström, H. (2005). *Pioneers in entrepreneurship and small business research* (Rapport technique). Springer.
- Lang, L. H., & Stulz, R. (1992). Contagion and competitive intra-industry effects of bankruptcy announcements : An empirical analysis. *Journal of financial economics*, 32(1), 45–60.
- Lasch, E, Robert, F., & Le Roy, F. (2013). Regional determinants of ict new firm formation. *Small Business Economics*, 40(3), 671–686.
- LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- Li, Y., & Ryan, L. (2002). Modeling spatial survival data using semiparametric frailty models. *Biometrics*, 58(2), 287–297.
- Littunen, H. (2001). *The birth and success of new firms in a changing environment*. 261 p. Yhteenveto.
- Liu, X. (2012). *Survival analysis : models and applications*. John Wiley & Sons.
- Lovely, M. E., Rosenthal, S. S., & Sharma, S. (2005). Information, agglomeration, and the headquarters of us exporters. *Regional Science and Urban Economics*, 35(2), 167–191.
- Lynch, S. M. (2007). Modern model estimation part 1 : Gibbs sampling. In *Introduction to applied bayesian statistics and estimation for social scientists* (pp. 77–105). Springer.
- Madrid, M. J. A., García-Tabuenca, A., & Gálvez, C. S. (2015). R & d investments and firm survival across regions. In *Innovation and entrepreneurship in the global economy*. Edward Elgar Publishing.

- Malecki, E. J. (1993). Entrepreneurship in regional and local development. *International regional science review*, 16(1-2), 119–153.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the national cancer institute*, 22(4), 719–748.
- Manzato, G. G., Arentze, T. A., Timmermans, H. J., & Ettema, D. (2011). Exploration of location influences on firm survival rates with parametric duration models. *Transportation research record*, 2245(1), 124–130.
- Marshall, A. (1890). Principles of economics macmillan. *London (8th ed. Published in 1920)*.
- Mas-Verdú, F., Ribeiro-Soriano, D., & Roig-Tierno, N. (2015). Firm survival : The role of incubators and business characteristics. *Journal of Business Research*, 68(4), 793–796.
- Mata, J., & Portugal, P. (1994). Life duration of new firms. *The Journal of Industrial Economics*, 227–245.
- Maté-Sánchez-Val, M., López-Hernández, F., & Fuentes, C. C. R. (2018). Geographical factors and business failure : An empirical study from the madrid metropolitan area. *Economic Modelling*, 74, 275–283.
- Maurel, E., & Sédillot, B. (1999). A measure of the geographic concentration in french manufacturing industries. *Regional Science and Urban Economics*, 29(5), 575–604.
- Mazzotta, F., Maietta, O. W., et al. (2017). *Firm survival and innovation : Knowledge context matters!* (Rapport technique). Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.
- Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6), 1695–1725.
- Mengistae, T. (2006). Competition and entrepreneurs' human capital in small business longevity and growth. *The Journal of Development Studies*, 42(5), 812–836.
- Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society : Series B (Methodological)*, 10(2), 243–251.
- Moran, P. A. (1950). A test for the serial independence of residuals. *Biometrika*, 37(1/2), 178–181.

- Muñoz-Bullón, F., & Cueto, B. (2011). The sustainability of start-up firms among formerly wage-employed workers. *International Small Business Journal*, 29(1), 78–102.
- Myriam Baron, B. (2013). *Modèle de von thünen*. <http://www.hypergeo.eu/spip.php?article566#>. (Accessed :2017-10-13)
- Naldi, L., Nilsson, P., Westlund, H., & Wixe, S. (2021). Amenities and new firm formation in rural areas. *Journal of Rural Studies*, 85, 32–42.
- Neal, Z. P. (2011). The causal relationship between employment and business networks in us cities. *Journal of Urban Affairs*, 33(2), 167–184.
- Neffke, F., Henning, M., Boschma, R., Lundquist, K.-J., & Olander, L.-O. (2011). The dynamics of agglomeration externalities along the life cycle of industries. *Regional studies*, 45(1), 49–65.
- Neffke, F. M., Henning, M., & Boschma, R. (2011). The impact of aging and technological relatedness on agglomeration externalities : a survival analysis. *Journal of Economic Geography*, 12(2), 485–517.
- Nelson, R., & Winter, S. (1982). *An evolutionary theory of economic change* 1982 cambridge. MA : *Belknap Press of Harvard University*.
- Nijkamp, P., Abreu, M. A., et al. (2009). *Regional development theory*. Vrije Universiteit, Faculty of Economics and Business Administration.
- Nogueira, M. Á., Fernández-López, S., Calvo, N., & Rodeiro-Pazos, D. (2018). Firm characteristics, financial variables and types of innovation : influence in spanish firms' survival. *International Journal of Entrepreneurship and Innovation Management*, 22(1-2), 57–79.
- OCDE. (2017). *Panorama de l'entrepreneuriat 2016*. Consulté sur <https://www.oecd-ilibrary.org/content/publication/9789264266346-fr> doi: <https://doi.org/https://doi.org/10.1787/9789264266346-fr>
- Ono, Y. (2006). What do census data tell us about headquarters location? *Economic Development Quarterly*, 20(2), 129–141.
- Pérez, S. E., Llopis, A. S., & Llopis, J. A. S. (2004). The determinants of survival of spanish manufacturing firms. *Review of Industrial Organization*, 25(3), 251–273.
- Perraud, A., MOATI, P., & COUVERT, N. (2008). La survie des entreprises dépend-elle du territoire d'implantation?

- Perreur, J. (1988). La localisation des unités de production. In *Analyse économique spatiale* (pp. 93–158). Presses Universitaires de France.
- Peto, R., & Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society. Series A (General)*, 185–207.
- Ponsard, C. (1988). *L'analyse économique spatiale*.
- Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika*, 65(1), 167–179.
- Regions in Europe, 2020 edition. (s. d.). *Statistics visualised*.
<https://ec.europa.eu/eurostat/fr/web/regions/publications>. (Consulté le 10-08-2021)
- Renski, H. (2011). External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science*, 90(3), 473–502.
- Resende, M., Cardoso, V., & Façanha, L. O. (2016). Determinants of survival of newly created smes in the brazilian manufacturing industry : an econometric study. *Empirical Economics*, 50(4), 1255–1274.
- Roberts, G. O., Gelman, A., Gilks, W. R., et al. (1997). Weak convergence and optimal scaling of random walk metropolis algorithms. *The annals of applied probability*, 7(1), 110–120.
- Roberts, G. O., Rosenthal, J. S., et al. (2001). Optimal scaling for various metropolis-hastings algorithms. *Statistical science*, 16(4), 351–367.
- Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, 94(5), 1002–1037.
- Rosenthal, J. S., et al. (2011). Optimal proposal distributions and adaptive mcmc. *Handbook of Markov Chain Monte Carlo*, 4(10.1201).
- Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of urban economics*, 50(2), 191–229.
- Ruud, P. (2000). *An introduction to classical econometric theory* (Rapport technique). Oxford University Press.
- Schröder, P. J., & Sørensen, A. (2012). Firm exit, technological progress and trade. *European Economic Review*, 56(3), 579–591.

- Seninger, S. F. (1989). *Labor force policies for regional economic development : the role of employment and training programs*. Praeger Publishers.
- Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy. *Small business economics*, 33(2), 141–149.
- Shearmur, R. (2010). Space, place and innovation : a distance-based approach. *The Canadian Geographer/Le Geographe canadien*, 54(1), 46–67.
- Shu, C., & Simmons, S. A. (2018). Firm survival in traded industries : does localization moderate the effects of founding team experience? *Small Business Economics*, 50(3), 643–655.
- Smith, T. E., & LeSage, J. P. (2004). A bayesian probit model with spatial dependencies. In *Spatial and spatiotemporal econometrics*. Emerald Group Publishing Limited.
- Soparnot, R. (2012). *Organisation et gestion de l'entreprise-2e édition*. Dunod.
- Staber, U. (2001). Spatial proximity and firm survival in a declining industrial district : The case of knitwear firms in baden-wuerttemberg. *Regional Studies*, 35(4), 329–341.
- Stearns, T. M., Carter, N. M., Reynolds, P. D., & Williams, M. L. (1995). New firm survival : industry, strategy, and location. *Journal of business venturing*, 10(1), 23–42.
- Stern, R., & Aldrich, H. (1980). The effect of absentee firm control on local community welfare : A survey. *The Economics of Firm Size, Market Structure and Social Performance*. Washington, DC : USGPO.
- Stinchcombe, A. L. (1965). Organizations and social structure. *Handbook of organizations*, 44(2), 142–193.
- Storey, D. J., & Wynarczyk, P. (1996). The survival and non survival of micro firms in the uk. *Review of Industrial Organization*, 11(2), 211–229.
- Strauss-Kahn, V., & Vives, X. (2009). Why and where do headquarters move? *Regional Science and Urban Economics*, 39(2), 168–186.
- Syversveen, A. R. (1998). Noninformative bayesian priors. interpretation and problems with construction and applications. *Preprint statistics*, 3(3), 1–11.
- Tarone, R. E., & Ware, J. (1977). On distribution-free tests for equality of survival distributions. *Biometrika*, 64(1), 156–160.

- Thamrin, S. A., Jaya, A. K., Mengersen, K., et al. (2021). Bayesian spatial survival modelling for dengue fever in makassar, indonesia. *Gaceta Sanitaria*, 35, S59–S63.
- Thompson, P. (2006). Patent citations and the geography of knowledge spillovers : Evidence from inventor-and examiner-added citations. *The Review of Economics and Statistics*, 88(2), 383–388.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography*, 46(sup1), 234–240.
- Van Kranenburg, H. L., Palm, F. C., & Pfann, G. A. (2002). Exit and survival in a concentrating industry : The case of daily newspapers in the netherlands. *Review of Industrial Organization*, 21(3), 283–303.
- Van Stel, A., & Suddle, K. (2008). The impact of new firm formation on regional development in the netherlands. *Small Business Economics*, 30(1), 31–47.
- Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3), 439–454.
- Vergin, R. C., & Rogers, J. D. (1967). An algorithm and computational procedure for locating economic facilities. *Management Science*, 13(6), B–240.
- Vivel-Búa, M., & Lado-Sestayo, R. (2021). Contagion effect on business failure : A spatial analysis of the hotel sector. *Journal of Hospitality & Tourism Research*, 10963480211023804.
- Wennberg, K., & Lindqvist, G. (2010). The effect of clusters on the survival and performance of new firms. *Small Business Economics*, 34(3), 221–241.
- Witt, U. (2016). *Rethinking economic evolution : Essays on economic change and its theory*. Edward Elgar Publishing.
- Zhang, L., Zhang, H., & Yang, H. (2018). Spatial distribution pattern of the headquarters of listed firms in china. *Sustainability*, 10(7), 2564.
- Zhou, H., & Hanson, T. (2015). Bayesian spatial survival models. *Nonparametric Bayesian Inference in Biostatistics*, 215–246.