



**HAL**  
open science

# L'impact des infrastructures de transport sur les décisions de localisation et la longévité des entreprises

Marie-Pier Champagne

► **To cite this version:**

Marie-Pier Champagne. L'impact des infrastructures de transport sur les décisions de localisation et la longévité des entreprises. Economies et finances. Université Bourgogne Franche-Comté; Université Laval (Québec, Canada), 2024. Français. NNT: 2024UBFCG003 . tel-04910165

**HAL Id: tel-04910165**

**<https://theses.hal.science/tel-04910165v1>**

Submitted on 24 Jan 2025

**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.

## THÈSE DE DOCTORAT

*Pour obtenir le grade de :*

DOCTEUR DE L'UNIVERSITÉ DE BOURGOGNE

UFR Droit, Sciences Économiques et Politique

École doctorale n°593 Droit, Gestion, Économie et Politique (DGEP)

DOCTORAT DE SCIENCES ÉCONOMIQUES

Réalisé en cotutelle

*Présentée par :*

Mme CHAMPAGNE Marie-Pier

## **L'IMPACT DES INFRASTRUCTURES DE TRANSPORT SUR LES DÉCISIONS DE LOCALISATION ET LA LONGÉVITÉ DES ENTREPRISES**

Co-direction internationale de thèse :      Jean DUBÉ  
Diègo LEGROS

Thèse présentée et soutenue à Dijon, le 3 juillet 2024

*Soutenue publiquement devant un jury composé de :*

Mme BREUILLÉ Marie	Directrice de recherche, INRAE	Présidente du jury
M. BOUZOUINA Louafi	Directeur de recherche, ENTPE	Rapporteur
Mme TRAVERS Muriel	Professeure, Université de Nantes	Rapporteur
M. VILLENEUVE Dominic	Professeur, Université Laval (Québec)	Membre du jury
M. DUBÉ Jean	Professeur, Université Laval (Québec)	Directeur de thèse
M. LEGROS Diègo	Professeur, Université de Bourgogne	Directeur de thèse

**Titre :** L'impact des infrastructures de transport sur les décisions de localisation et la longévité des entreprises

**Mots clés :** Infrastructure de transport ; Localisation ; Commercial ; Méta-analyse ; Modèle de survie

**Résumé :** Les infrastructures de transport sont au cœur des villes et de leur développement. La distribution spatiale des activités économiques dépend de nombreux facteurs et tend à être influencée par les caractéristiques intrinsèques de l'entreprise, mais aussi par la présence d'économies d'agglomération et d'infrastructures de transport. La théorie de la localisation présente ces facteurs comme des piliers sur lesquels reposent les décisions de localisation des entreprises, et ultimement la structure urbaine des villes. Cette thèse s'intéresse particulièrement aux impacts qu'ont ces infrastructures sur les décisions de localisation des entreprises et leurs performances.

La théorie néo-classique économique présente les décisions de localisation d'une entreprise comme le résultat d'un problème d'optimisation de la fonction de profits. Dans une perspective de développement commercial et de pérennité économique, l'étude de la performance des entreprises, en fonction d'une localisation jugée optimale, s'avère intéressante et complémentaire aux études de localisation.

Une analyse approfondie de la littérature, basée sur une revue systématique, permet de réaliser un portrait de la littérature inhérente au rôle des transports dans les décisions de localisation des entreprises. La méta-analyse de ces résultats suggère que les effets d'une infrastructure de transport sur les décisions de localisation des entreprises sont positifs, ou dans le pire des cas, non-significatifs. L'évaluation des impacts causaux des infrastructures de transport en commun sur la probabilité de survie des entreprises représente une originalité de ce projet de thèse. Les résultats obtenus sont complémentaires aux divers angles déjà abordés dans la littérature.

Empiriquement, un modèle de survie de Cox est appliqué à deux études de cas, soit le métro de Montréal (Canada) et le tramway de Dijon (France). Le métro de Montréal a connu sa plus récente extension en 2007 avec le prolongement de la ligne orange. Cette dernière permet de connecter l'île de Montréal à sa banlieue nord, la ville de Laval. Le tramway de Dijon, quant à lui, est proposé en 2008 comme une alternative durable à un réseau d'autobus saturé. Les deux lignes de tramway ont été mises en service en 2012 et desservent la ville de Dijon d'est en ouest (T1), et du nord au sud (T2).

Une étude approfondie des effets d'anticipation et d'ajustement est effectuée à partir des différentes périodes (annonce, construction, et mise en opération) menant à l'instauration de ces systèmes de transport collectif. Les impacts sont évalués en fonction de la distance aux stations afin d'identifier les zones les plus touchées par la mise en place de ces services.

Les résultats suggèrent que les impacts se matérialisent majoritairement après la mise en opération des services de transport en commun. Dans un contexte nord-américain, un impact positif sur la probabilité de survie est observé pour une distance allant jusqu'à 1 250 mètres de la station la plus proche. Dans un contexte européen, basé uniquement sur les petites entreprises, les résultats identifient un impact positif, cependant, il semble plus concentré spatialement, soit à moins de 500 mètres des stations.

**Title :** The impact of transport infrastructure on firm's location decision and longevity

**Keywords :** Transport infrastructure; Location; Commercial; Meta-analysis; Survival model

**Abstract :** Transport infrastructures are fundamental to cities and their development. The spatial distribution of economic activities depends on numerous factors and tends to be influenced by the intrinsic characteristics of the firms, the presence of agglomeration economies and transport infrastructures. Location theory presents these factors as pillars upon which commercial location decisions are based, ultimately shaping the urban structure of cities. This thesis particularly focuses on the impacts of transport infrastructures on firms' location decisions and their performances.

The neoclassical economic theory presents a firm's location decisions as the outcome of the profit function optimization problem. From a perspective of commercial development and economic sustainability, studying firms' performance, based on an optimal location, proves to be both interesting and complementary to location studies.

An analysis of the literature, based on a systematic review, produces an exhaustive summary of the literature regarding the role of transportation on firms' location decisions. The meta-analysis of these results suggests that the effects of transportation infrastructure on firms' location decisions are positive, or in the worst case, non-significant.

The originality of this thesis relies on the evaluation of the causal impact of public transit on firms' probability of survival. The obtained results are complementary to the literature on the impact of public transit from a commercial perspective and, therefore, insert themselves easily into the existing literature.

Empirically, a Cox survival model is applied to two case studies: the Montréal Metro (Canada) and the Dijon Tramway (France). On the one hand, the Montréal Metro experienced its most recent extension in 2007 with the extension of the Orange Line, connecting the Island of Montréal to its northern suburb, the city of Laval. The Dijon Tramway, on the other hand, was proposed in 2008 as a sustainable alternative to a saturated bus network. Both tram lines were put into service in 2012, serving the city of Dijon from east to west (T1) and from north to south (T2).

To account for potential anticipation and adjustment effects, the studies are conducted using the temporal periods (announcement, construction, and operation) leading to the implementation of these public transit systems. The impacts are evaluated based on proximity to the public transit stations to identify the spatial areas that were mostly affected by the opening of these services.

Results suggest that the impact materialized mainly after the opening of the public transit services. In a North American context, a positive impact on firms' survival probability is observed within 1,250 meters of the closest station. In a European context, based solely on small businesses, the results also identify a positive impact, however, it appears to be more spatially concentrated, i.e., within 500 meters of the stations.

# Table des matières

Résumé .....	i
Abstract.....	ii
Liste des figures.....	v
Liste des tableaux .....	vi
Liste des abbreviations .....	vii
Remerciements .....	viii
Avant-propos .....	ix
Introduction générale.....	1
CHAPITRE 1 : The impact of transport infrastructure on firms' location decision: A meta-analysis based on a systematic literature review .....	8
Résumé.....	8
Abstract .....	9
Introduction .....	10
Conceptual and theoretical background .....	11
Methodology .....	14
Descriptive Statistics.....	16
Meta-Analysis .....	24
Concluding remarks .....	27
CHAPITRE 2 : Cadre théorique.....	30
Maximisation des profits.....	30
Transport et fonction de profit .....	32
Économies d'agglomération.....	34
Conclusion partielle .....	36
CHAPITRE 3 : Cadre méthodologique.....	38
Relations statistiques.....	38
Problème d'endogénéité des variables explicatives .....	39
Analyses causales.....	41
Le contrefactuel.....	41
Expériences randomisées .....	42
Endogénéité du traitement.....	43
Inférence causale.....	43
Ajustement statistique : Méthode d'appariement .....	43
Quasi-Expérience : Approche de Différence-en-Différences.....	45
CHAPITRE 4 : Standing strong? The causal impact of metro stations on service firms' survival .....	51
Résumé.....	51
Abstract .....	52

Introduction .....	53
Related literature .....	54
Methodology .....	58
Case study and data .....	60
Montréal Metro .....	60
Data and Descriptive Statistics.....	63
Results .....	68
Model 1 – All Economic Sectors .....	69
Model 2 – Specific Economic Sectors .....	75
Conclusion.....	76
<b>CHAPITRE 5 : Revitalizing Commute: The Impact of Bus-to-Tramway Shift on Small Businesses’ Survival.....</b>	<b>79</b>
Résumé.....	79
Abstract .....	80
Introduction .....	81
Related literature .....	82
Methodology .....	85
Case study and data.....	88
Dijon’s public transit network.....	88
Data .....	89
Results .....	96
Robustness tests .....	101
Conclusion.....	109
Conclusion générale .....	111
Bibliographie .....	115
Annexe du Chapitre 1 .....	130
Annexe du Chapitre 4.....	146
Annexe du Chapitre 5 .....	165

## Liste des figures

Figure 1.1 - Selection process of papers.....	16
Figure 1.2 - Concentration of papers per year .....	17
Figure 1.3 - Geographical distribution of the case studied .....	17
Figure 1.4 - Transport infrastructure studied in the selected papers.....	19
Figure 1.5 - Economic sectors studied in the selected papers.....	20
Table 1.4 - Meta-analysis results .....	26
Figure 2.1 - Distribution non-uniforme des activités économiques.....	30
Figure 3.1 - Représentation graphique de l’approche de différence-en-différences .....	47
Figure 3.2 - Impact par période et cumulé dans la spécification DiD .....	49
Figure 4.1 - The submitted extensions of the orange and blue lines in October 1998.....	61
Figure 4.2 - Temporal evolution of the extension of the orange metro line in Laval .....	62
Figure 5.1 - Dijon’s BRT network before and after the tramway.....	89
Figure 5.2 - Tramway lines and unchanged BRT lines .....	93
Figure 5.3 - Spatial distribution and group division based on critical distance .....	95
Figure 5.4 - Complete studied timeline in relation to Dijon’s tramway .....	103
Figure A1.3 - Evaluating the potential publication bias .....	141
Figure A4.2.1 - Kaplan-Meier survival plots, for all economic sectors.....	147
Figure A4.2.2 - Kaplan-Meier survival plots, for retail activities .....	148
Figure A4.2.3 - Kaplan-Meier survival plots, for FIRE and KIBS.....	149
Figure A4.2.4 - Kaplan-Meier survival plots, for local activities .....	150
Figure A4.3 - Cox-Snell Residuals (Model 1).....	151
Figure A4.4 - Cox-Snell Residuals (Model 2).....	152
Figure A5.2 - Cox-Snell residuals: Testing the goodness-of-fit of the model.....	167

# Liste des tableaux

Table 1.1 - Keywords and their derivation .....	15
Table 1.2 - Data type used based on the methodological approach.....	20
Table 1.3 - Summary of the selected papers' results: transport infrastructure's effects on firms' location decision based on the economic sectors and the type of infrastructure .....	23
Table 4.1 - Firms' opening and closure (descriptive statistics) .....	64
Table 4.2 - Distribution among the interaction variables (descriptive statistics).....	66
Table 4.3 - Cox regression results for all economic sectors (Model 1) .....	71
Table 4.4 - Cox regression results for all specific economic sectors (Model 2) part 1.....	73
Table 4.4 - Cox regression results for all specific economic sectors (Model 2) part 2.....	74
Table 5.1 - Number of closed and censored firms in the sample.....	94
Table 5.2 - Cox model results (part 1).....	99
Table 5.2 - Cox model results (part 2).....	100
Table 5.3 - Cox model results using distance-based specification for larger critical distances .....	102
Table 5.4 - Cox model results for a shorter period, January 2006 – December 2015 inclusively (part 1) ..	105
Table 5.4 - Cox model results for a shorter period, January 2006 – December 2015 inclusively (part 2) ..	106
Table 5.5 - Cox model results for a shorter period, January 2006 – December 2017 inclusively (part 1) ..	107
Table 5.5 - Cox model results for a shorter period, January 2006 – December 2017 inclusively (part 2) ..	108
Table A1.1 - Selected papers used in the systematic review and meta-analysis.....	130
Table A1.2 - Main conclusions of the 52 selected papers .....	133
Table A1.4.1 - Ordered-logit regression, pairing negative and significant effects with non-significant effects. ....	142
Table A1.4.2 - Ordered-logit regression, using positive and non-significant effects (discarding the negative effect) .....	143
Table A1.5.1 - Meta-analysis results, using cross-product multiplication accounting for freight transportation .....	144
Table A1.5.2 - Meta-analysis results, using cross-product multiplication accounting for passengers' transportation.....	145
Table A4.1 - Repartition of economic sectors in 15 categories .....	146
Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 1 .....	153
Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 2 .....	155
Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 3 .....	157
Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 1 .....	159
Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 2 .....	161
Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 3 .....	163
Table A5.1 - Placebo test: Testing the parallel trend assumption (part 1).....	165
Table A5.1 - Placebo test: Testing the parallel trend assumption (part 2).....	166



## Liste des abbreviations

AMM	Alonso-Muth-Mills
ATE	<i>Average Treatment Effect</i>
ATT	<i>Average Treatment Effect on the Treated</i>
BRT	<i>Bus Rapid Transit</i>
CAQ	Coalition Avenir Québec
CBD	<i>Central Business District</i>
CRSH	Conseil de Recherches en Sciences Humaines
DiD	Différence-en-différences / <i>Difference-in-differences</i>
EAC	<i>Economic Activity Code</i>
FIRE	<i>Finance, Insurance and Real Estate</i>
GES	Gaz à Effet de Serre
GMM	<i>Generalized Method of Moments</i>
HR	<i>Hazard Ratio</i>
HI	<i>Herfindahl Index</i>
IBSS	<i>International Bibliography of the Social Sciences</i>
IF	<i>Impact Factor</i>
iid	Indépendant et identiquement distribué
KIBS	<i>Knowledge Intensive Business Services</i>
Lianes	Ligne à niveau élevé de service
LQ	<i>Location Quotient</i>
LRT	<i>Light Rail Transit</i>
NEQ	Numéro des Entreprises du Québec
OLS	<i>Ordinary Least Square</i>
PH	<i>Proportional Hazards</i>
PHA	<i>Proportional Hazard Assumption</i>
REQ	Registraire des entreprises du Québec
SEM	<i>Spatial Error Model</i>
SLX	<i>Spatial lagged explanatory variable model</i>
STM	Société des Transports de Montréal
TWFE	<i>Two-Way Fixed Effect</i>
VIF	<i>Variance Inflated Factor</i>
WTP	<i>Willingness-to-pay</i>

## Remerciements

Avant toute chose, j'aimerais remercier mes directeurs de recherche, M. Jean Dubé et M. Diego Legros. Leur support constant, leur efficacité et leurs excellents conseils m'ont permis de mener à bien ce projet de thèse. Je les remercie également de la liberté de recherche dont j'ai pu bénéficier et de la confiance dont ils m'ont témoigné. Je me considère choyée d'avoir été aussi bien encadrée tout au long de mon parcours doctoral. Une part de ma réussite leur revient définitivement.

Je tiens à remercier spécialement M. Jean Dubé, qui m'a encadré dès la maîtrise à titre de co-directeur. Ses retours rapides et perspicaces, et son optimisme ont fait une différence indéniable à la fois dans mon développement en tant que chercheuse, mais aussi dans la qualité de mes projets de recherche. Je termine en soulignant l'importance qu'ont joué ses multiples lettres de recommandation dans l'obtention de mes bourses de recherche.

Je profite également de cette section pour remercier le Conseil de Recherches en Sciences Humaines (CRSH) pour le support financier accordé par l'intermédiaire de la bourse d'études supérieures du Canada Vanier. Cette généreuse bourse de recherche m'a permis de me consacrer pleinement à cet ambitieux projet qu'est le doctorat.

Enfin, j'aimerais remercier chaleureusement mon copain, Christian Landry, pour son soutien et ses judicieux conseils durant mon processus de rédaction. Son regard externe et ses commentaires pertinents ont clairement permis d'élever le niveau final de mon projet de thèse. Je le remercie aussi pour sa patience et sa compréhension, surtout durant les dernières semaines de rédaction. Pour terminer, je remercie ma famille, mon père Yves, ma mère Julie et mes deux sœurs, Audrey et Catherine, ainsi que tous mes amis qui m'ont encouragé et soutenu, de près ou de loin, pendant mon parcours.

Une mention particulière à mon partenaire de recherche, Diego Andres Cardenas Morales, nos discussions informelles illustrent à merveille l'importance du partage des connaissances et la nécessité du travail en présentiel.

Merci à tous. Sur ce, je vous souhaite une bonne lecture.

## Avant-propos

Ce projet de thèse prend la forme d'une thèse par article et compte trois articles desquels je suis première auteure. Le processus de modélisation et la rédaction de tous les articles ont été réalisés sous la supervision de mes directeurs de recherche.

- 1) *The impact of transport infrastructure on firms' location decision: A meta-analysis based on a systematic literature review.*

L'article a été soumis au journal *Transport Policy* le 27 avril 2021. À la demande de l'éditeur, des modifications ont été effectuées en mai 2022, puis en novembre 2022. L'article a été officiellement publié en février 2023 et un seul coauteur figure sur cette publication, M. Jean Dubé (co-directeur, Université Laval).

- 2) *Standing strong? The causal impact of metro stations on service firms' survival.*

L'article a été soumis au journal *Transportation Research Part A : Policy and Practice* le 2 février 2023. À la demande de l'éditeur, des modifications ont été effectuées en novembre 2023, puis en janvier 2024. L'article a été officiellement publié en mars 2024. Les deux coauteurs de cet article sont les directeurs de ce projet de thèse, M. Jean Dubé (Université Laval) et M. Diègo Legros (Université de Bourgogne).

- 3) *Revitalizing Commute: The Impact of Bus-to-Tramway Shift on Small Businesses' Survival.*

L'article a été soumis au journal *Small Business Economics* en avril 2024. Comme pour le second article, les deux coauteurs sont les directeurs de ce projet de thèse, M. Jean Dubé (Université Laval) et M. Diègo Legros (Université de Bourgogne).

# Introduction générale

Les dynamiques collectives sont le reflet des décisions individuelles. Par leurs choix, les agents économiques influencent la répartition spatiale des populations et des activités, affectant directement le développement local et le portrait global de la société. Dans une perspective d'aménagement du territoire, les décisions de localisation des individus et des entreprises s'avèrent largement influencées par les infrastructures de transport (Fujita, 1989).

Ces infrastructures augmentent l'accessibilité, facilitent les déplacements et diminuent les coûts associés au transport (Glaeser et al., 2008; Glaeser et Kohlhase, 2004). D'une part, les ménages cherchent à se localiser dans des lieux qui bénéficient de plus d'opportunités d'emploi et de consommation. La localisation des infrastructures de transport influence la structure des déplacements et se traduit généralement par une plus grande volonté de payer. Or, l'offre de sol est fixe (par définition), ce qui implique un ajustement limité (du côté de l'offre) face à une hausse de la demande pour ces localisations. Cette demande excédentaire est internalisée par le marché et engendre la hausse des prix immobiliers (Schoenmaker et Van der Vlist, 2015). D'autre part, les entreprises sont attirées par les infrastructures de transport pour des raisons similaires, telles que profiter d'une meilleure accessibilité et diminuer ses coûts de transport. Une localisation près du centre de son marché et facile d'accès pour les travailleurs et les consommateurs s'avère être un avantage non-négligeable pour la performance des entreprises (Fujita, 1989).

Évidemment, il existe une relation d'interdépendance entre les choix de localisation des entreprises et des ménages. Les entreprises représentent des opportunités d'emploi et de consommation pour les ménages. À l'inverse, les ménages forment, pour les entreprises, un bassin de consommateurs et de main-d'œuvre. Favoriser le contact entre les agents économiques est mutuellement bénéfique et largement facilité par la présence d'infrastructure de transport. Une meilleure accessibilité se reflète dans le marché de l'emploi, dans les opportunités de consommation des ménages, ainsi que dans les parts de marché des entreprises. Cette relation entre les ménages et les entreprises est internalisée dans les décisions de localisation des agents économiques. Néanmoins, le développement précis de cette relation découle du paradoxe de l'œuf et de la poule, où la décomposition des

effets est ardue, surtout lorsque l'intérêt est porté sur les dynamiques locales (Dubé et al., 2017; Hoogstra et al., 2017).

La présence d'infrastructures de transport connectant les périphéries au centre-ville joue un rôle structurant dans l'expansion des villes (Dubé et al., 2023a) et l'émergence de centres secondaires, stimulant l'apparition de villes dites polycentriques (Baum-Snow, 2007a, 2007b).

Les infrastructures de transport génèrent des effets externes pour les agents économiques et un arbitrage se forme entre les externalités positives et négatives. D'une part, les infrastructures de transport diminuent les coûts de transport et augmentent l'accessibilité. Elles peuvent également engendrer une plus grande visibilité et un meilleur achalandage, en augmentant les opportunités de consommation pour les ménages (Glaeser et Kohlhase, 2004; Chatman et Noland, 2011). D'autre part, la popularité des infrastructures de transport peut mener à une saturation du système, engendrant certaines nuisances telles qu'une hausse de la pollution (visuelle, sonore et de l'air), du stress, des troubles de santé, etc. (Beaudoin et al., 2015; Brugge et al., 2007; Zhong et al., 2017).

Dans une optique strictement commerciale, le secteur d'activités de l'entreprise et le type d'infrastructure de transport ont un impact sur l'effet net observé, soit la différence entre les avantages et les inconvénients. Par exemple, une entreprise manufacturière qui demande énormément de capital physique a intérêt à se localiser en périphérie des centres, où les terrains sont moins chers. De plus, la proximité aux autoroutes, chemins de fer et généralement aéroports représente un attrait important pour ces localisations périphériques (Holl, 2004a; 2004b). À l'inverse, une entreprise de service préfère se localiser au centre-ville afin d'être facilement accessible pour sa clientèle, mais aussi pour ses travailleurs. Le secteur tertiaire est dépendant du capital humain. Ainsi, les entreprises de ce secteur doivent être facilement accessibles, et donc les centres-villes, généralement desservis par des réseaux de transport collectif, s'avèrent plus attrayants (Credit, 2018; 2019).

La forte densité commerciale représente aussi un avantage pour les entreprises. Elle génère des économies d'agglomération, soit des externalités positives qui émergent de la proximité entre les entreprises. Le partage du bassin de main-d'œuvre (qualifiée) et de consommateurs,

des connaissances (*knowledge spillovers*) et des aménités locales sont des exemples d'externalités positives découlant de la concentration spatiale d'activités économiques, semblables (Marshall, 1891) ou diversifiées (Jacobs, 1969).

En somme, les décisions de localisation des entreprises sont influencées par trois facteurs prédominants; le type d'infrastructure de transport, son secteur d'activités et la présence d'agglomération spatiale. En ce sens, l'étude de l'impact des infrastructures de transport, au sens large, sur la distribution spatiale des activités commerciales est importante afin de comprendre (et potentiellement influencer) le développement et la structure des villes et des régions.

Ainsi, une première question de recherche s'impose : quels sont les impacts des infrastructures de transport sur les décisions de localisation des entreprises ? Le premier chapitre de cette thèse cherche à répondre à cette question à l'aide d'études recensées dans la littérature. Une analyse descriptive et une méta-analyse permettent d'identifier les principaux constats existants. Une attention particulière est portée au type d'infrastructure et aux secteurs d'activités étudiés, mais aussi sur les données mobilisées et les différentes approches méthodologiques.

La distribution spatiale des activités commerciales ainsi que l'offre de transport influencent les dynamiques locales. Par exemple, le phénomène d'étalement urbain est largement stimulé par les infrastructures autoroutières mises en place. Dans une perspective de développement durable, l'omniprésence des autoroutes s'avère problématique, notamment pour les externalités négatives que ce type d'infrastructure génère. Un trop grand flux d'automobiles, menant inévitablement à des problèmes de congestion, s'avère grandement nocif pour l'environnement, mais également pour la santé des individus (Brugge et al., 2007; Beaudoin et al., 2015).

L'élargissement ou l'extension du réseau routier peut sembler être une solution simple aux problèmes de congestion. Or, contrairement à la pensée populaire, ajouter des axes routiers n'est pas une solution permanente, il s'agit au mieux d'un ajustement temporaire au problème de congestion (Downs, 1962; Duranton et Turner 2011; 2012). En effet, la loi fondamentale de la congestion stipule qu'une hausse de la capacité routière entraîne une augmentation plus

que proportionnelle du nombre de déplacements (demande induite). Une saturation du réseau routier bonifié est à prévoir quelques années plus tard (Downs, 1962). Les mécanismes sous-jacents à cette loi sont relativement intuitifs : l'ajout d'axes routiers engendre une diminution des coûts (du temps) de transport, pour un déplacement donné, et stimule l'étalement urbain. En d'autres termes, un plus grand réseau routier encourage les agents économiques à se relocaliser en périphérie des centres urbains, où les prix immobiliers sont généralement plus faibles (Alonso, 1964; Muth, 1969; Mills, 1972).

En somme, l'extension d'un réseau autoroutier n'est pas une solution viable, à long terme, au problème de congestion engendré par la concentration spatiale des populations et des activités économiques. Par ailleurs, les externalités négatives, causées par la saturation des réseaux routiers en milieu urbain, génèrent des enjeux de société où la protection de la planète doit être mise de l'avant.

Dans cette optique, une infrastructure de transport en commun peut être une piste de solution intéressante. Néanmoins, de nombreux débats de société émergent quant à la mise en place de réseaux structurants de transport collectif. En effet, ces projets d'envergure peuvent être source de mécontentement, surtout dans un contexte nord-américain, où une importante dépendance à l'automobile est observée. La mise en place d'infrastructure de transport en commun peut également s'avérer onéreuse et les chantiers de construction d'importantes nuisances urbaines.

Afin d'informer les décideurs publics et d'appuyer leur prise de décision, une compréhension approfondie des impacts des infrastructures de transport en commun est nécessaire. Il est donc primordial d'analyser et de quantifier leurs impacts (économiques, sociaux, environnementaux, etc.), afin de maximiser les retombées positives, en fonction des objectifs ciblés. Les objectifs peuvent être multiples et varier d'un endroit à l'autre. Ils peuvent viser, par exemple, la densification, la décongestion routière, et/ou la revitalisation de centres urbains.

Dans cette logique, la suite de la thèse s'articule spécifiquement autour des transports collectifs et du développement commercial. L'étude de l'impact des infrastructures de transport en commun a été effectuée sous plusieurs angles dans la littérature : création

d'emploi (Yao et Hu, 2020; Credit 2018; 2019; Matas et al., 2020), (re)localisation d'entreprises (Mejia-Dorantes et al., 2012; Beckerich et al., 2016; Iseki et Jones, 2018; Champagne et Dubé, 2023), occupation du sol et constructions immobilières (Hurst et West, 2014; Champagne et al., 2022), prix immobiliers (Dubé et al., 2011; 2013; 2014; Cordera et al., 2019), économies d'agglomération (Chatman et al., 2016; Chatman et Noland, 2011), etc. Cependant, la littérature inhérente à la performance commerciale est encore relativement peu documentée, surtout dans une optique de survie. Ainsi, l'étude des impacts des infrastructures de transport en commun sur la longévité des entreprises s'avère pertinente et s'insère aisément dans la littérature.

Une seconde question de recherche émerge : quel est l'impact des infrastructures de transport en commun sur la performance des entreprises? Et plus spécifiquement sur leur survie ? Sous l'hypothèse que les stations de transport en commun augmentent la visibilité et l'achalandage pour les commerces, une plus grande probabilité de survie devrait être observée pour les entreprises localisées à proximité. Afin de tester cette hypothèse, un modèle de survie de Cox est appliqué à deux études de cas distinctes : (1) le cas du métro de Montréal, et (2) le cas du tramway de Dijon.

Le métro de Montréal a connu l'extension de la ligne Orange en 2007. Trois nouvelles stations ont permis de connecter la ville, et l'île, de Laval à l'île de Montréal. Il s'agit des premières stations de métro à desservir la banlieue nord hors de l'île de Montréal. L'annonce simultanée des extensions de la ligne Orange et de la ligne Bleue, en 1998, crée un scénario d'étude intéressant, puisque seule l'extension de la ligne Orange s'est concrétisée. Le report du projet de la ligne Bleue permet ainsi de créer un contrefactuel intéressant pour l'analyse d'impacts causaux. Ce scénario laisse place à un comparatif crédible et à l'étude de l'impact du prolongement d'une ligne de métro sur la survie des entreprises.

Proposé en 2008, le tramway de Dijon visait à réduire la congestion routière et encourager le développement immobilier. Il est officiellement mis en service à la fin de l'année 2012, après le début de sa construction en 2010. Ce projet de transport structurant compte 2 lignes qui couvrent le territoire d'est en ouest et du nord au sud. Il dessert les communes de Dijon, Chenôve (au sud) et Quetigny (à l'est). La substitution d'un service d'autobus rapide par un réseau de tramway est un contexte d'étude intéressant, dans l'optique où les effets étudiés



sur la performance des entreprises sont relatifs à une amélioration de l'offre de transport en commun. Les lignes d'autobus articulées inchangées et en opération avant même l'introduction du tramway constitue un contrefactuel crédible pour l'identification d'un potentiel impact causal.

Cette thèse par article est divisée en six (6) chapitres. Le premier chapitre présente une revue de la littérature liée aux impacts des infrastructures de transport sur les décisions de localisation des entreprises. Ce chapitre propose, d'une part, une revue systématique de la littérature, et d'autre part, une méta-analyse des résultats empiriques contenus dans les articles étudiés. Les principales conclusions de la littérature sur les décisions de localisation des entreprises, en fonction du secteur d'activités et du type d'infrastructure étudié, sont discutées dans ce chapitre. Enfin, le type de données et l'approche méthodologique mobilisés dans la littérature sur les impacts des infrastructures de transport sur les décisions de localisation des entreprises sont également abordés.

Le second chapitre introduit le cadre théorique autour duquel s'articule ce projet de thèse. Basé sur les fondements économiques néo-classiques, ce chapitre introduit le problème d'optimisation d'une entreprise, par la maximisation de ses profits, et par dualité la minimisation de ses coûts. La localisation d'une entreprise est le reflet, en théorie, de ce processus d'optimisation. Outre les caractéristiques intrinsèques de l'entreprise, les coûts de transport et la présence d'économies d'agglomération sont des facteurs d'influence permettant d'expliquer les choix de localisation des entreprises.

Un cadre méthodologique constitue le troisième chapitre de la thèse. Afin d'étudier une relation de cause à effet entre les infrastructures de transport et la performance des entreprises, des stratégies d'identification causale doivent être mobilisées. Les analyses économétriques communes ne permettent pas nécessairement d'identifier des effets causaux. Les principaux problèmes liés à l'endogénéité des variables et à une assignation non-aléatoire du traitement sont discutés dans ce chapitre. Ce dernier point est d'autant plus important à adresser, puisque les politiques publiques sont rarement assignées de manière aléatoire. La spécification de type différence-en-différences, mobilisée pour la suite des analyses empiriques, est présentée en détail.

Le chapitre quatre propose une première analyse causale de la performance des entreprises autour du métro de Montréal. Il s'agit d'une application empirique, sous la forme d'un article scientifique, évaluant l'impact de l'extension d'une ligne de métro sur la survie des entreprises à Montréal (Canada).

Le cinquième chapitre présente une seconde application empirique sur l'impact causal d'une infrastructure de transport en commun sur la performance des entreprises. Plus précisément, ce chapitre étudie l'impact de la substitution d'un réseau d'autobus rapide par un service de tramway sur la survie des entreprises à Dijon (France).

Enfin, le dernier chapitre conclut la thèse. Ce chapitre synthétise les points les plus pertinents abordés dans la thèse. Les principaux résultats obtenus dans les applications empiriques sont également résumés.

# **CHAPITRE 1 : The impact of transport infrastructure on firms' location decision: A meta-analysis based on a systematic literature review**

## **Résumé**

Cet article propose d'explorer la littérature existante sur l'impact des infrastructures de transport sur les décisions de localisation des entreprises. À l'aide de trois bases de données distinctes (*EconLit*, *International Bibliographie of the Social Sciences* et *Web of Science*), cette étude identifie 338 articles potentiels, dont 52 sont retenus pour l'analyse. L'article présente une analyse descriptive complète des articles sélectionnés, ainsi qu'une méta-analyse des résultats identifiés dans les articles sélectionnés. Les résultats de la méta-analyse suggèrent un impact plus positif et significatif des infrastructures de transport sur les activités financières, d'assurance et immobilières (FIRE). La méta-analyse met également en évidence l'importance de considérer les dimensions spatiales et temporelles dans l'analyse des impacts des infrastructures de transport sur les décisions de localisation commerciale. L'omission d'une de ces dimensions pourrait conduire à une surestimation de l'effet, et donc à un biais dans les résultats observés.

**Mots-clés :** Économie des transports; Infrastructure de transport; Décisions de localisation; Décisions commerciales; Aménagement du territoire; Méta-analyse.

## **Abstract**

This paper proposes to explore the existing literature on the impact of transportation infrastructure on firms' location decision. Using three distinct databases (EconLit, International Bibliography of the Social Sciences, and Web of Science), the research identifies 338 potential papers, with 52 of them retained for the analysis. The study presents a complete descriptive analysis of the selected papers, as well as a meta-analysis based on all the reported statistics. From a quantitative perspective, the meta-analysis results suggest a more positive and significant impact of transport infrastructure on finance, insurance, and real estate activities (FIRE). The meta-analysis also highlights the importance of taking into account the spatial and temporal dimensions of location decisions when evaluating the impact of transport infrastructure. The omission of one of those dimensions might lead to an overestimation of the effect, and therefore bias in the observed results.

**Keywords:** Transportation Economic; Transport Infrastructure; Location Decision; Firms' Decision; Land Use Patterns; Meta-Analysis.

## **Introduction**

The development and implementation of new transport infrastructure has a direct influence on the location decisions of firms and households. On the one hand, transport infrastructure might raise accessibility, reduce transportation costs and attract people (centripetal force). Therefore, firms might be attracted by new transport infrastructure as it affects transport costs and thus total production costs (Credit, 2018). On the other hand, concentrating people, cars and trucks within a specific area also generates negative externalities, such as higher traffic, air pollution, noise pollution, etc., which translates into a diffusion (or centrifugal) process to avoid negative impact. The impact of transport infrastructure is uncertain for firms and households, since new transport infrastructure introduces those two opposite forces (Higgins and Kanaroglou, 2018). The differences between positive and negative impacts, i.e., the net impact, depends on the forces at play between these two opposite effects.

So far, the literature on the impact of new transport infrastructure has been widely studied from different perspectives such as economic output, productivity, wages and the real estate market. For example, the hedonic pricing model has been a popular approach to evaluate the willingness-to-pay (WTP) to be located close to new transport infrastructure. Moreover, the empirical analysis using the revealed preference method has been mainly devoted to the residential market, while applications on commercial and industrial transactions have also been developed (Debrezion et al., 2007; Mohammed et al., 2013).

However, few analyses investigate the impact of new transport infrastructure on firms' location decisions. Some work has been done regarding the movement of manufacturing activities (Chatman et al., 2016; Holl, 2004a; 2004b; 2004d), which suggests that manufacturing activities tend to cluster on the edge of larger and medium cities, where transport infrastructure is available.

How do firms' location decision react to a change in spatial environment, especially related to the development or modification of transport infrastructure? How can such change affect the spatial distribution of economic activities? These questions are not without impact for policy makers and public authorities when it comes time to propose some modification of the living environment, especially from a sustainable development perspective.

The aim of the paper is to propose a global review of existing empirical evidence from the impact of development and change in transport infrastructure on firms' location decision. To do so, a two-step procedure is proposed. First, a systematic literature review is conducted to identify the pertinent literature. The systematic literature review is based on research covering three distinct databases: i) EconLit; ii) International Bibliography of the Social Science (IBSS); and iii) Web of Science. Using specific keywords, a total of 338 papers were isolated. Of this total, 52 papers have been extensively analyzed.<sup>1</sup> Second, a meta-analysis is developed based on all the results available within the identified papers, with special attention devoted to the contextual and methodological variables that may influence the impact observed.

The remainder of the paper is divided into five sections. The next section is devoted to a review of the conceptual and theoretical background on the role of transport infrastructure on location decision, i.e., the accessibility and its impact on firms, the economic benefits or activities, and the negative externalities that might be introduced. The third section describes the different stages of the systematic literature review. The fourth section is a descriptive analysis of the selected papers. The fifth section presents the meta-analysis and the inherent results, while the last section concludes the paper.

### **Conceptual and theoretical background**

The very first location theory, based on the work of von Thünen (1826), points to the importance of transportation costs in location decisions. The general theory of industrial location, proposed by Weber (1909), suggests that firms choose their optimal location by minimizing total transportation costs related to input and output, and their respective weights. The central place theory, introduced by Christaller (1933), also underlines the role of transport cost on final sale price, naturally pushing economic agents to cluster and concentrate over space.

Lösch (1954) suggests that the spatial distribution of economic activities (supply) lies on the (spatial) distribution and the frequency of the demand. As such, local economic activities,

---

<sup>1</sup> The complete list of the papers selected is provided in *Annexe du Chapitre 1* (Table A1.1).

such as bakery and barber shops, can be found in any city since the demand is frequent. Specialized service activities, such as banking and lawyers, are usually found in larger cities, as the demand for such services is less frequent. The model explains why a country does have a hierarchical distribution of the cities with a different mix of goods and services supply. It also explains why countries have many small and medium cities and fewer large metropolitan areas.

The Alonso-Muth-Mills (AMM) model (Alonso, 1964; Muth, 1969; Mills, 1972), articulated on a monocentric city, is among one of the first models to introduce location decision based on utility/profit maximization and optimization. The model assumes that most of the economic activities are located within the Central Business District (CBD) which is, by definition, the major attractive node, i.e., where workers and consumers converge throughout the day (Glaeser et al., 2008; Kulish et al., 2012). Based on economic constraints, such as revenue, expenses, and individual preferences, the economic agent decides to localize where their profit/utility is maximized.

The location decision of individual firms also shapes spatial concentration of economic activities and generates agglomeration economies, a classic concept first introduced by Marshall (1891). On the one hand, spatial concentration of similar economic activities leads to agglomeration economies, which are enhanced by proximity to other firms, e.g., sharing of skilled labor, knowledge, specialized equipment, and consumers (Chatman and Noland, 2011). On the other hand, the clustering of different economic activities within cities contributes to the diversification of the metropolitan economy. This diversification with different, but complementary, characteristics and activities, leads to the concept of urbanization economies, which also introduces production advantages (Jacobs, 1969). The presence of economies of agglomeration have been empirically confirmed (Chatman et al., 2016; Weisbrod et al., 2014; Holl, 2004d). However, the results are still heterogenous within the literature (Beaudry and Schiffauerova, 2009; Melo et al., 2009).

Spatial concentration of the economic agents can also induce negative externalities, which in turn results in the spreading out of some economic activities (Henderson and Ono, 2008). High spatial concentration may increase negative externalities such as congestion, automobile pollution, higher crime rates, and others (Credit, 2018). The increase in spatial

concentration implies a higher density within the CBD, which can have a negative impact on the quality of life and health of the residents and workers (Simeonova et al., 2021; Zhong et al., 2017).

Transport infrastructure can help to reduce potential negative externalities by facilitating the movement of goods and people. Better accessibility influences the spatial distribution of the consumers/workers and firms/producers, who choose to locate around or within those transport-generated centers (Glaeser and Kohlhase, 2004). A reduction in transportation costs is usually internalized into the real estate market by increasing the prices (or rent) of the more desirable locations (Duncan, 2011; Mohammad et al., 2013; Li and Huang, 2020). Transport infrastructure influences and shapes the urban dynamics and location decisions of economic agents, depending on their intrinsic characteristics, such as cost structure, its specific needs, and their respective clienteles (Fujita, 1989).

On the one hand, firms in the service industry choose to locate within the city center (in metropolitan areas) because it represents the center of the market, which facilitates face-to-face contact (Schuetz, 2015). For knowledge-intensive business services (KIBS), firms choose to locate where competing (and similar) firms spatially agglomerate, which potentially produces spatial spillover effect (Wong and He, 2002). The combination of resources and knowledge leads to innovation, which in turn helps expand the productivity of the industry (Grossman and Helpman, 1994; Gibbons et al., 2019). A reduction in transportation costs can also improve externalities related to agglomeration economies (Coughlin and Segev, 2000; Melo et al., 2010). The knowledge spillover is usually spatially concentrated, since ideas travel easier over small geographic distances, because of their informal nature. All those factors point in the same direction: the reinforcement of agglomeration economies, which raises productivity and benefits every firm inside those clusters (Gibbons et al., 2019; Weisbrod et al., 2014; Baptista and Swann, 1998). For such economic activities, proximity to public transport may thus be an interesting advantage compared to other types of transport infrastructures (Credit, 2018; 2019).

On the other hand, manufacturing, logistics and wholesale firms locate where land is relatively affordable (Holl, 2004a). Glaeser and Kohlhase (2004) identify a non-monotonic relation between the distance to CBD and manufacturing firms' location, meaning that the



number of manufacturing plants is highest in mid-density regions (Henderson and Ono, 2008). Manufacturing firms may also want to be close to a highway and railway, because they connect the suburbs to the CBD and facilitate the transport of their inputs and outputs (Holl, 2004a; Verhetsel et al., 2015).

Overall, transport infrastructures have (and still) played an important role in the development of the traditional city center (CBD) and on the development of other economic centers, which now shape many (polycentric) North American cities (Baum-Snow, 2007a; 2007b; Bogart and Ferry, 1999; Giuliano and Small, 1991; McDonald and McMillen, 1990).

### **Methodology**

To explore the existing literature on the impact of transportation infrastructure on firms' location decision, a systematic bibliographic review (Figueiredo et al., 2017; Munn et al., 2018; Shelby and Vaske, 2008) is undertaken. Systematic bibliographic research is a logical process that first selects all possible relevant papers based on a set of appropriate keywords. The use of Boolean operators ("AND" and "OR") is critical in such exercise.

The main keywords used for this study are: *transport*, *firm*, *location* and *infrastructure* (Table 1.1). The four keywords have been used with the "AND" option. To cover the widest range possible, the use of synonyms was considered. The truncation of the words, based on the asterisk option (\*), is used to include all possible derivations. For example, by using *transport\**, the research will include its derivations such as *transportation*, *transportation cost* and *transportation network*. The last keyword, *transit\** was also used with the "OR" option for the keyword *transport* and includes derivations such as *public transit*, *transit-oriented development*, and *urban transit*. The keyword *firm* also used the asterisk option (\*) to include words such as *firm-level* and the words *business* and *enterprise* were used as synonyms. These substitutes were used with the option "OR." The keyword *location* used the word *spatial location*, *accessibility* and *proximity* as synonyms. Finally, the words *highway*, *motorway*, *railway*, *station*, *port* and *airport* were used as substitutes, with the "OR" option, for the keyword *infrastructure*. The list of relevant papers is created with the use of keywords linked to the research.

To identify relevant papers, the bibliometric research is based on three distinct databases: i) EconLit; ii) International Bibliography of the Social Science (IBSS); and iii) Web of Science. These databases cover a wide range of papers. The systematic bibliometric research was conducted until December of 2020, and was limited to English publications, peer-reviewed, and did not consider articles classified as reviews. After having isolated all the papers, a first selection round was done by relevance of the titles. The titles must target the impact of transport infrastructure on the firm’s location decision.

**Table 1.1 - Keywords and their derivation**

Transport*		Firm*		Location		Infrastructure
Transportation		Firm-level		<b>OR</b>		<b>OR</b> Highway
Transportation network		<b>OR</b>		Spatial location		<b>OR</b> Motorway
Transportation cost	AND	Business	AND	<b>OR</b>	AND	<b>OR</b> Railway
Transport infrastructure		<b>OR</b>		Accessibility		<b>OR</b> Station
<b>OR</b>		Enterprise		<b>OR</b>		<b>OR</b> Port
Transit*				Proximity		<b>OR</b> Airport

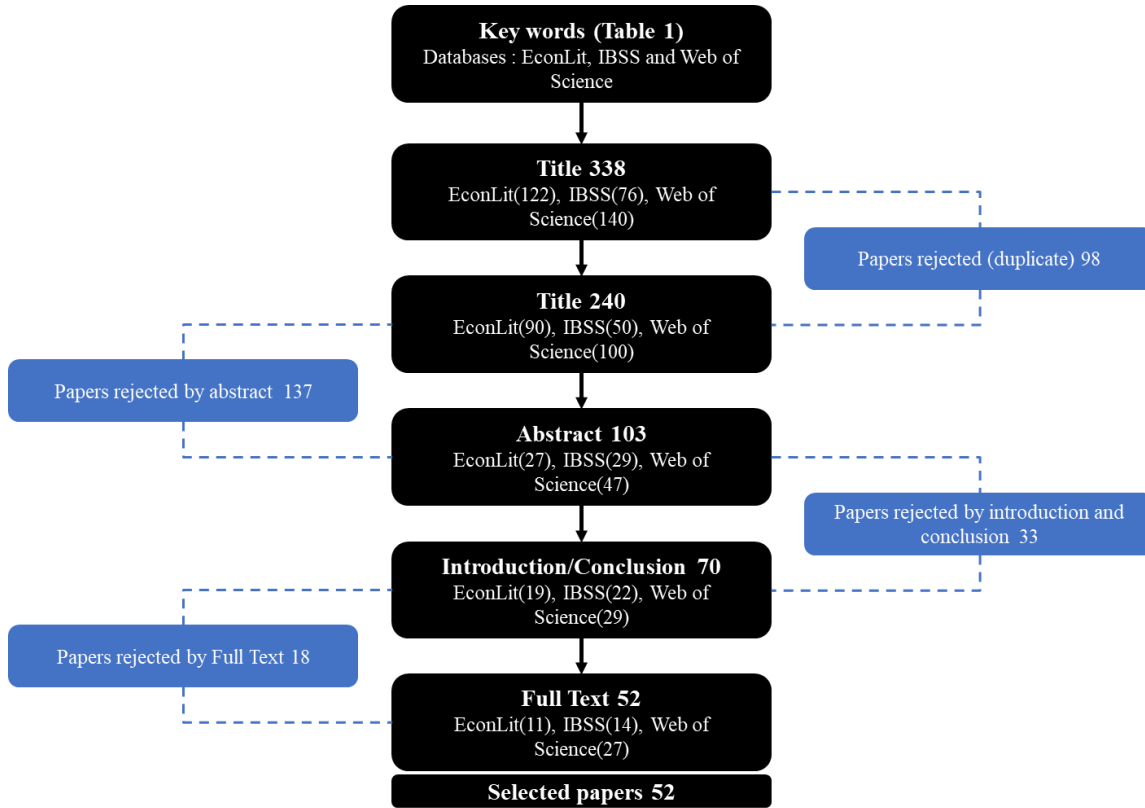
For the second step, the papers are selected based on the pertinence of their abstract and the theme of the paper. One by one the abstracts are read and analyzed. The papers with relevant abstracts are kept, while the others are rejected. The same process is repeated with the introduction and the conclusion of the papers. If the same paper has been selected more than once, in different databases, the duplicates are discarded.

Finally, the selected papers are read and the most relevant are retained for further analysis. During this last step of the selection process, the focus is on the link between transport infrastructure, in general, and the location decision of the firms around those infrastructures.

The first selection returns a total of 338 papers (Figure 1.1). Of this total, 98 were rejected because they were duplicates, and 137 papers were not considered because they did not match the subjects based on the abstract. After a first round, 33 papers were rejected based on the content of their introduction and conclusion. Finally, after a full reading of the papers, 18 were rejected because they did not match the topic. The final data set consists of 52 published papers.<sup>2</sup>

<sup>2</sup> The data type, methodology used, transport infrastructure studied, and main conclusion of each paper are summarized in *Annexe du Chapitre 1*. For each paper, only the main approach is reported, meaning that even if

**Figure 1.1 - Selection process of papers**



### Descriptive Statistics

The analysis of the papers identified in the systematic literature review suggests that the development of the impact of transport infrastructure on the location decision of the firms is an emerging topic. Half of the selected papers (26 out of 52) were written in the last 5 years (2015–2020), while 15 were written between 2000 and 2009 and only 2 published before 2000 (Figure 1.2).

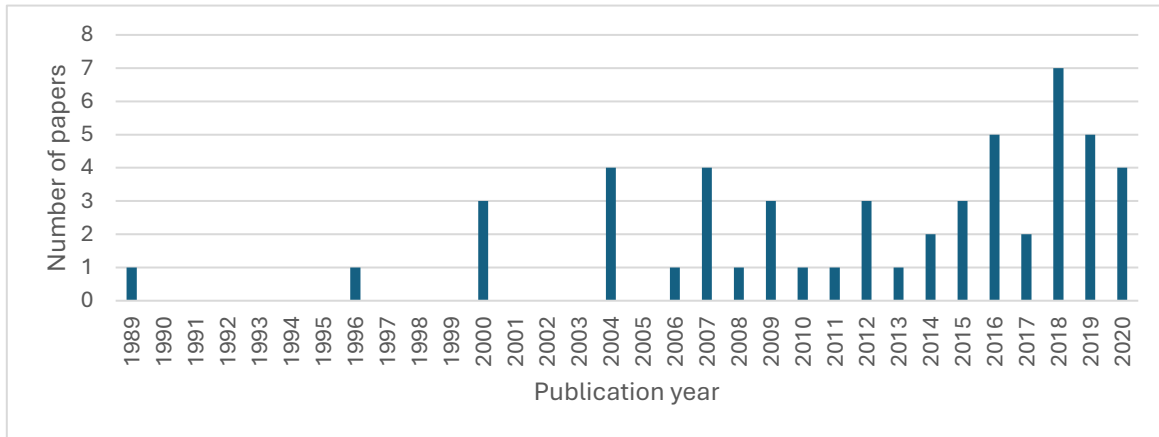
The development of the cities is still an evolving topic and factors influencing the dynamic mechanisms, such as transport infrastructure, will continue to emerge in the coming decades. Thus, even if the literature is relatively circumscribing, it nevertheless remains that understanding its impact on firms’ location decisions could help public authorities to think

---

other methodological approaches may have been introduced by the authors, only one approach was chosen and noted in Table A1.2. The methodological approach was selected based on its relative importance in the paper.

about a way to anticipate potential external effects and try to limit the negative collateral effects.

**Figure 1.2 - Concentration of papers per year**

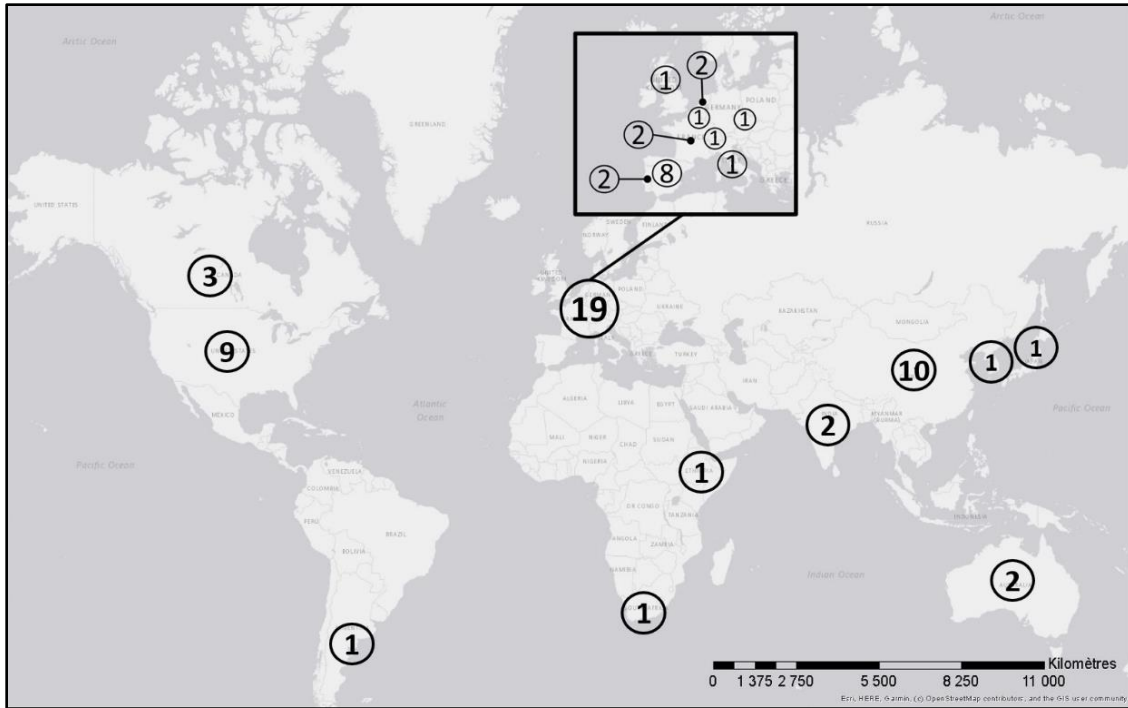


As noted in the *conceptual and theoretical background* section, the impact of transport infrastructure on the firms’ location decision includes a spatial component. Thus, the location of the different case studies is analyzed. Based on the spatial distribution of the case studies, the relative importance of Europe, China, and the United States is clear: 38 out of the 52 case studies originated from one of these three areas (Figure 1.3).

To ensure the quality of the selected papers, we chose peer-reviewed articles and turned our attention to the journals in which they were published. To evaluate the quality of the journals we use their impact factor (IF).<sup>3</sup> The most popular journals, which are *Journal of Transport Geography* (IF: 4.986) and *Transportation Research Record* (IF: 1.560), published a total of 4 papers each. Other popular journals identified are *Transportation Research: Part A: Policy and Practice* (IF: 5.594), *Papers in Regional Science* (IF: 2.858) and *Journal of Regional Science* (IF: 2.713). Three articles were selected in each of these journals.

**Figure 1.3 - Geographical distribution of the case studied**

<sup>3</sup> The 2021 impact factors were used, and all impact factors were chosen from the “Journal Impact Factor List 2021” (<https://impactfactorforjournal.com/jcr-2021/>). This list is provided by the Journal Citation Report and published by Clarivate Analytics, a Web of Science Group. This list includes the impact factors of close to 13 000 journals covering 27 research disciplines.



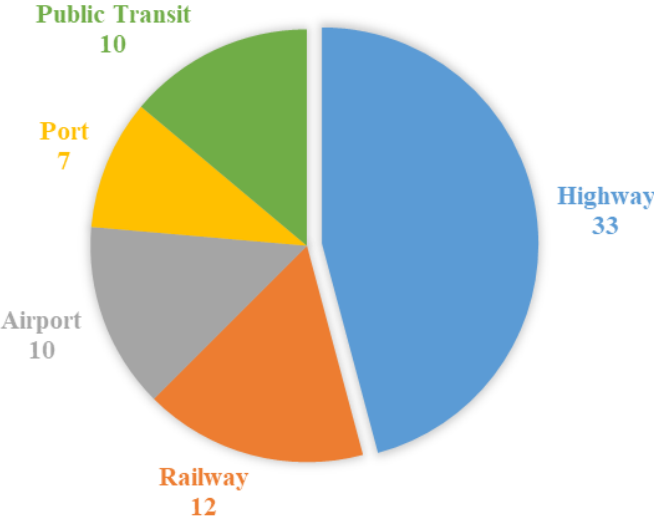
Based on the Journal Citation Report, a score over 3 is a good score, while an impact score of less than one is considered average. None of the selected papers were published in average journals since our lowest ranked journal is *Japan and the World Economy* (IF: 1.206), while the highest is *Technovation* (IF: 6.606). The weighted average impact factor of the journals in which the selected papers were published is 3.441, suggesting that, on average, the selected papers were published in journals with a good score.

The literature on the impact of transport infrastructures on firms' location decision can be analyzed regarding the type of infrastructure. For analysis purposes, the transport infrastructures have been separated into five categories: *Highway, Railway, Airport, Port* and *Public Transit* (Figure 1.4). Railways are defined by the coverage of the rail network, which must connect at least two distinct regions. Public Transit encompasses every means of public transportation, such as metro/subway, tramway, and bus, even if their infrastructures can differ. Commuter trains are also included in public transit since they connect the city core to its suburbs, serving a single region.

Highways are the most common infrastructure analyzed with 33 papers, which represent a little more than half of the reported papers. Railways, airports and public transit infrastructure are studied in 12, 10 and 10 papers, respectively, while maritime ports are less common,

being included in only 7 papers. Some papers evaluated more than one type of transport infrastructure, which explains why the number of transport infrastructures studied exceeds the number of selected papers.

**Figure 1.4 - Transport infrastructure studied in the selected papers**



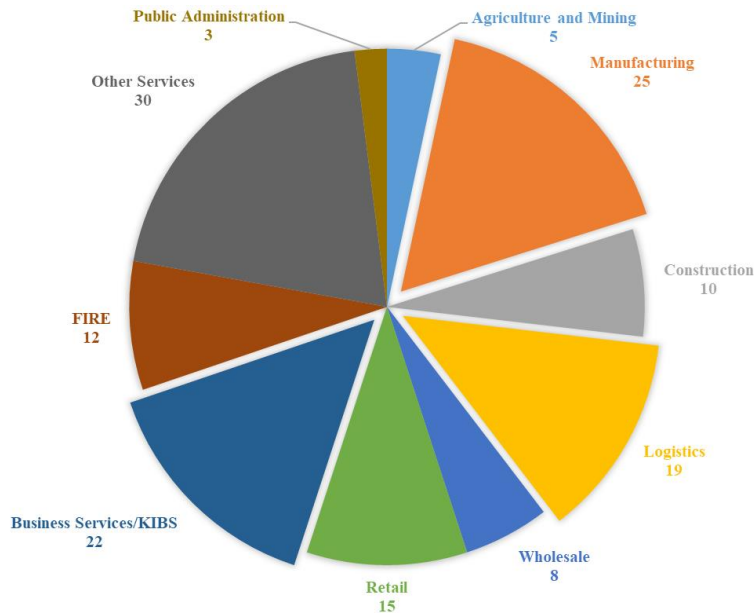
Location decisions are also investigated based on different economic activities. The literature mainly focuses on manufacturing firms. In fact, almost half of the selected papers (25 out of 52) are devoted to this sector (Figure 1.5).

The impact of transport infrastructure on logistics firms’ location decision is also common, based on the selected papers. This economic sector includes various modes of transportation, warehousing, and arrangement of freight/cargo transport. In our sample, logistics firms are evaluated in 19 out of the 52 selected papers.

The services sector is also widely studied, but it is often disaggregated into two distinct classes: i) the standard services, such as retail (15) and wholesale (8); and ii) the knowledge economy, such as business services (14), finance, insurance, and real estate (FIRE) (12) and knowledge-intensive business services (KIBS) (8). As for the transport infrastructure, most of the selected papers (30 out of 52) studied more than one economic sector.

The selected papers use a wide range of methodological approaches and types of data, which clearly underlines the heterogeneity of the applications (Table 1.2).

**Figure 1.5 - Economic sectors studied in the selected papers**



\* “Logistics” also includes *Transportation* and *Warehousing* activities. “Other services” refers to *Health and Social Services, Education, Food and Accommodation, and Arts and Entertainment*.

**Table 1.2 - Data type used based on the methodological approach**

	Cross-Sectional	Cross-Sectional Pooled	Longitudinal Data (Panel)	Time Series	Survey/Stated Preferences	Total
Logit Model	5	5	2	2	2	16
Count Model	2	6	4	0	0	12
Descriptive Analysis	0	1	1	0	6	8
Linear Model	1	1	2	0	1	5
Spatial Econometrics Model	3	0	0	0	0	3
Other Models	4	0	2	0	2	8

The most common methodological approaches used to evaluate the impact of transport infrastructure on firms’ location decision are the logit models and count data models (Poisson/negative binomial). Logit models (conditional, nested, multinomial, etc.) are used to analyze the probabilities that a firm chose a certain location (such as municipalities) as a function of the presence of a transport infrastructure located nearby. The different variations of the logit models were also popular to evaluate how the transport infrastructure influenced the probability of firms locating nearby.

Count data models, such as Poisson and negative binomial regression, are also popular in selected papers. The models are mostly used to explain the number of new firms that emerge in a certain location, which can be influenced by (new) transport infrastructure (independent variable). For those models, longitudinal and pooled cross-sectional data were mainly used.

Other models are used, such as the descriptive analysis, linear models, and spatial econometric models. Descriptive analyses are generally associated with stated preferences and other survey methods. Those approaches aim to evaluate the different factors that may influence the location decisions of the firms, mainly from a firm/employer perspective.

Linear models, such as Ordinary Least Square (OLS) regression and Generalized Method of Moments (GMM) regression, were mostly used to study firm growth in a treated/control area. Difference-in-differences (DiD) specifications are popular to identify the causal impact of transport infrastructures.

Finally, spatial econometrics models are used when the geographical dimension of the data is explicitly considered. Spatial econometrics is useful to take into account spatial spillovers and externalities (LeSage and Pace, 2009). This is the main interest of this approach as compared to classical econometric models. The spatial Durbin model was the most common spatial model (2 out of 3). The remaining models were the Spatial Error model (SEM) and Spatial Lagged Explanatory Variable model (SLX) used by Wu et al. (2019).

From a qualitative perspective, the main conclusions of the potential effects of transport infrastructure on firm's location decision appear to be highly heterogeneous (Table 1.3). The magnitude of the effect appears to be related to the type of transport infrastructure and economic sector. In the 52 papers identified, 308 results were extracted, while 11 of them were ultimately dismissed, due to their subjective nature. Those were mainly coming from surveys/stated preferences with not much information about positive or negative impact on firms' location decision.

Manufacturing activities appear to be positively influenced by the development and location of highway, railway, and airport infrastructure. For firms in this sector, along a major transport infrastructure to ship goods and receive inputs, this seems to be the main factor that affects location decisions. Logistics, transport, and warehouse firms also appear to be more



influenced by the proximity to highways, airports, and ports. This is logical, as most of the transport of merchandise is made as a dispatch of less important quantities and a complementary transport logistic to national and international transport.

The fact that the services activities, such as FIRE, KIBS, Education, Health and Social Services, Food and Accommodation and Arts and Entertainment, appear to be positively affected by the presence of highways and public transit reveals that firms seek to maximize their potential clients by maximizing collective accessibility. It is thus logical that they seek proximity to these transport infrastructures. Moreover, both infrastructures allow a relatively direct and fast access to firms for employees and clients. This is particularly true for FIRE, KIBS, Food and Accommodation as well as Arts and Entertainment. For those firms, transport costs are partly assumed by the employees and the clients, which points to research for a central place within the city. They are looking for places with transport facilities related to the traveling of individual agents instead of resources and goods. Those conclusions can also be extended to railways, which seem to have a positive impact on FIRE, KIBS and Food and Accommodation activities.

Finally, very few negative results were identified independently of the type of transport infrastructure and the economic sector. Therefore, transport infrastructures do appear to have, at most, a positive impact on firms' location decision, or at least, no significant impact.

**Table 1.3 - Summary of the selected papers' results: transport infrastructure's effects on firms' location decision based on the economic sectors and the type of infrastructure**

	Highway			Railway			Airport			Port			Public transit		
	+	n.s.	-	+	n.s.	-	+	n.s.	-	+	n.s.	-	+	n.s.	-
Agriculture/Mining		6			3			1						1	
Construction	5	5	1	2	1		1						2	1	
Manufacturing	14	11		6	2		4	1		1			4	3	1
Logistics, Transportation, and Warehousing	13	8		4	3		4	2		6			3		
Wholesale	5	5		2	1		2						1	1	
Retail	7	6		2	1		2						4	3	2
KIBS/Business Services*	8	7		5	1		5	1					7	5	1
Finance, Insurance, Real Estate	9	3		3			1	1					4	1	1
Health and Social Services	4	4	1	1	1		2						3		1
Education	4	3		1	1		1	1					2		
Food and Accommodation	7	5		3			2						4	1	1
Arts and Entertainment	2	1	1				1	1					3	1	
Public Administration	3	1			1		1	1					1		
Total	81	65	3	29	15	0	26	9	0	7	0	0	38	17	7

Legend: All results presented were significant at the 5% level or lower.

\* Professional, scientific, technical, administrative and support service activities.

n.s.: non-significant results.

## Meta-Analysis

The meta-regression analysis presented in this section aims to evaluate the impact of transport infrastructure on firms' location, while studying the role of different variables (contextual and methodological) on the observed results.

To do so, the results are dichotomized in order to conduct a “sign and significant” approach (Duvivier, 2019; van Huizen and Plantenga, 2018; de Groot et al., 2016; Card et al., 2010; Horváthová, 2010). It allows the study of a large range of results compared to more common approaches for meta-analysis, such as effect-size.

The 297 results (at a 5% significant level) presented in the 52 selected papers are used as a categorical dependent variable. Using an ordered-logit model (Equation 1.1), the results are divided into three categories: (i) positive and significant results ( $Y_i = 1$ ); (ii) non-significant results ( $Y_i = 0$ ); and (iii) negative and significant results ( $Y_i = -1$ ).

$$P(Y_i = j|X_i) = F(X_i\beta) + \varepsilon_i \quad (1.1)$$

Where  $j = \{1, 0, -1\}$  and  $F(X_i\beta)$  is the cumulative logistic distribution (Williams, 2016; Stata, 2022; Hailpern and Visintainer, 2003). The explanatory variables ( $X$ ) can be separated into two categories: contextual and methodological (Duvivier, 2019). On the one hand, the contextual variables refer to the different characteristics relative to the case study, such as economic sector, type of transport infrastructure and location in which the case study took place. As presented before, those variables are expected to influence the observed results. On the other hand, the methodological variables include the econometric approaches and data type, which are also likely to influence the reported results. A likelihood-ratio test was conducted and confirmed that the selected model and variable are suited for the meta-analysis.

The main issue related to the “sign and significance” approach is the lack of mean to address and correct for the publication bias. However, no publication bias/selection was found in the 297 observed results.<sup>4</sup> A second issue that can arise from the data is the small proportion of

---

<sup>4</sup> A funnel plot is presented in *Annexe du Chapitre 1* (Figure A1.3), with the logarithmic function of the odd ratios on the x-axis and the inverse standard-error on the y-axis (Hoffman, 2019; Sterne and Harbord, 2004). The distribution of the (log) odd ratios is allocated symmetrically around the vertical line and within the 95%

negative results (Table 1.3), which may impact the results of the meta-analysis. To address this potential bias, robustness tests are conducted in which we address the small number of negative observations. For the first test, the dependent variable is divided into two categories; (i) positive and significant results ( $Y_i = 1$ ); and (ii) non-significant results and negative and significant results ( $Y_i < 1$ ). For the second robustness test, the negative results are simply dismissed ( $Y_i = \{0, 1\}$ ). The results of the robustness tests are similar to those of the first regression (Table 1.4) and are presented in *Annexe du Chapitre 1* (Table A1.4.1, and Table A1.4.2).

The results of the meta-analysis (Table 1.4) suggest that the economic sector is mostly non-significant, relatively to a manufacturing firm, to explain the sign and significance. However, a statistically significant impact is noted in the Finance, Insurance and Real Estate (FIRE) sector. The results suggest that the probability of identifying a positive effect of transport infrastructure on firms' location decision in FIRE activities is 2.5 times more important, compared to manufacturing firms. The results also indicate that the impact of the Agriculture and Mining sector is significant and that the probability of observing positive effects in this sector is approximately 80% less than firms operating in the manufacturing sector.

Regarding transport infrastructure, the railway system is the only one to show positive results and significant odd ratio as compared to the highway. For other transport infrastructures (Airport, Port and Public Transit), the effects appear to be not statistically different from those found around highways. Thus, railways seem to influence more firms' location decisions than highways, airports, ports, and public transit. The probability of observing positive effects is 1.8 times higher for railways, compared to highways.

As presented in the *conceptual and theoretical background* section, transport infrastructure has an impact on firms' location decisions, depending on their intrinsic characteristics and economic sector, which influence their specific needs and clientele. On the one hand, manufacturing or logistics firms use transport infrastructure mainly to move their physical

---

confidence interval. Therefore, the full sample shows no publication bias (Egger et al., 1997). This lack of bias may be explained by the inclusion of all results presented in the papers, and not only the one discussed by the authors.

input or output. On the other hand, services firms must be easily accessible by consumer and workers to operate.

**Table 1.4 - Meta-analysis results**

Variable		Odd ratio	Std. Error	sign.
Sector	Manufacturing	Reference	--	
	Agriculture and Mining	0.2060	0.1529	**
	Construction	1.4172	0.9040	
	Logistics	2.1994	1.2584	
	Wholesale	1.2283	0.8271	
	Retail	0.9892	0.5777	
	Business Services	0.8648	0.5059	
	FIRE	3.4702	2.2244	*
	KIBS	4.4027	4.1227	
	Education/Health/Food and Accommodation/Arts and Entertainment	1.5910	0.7829	
Public Administration	1.1144	1.1156		
Transport	Highway	Reference	--	
	Railway	2.7649	1.2840	**
	Airport	0.9856	0.5612	
	Port	8368311	1.48e10	
	Public Transit	1.0989	0.4825	
Location	Other	Reference	--	
	United States	0.0522	0.0573	***
	Europe	0.1666	0.1633	*
	Asia	0.1746	0.1580	*
Methodology	Other Models	Reference	--	
	Count Model	4.8424	2.6687	***
	Logit Model	1.1960	0.7647	
	Descriptive Analysis	0.1769	0.1518	**
	Linear Model	2.4115	3.2804	
	Spatial Econometric Model	0.3331	0.4180	
Data	Cross-Sectional	Reference	--	
	Pooled Cross-Sectional	0.0672	0.0789	**
	Longitudinal/Panel	0.0429	0.0498	***
	Stated Preferences	0.2117	0.2895	
	Time Series	6562596	5.50e9	

The option [or] in Stata was used in order to obtain the odd ratio as coefficient.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

To account for such possibility, a dummy variable was introduced with a cross-product multiplication accounting for the nature (freight or passengers) and the type of transportation (see *Annexe du Chapitre 1*). The results suggest no significant differences between all the transport infrastructures, in relation to the freight transportation (Table A1.5.1) and passengers transportation (Table A1.5.2). However, those results should be interpreted with

caution, due to the limited number of observations and the weak number of observations different from 0 appearing in the interaction terms.

Furthermore, studies conducted in countries located outside of Europe, Asia and the United States seem to find more positive impact of transport infrastructure on firms' location decision. Results suggest that operating in countries such as Canada, Australia or other countries located in Africa or South America induces higher probabilities ranging from 83 to 95% to identify a positive effect, as compared to Europe, Asia, or the United States.

Regarding methodological choices, the results appear to be more positive and significant when the analysis is based on the use of count models, while they are less positive and significant when using "Descriptive Analysis," as compared to the "Other models." On the one hand, using a "Count model" implies a probability of observing positive effect 3.8 times higher, compared to the "Other models". On the other hand, a decrease of 82% in the probability of identifying positive effects of transport infrastructure on firms' location decision is noted when a "Descriptive Analysis" is conducted, in comparison with the "Other models."

Finally, the data type also appears to influence the conclusions. "Pooled Cross-Sectional" and "Longitudinal" data seem to produce more conservative results as compared to the reference category ("Cross-Sectional"). The results indicate that the impact of transport infrastructure on the location of firms is relatively lower when data takes into account the spatial and temporal dimension simultaneously. The probability to observe positive and significant effects of transport infrastructure on firms' location decisions are, respectively, 93% and 96% lower when using Pooled Cross-Sectional data and Longitudinal/Panel data as compared to Cross-Sectional data. For all other statistical modeling approaches, i.e., "Stated Preferences" and "Time Series," the results appear to be statistically equal, as no coefficients are statistically significant.

### **Concluding remarks**

Based on the papers published in this emerging literature (52 papers), a complete analysis of the publications was conducted. The compilation underlines a consensus on the impact of such infrastructure on the location decision of firms: in the worst-case scenario, the

infrastructure appears to play a non-significant role on location decision, as only 10 results report negative and significant effect out of the 297 observed results.

Moreover, the use made of transport infrastructure may be different (highways vs. public transportation) based on who assumes transportation costs (firms or workers), and thus, on the main activity of individual firms. Therefore, the observed impact of transport infrastructures on firms' location decision may depend on the infrastructure and the firm's economic sector.

On the one hand, the pattern is different for manufacturing firms since most of their input and output are materials (goods). Moreover, many firms are looking for large spaces, for their production process, that are well connected to the markets and inputs. Consequently, the expansion of highways is a major factor that influences those firms' location. On the other hand, services' economic activities use more human capital in the production process. As such, firms do not directly pay transport costs for their main input (assumed by the worker), nor even for part of the services they provide (assumed by the clients).

From a qualitative perspective, the impact of highways has been largely studied in the literature, and those impacts were mostly evaluated for manufacturing and logistics firms. For those sectors, the impacts of highways on their location decision tend to be positive, nevertheless non-significant results were also found.

The impact of highways was also studied for other sectors, such as KIBS, business services and FIRE. The descriptive analysis shows that the impacts of highways (and public transit) were positively dominant for FIRE activities, while being more ambiguous, between positive and non-significant effect, for KIBS and business services. Furthermore, a significant and positive impact of railways on FIRE, KIBS and business services activities emerges from the literature.

Regarding FIRE activities, similar conclusions were drawn from the meta-analysis. In fact, for the majority of the economic sectors the influence of transport infrastructure on firms' location decision is not statistically different from the impact observed on manufacturing firms. However, the probability of observing a positive effect of transport infrastructure on location is higher for FIRE activities than for manufacturing firms.

Another interesting conclusion of the meta-analysis is the impact of the selected data on the results. Pooled cross-sectional data and longitudinal (panel) data appear to obtain smaller results, compared to cross-sectional data. The main difference lies in the temporal dimension, which is ignored or relatively small for the cross-sectional data. The meta-analysis results imply that using cross-sectional data can potentially lead to a (positive) bias in the observed results.

The originality of the paper is to propose a comprehensive synthesis and meta-analysis of this emerging research field. Since the literature has recently expanded, this paper aims to establish some common bases to guide future research. To date, the recent work of Credit (2018; 2019) appears to be a good theoretical and empirical base to develop for further investigation into the direct impact of the development and implementation of transport infrastructure on firm's location decision.

The main limitations of this paper rely on the selected approach for the meta-analysis as well as the sample size available for the estimation. Using a "sign and significance" approach does allow a heterogeneous set of methodological approaches, but it can only indicate whether the expected impact will be positive, non-significant or negative. Therefore, the quantitative dimension of the observed results cannot be measured or estimated by the selected approach. Moreover, the small number of results, especially for non-significant and negative effect, sets a clear limit on the interpretation that can be made from such an exercise.

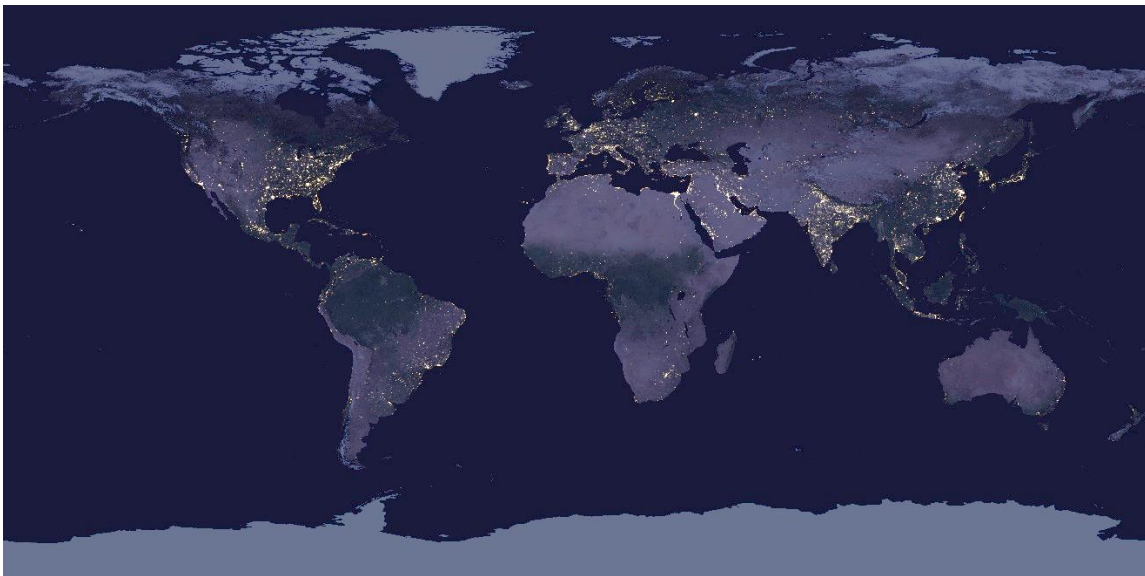
Finally, the explicit consideration of spatial spillover effect is rarely taken into account in the modeling exercise of the selected papers. However, as spatial concentration appears to be an important key factor of location decisions, it might also generate a spillovers (wider economic effects) that could be captured by spatial variables. The consideration of spatial econometric models might be a good way to capture such latent components. A literature review on the subject could be another interesting path to investigate for future research on the existing literature.



## CHAPITRE 2 : Cadre théorique

En économie urbaine et régionale, l'espace joue un rôle crucial pour expliquer les comportements économiques. L'espace n'est pas un univers neutre et les caractéristiques spatiales ne sont pas distribuées de manière uniforme (Figure 2.1). La répartition des ressources et des populations influence les décisions de localisation des entreprises, et ces décisions ont des effets de rétroactions sur les localisations des ménages et d'autres entreprises.

**Figure 2.1 - Distribution non-uniforme des activités économiques**



Source : NASA (2016)

Bien que l'étude de l'urbain et du régional repose en large partie sur l'approche néo-classique, les fondements classiques sont néanmoins révisés par l'introduction de variables additionnelles. De manière simple, la prise en compte de l'aspect spatial dans les décisions de localisation nécessite d'introduire de nouvelles variables dans les équations classiques telles que la fonction de profit des entreprises.

### **Maximisation des profits**

Les théories microéconomiques qui dictent les comportements des entreprises reposent sur l'optimisation de la fonction de profit. Selon l'approche néo-classique, une entreprise cherche à opter pour une localisation qui maximisera son profit (Marshall, 1891; Varian,

2015), défini par la somme des recettes moins la somme des dépenses liées au processus de transformation et de production (équation 2.1). D'un côté, les recettes d'un secteur productif donné,  $s$ , sont obtenues en multipliant la quantité produite d'un extrant ( $q_s$ ) par son prix de vente ( $p_s$ ). D'un autre côté, les coûts de production sont définis par la somme des quantités d'intrants mobilisés ( $x_{is}$ ) multipliées par les prix respectifs des intrants ( $w_{is}$ ). Habituellement, plusieurs intrants sont nécessaires à la production,  $i = 1, \dots, M$ . Les intrants peuvent être sous forme de capital physique (ressources, terrain) ou humain (travailleurs).

$$\pi_s = q_s p_s - \sum_{i=1}^M x_{is} w_{is} \quad (2.1)$$

Il est possible de voir cette équation avec un certain nombre de variables endogènes, c'est-à-dire celles qui sont déterminées par l'entreprise elle-même, mais aussi par des variables exogènes, c'est-à-dire hors du contrôle de l'entreprise. Dans une perspective individuelle et pour des entreprises de tailles modestes, il est concevable d'imaginer les prix comme des variables exogènes à l'entreprise. On dit alors que les entreprises sont des *price takers*. Ainsi, la maximisation des profits correspond, par dualité, à une minimisation des coûts de production.

Pour une quantité donnée d'extrants, une entreprise doit optimiser les combinaisons d'intrants, selon les possibilités de substitution, afin de maximiser ses profits. Par exemple, l'automatisation de certains emplois entraîne la substitution du capital physique au capital humain. La quantité de travailleurs, soit ses besoins en capital humain, dépend de l'extrant produit, mais aussi des substitutions techniques liées à la fonction de production et aux prix relatifs.

Les pressions urbaines sont susceptibles d'influencer les prix du terrain (immobilier), des intrants (capital physique et humain), mais aussi des extrants. Généralement exogènes aux entreprises, ces prix varient en fonction du marché dans lequel elles opèrent et évoluent. Par exemple, la taille d'une agglomération urbaine tend à influencer positivement les prix du foncier et les salaires.<sup>5</sup>

---

<sup>5</sup> Certaines exceptions peuvent entraîner des salaires plus élevés, telles que les primes d'éloignements, pour de petites agglomérations isolées géographiquement.

Les coûts fonciers sont également corrélés négativement avec la distance aux marchés. Cette relation provient de l'équilibre sur le marché immobilier, entre la demande et l'offre de biens immobiliers ou de terrains. L'attractivité d'un centre se reflète par une demande élevée du sol (Alonso, 1964; Muth, 1969; Mills, 1972). L'effet prix peut avoir un impact négatif sur l'attractivité de certaines activités, bien que la présence d'infrastructures de transport puisse influencer la structure de coûts.

### **Transport et fonction de profit**

Les infrastructures de transport jouent un rôle important dans la distribution spatiale des activités commerciales. En effet, les coûts de transport sont influencés par la quantité et la qualité des infrastructures de transport disponibles. Ces derniers se reflètent dans l'accessibilité et la connectivité du réseau de transport. Pour un trajet donné, une meilleure accessibilité diminue les coûts de transport de l'entreprise, soit par une diminution du temps de déplacements ou par un itinéraire alternatif plus direct. Ainsi, pour une même distance, une localisation bien desservie en infrastructures implique généralement des coûts de transport plus faibles (Glaeser et al., 2008).

Les infrastructures de transport engendrent également des effets collatéraux : ceux générés par les effets externes. Ces externalités peuvent être positives, mais aussi négatives, et sont généralement internalisées dans les valeurs immobilières. Par exemple, la mise en place d'une autoroute améliore la connectivité et l'accessibilité d'une localisation donnée. Toutefois, une nouvelle autoroute engendre aussi une augmentation de la congestion routière, du bruit et des émissions polluantes. L'effet net résulte de la différence entre les externalités positives et négatives observées.

En ce sens, et dans un contexte spatial, deux composantes importantes s'ajoutent à la fonction de coût pour une entreprise du secteur  $s$  : i) les coûts de transport des extrants ( $c_{qs}d_{qs}$ ); et ii) les coûts de transports des intrants ( $c_{is}d_{is}$ ), où  $c_{qs}$  et  $c_{is}$  représentent, respectivement, les coûts unitaires de transport de l'extrant et des intrants, alors que  $d_{qs}$  et  $d_{is}$  représentent, respectivement, la distance nécessaire pour livrer l'extrant et les intrants (équation 2.2).

$$\pi_{sl} = q_s p_s - \left[ q_s c_{qsl} d_{qsl} + \sum_{i=1}^M x_{is} w_{is} + \sum_{i=1}^M x_{is} c_{isl} d_{isl} \right] \quad (2.2)$$

L'équation de coût suggère un certain équilibre quant à la localisation en fonction des coûts de transport de l'extrait ( $c_{qs}$ ) et des intrants ( $c_{is}$ ), relativement à la localisation  $l$  de l'entreprise. Ainsi, un processus de transformation densifiant, qui augmente la masse du produit final, favorise une implantation plus orientée vers le marché ( $d_{qs} < d_{is}$ ), alors qu'un processus de transformation réducteur, qui réduit la masse de l'extrait, favorise une implantation orientée vers les ressources ( $d_{qs} > d_{is}$ ). Cette dichotomie conduit, selon la théorie de la localisation industrielle de Weber (1909), à une caractérisation des localisations qui sont soit axées vers le marché, soit axées vers la (les) ressource(s).

Évidemment, une localisation jugée optimale pour un secteur donné ne l'est pas nécessairement pour un autre. Le secteur d'activité influence les besoins, les préférences et les comportements d'une entreprise (Fujita et al. 1999). Ces caractéristiques spécifiques par secteur se reflètent dans leurs besoins d'espace (quantité de terrain), mais également dans leurs préférences pour certains types d'infrastructures de transport.

Par exemple, une entreprise manufacturière nécessitant beaucoup de capital physique a intérêt à faciliter le déplacement de ses intrants/extrants, afin de diminuer les coûts associés à leur transport. La proximité à de grandes infrastructures de transport, telles que des autoroutes, des aéroports et des gares ferroviaires, s'avère être un critère de localisation important pour ces entreprises. En se localisant en marge du marché, elles bénéficient généralement d'une diminution des coûts de terrain, pour une superficie donnée, tout en conservant une distance relativement faible au centre urbain. Ces besoins spécifiques se reflètent dans les décisions de localisation des entreprises de manufacture qui choisissent généralement des localisations périphériques (Holl, 2004a; 2004b).

Les entreprises de services cherchent des localisations facilement accessibles pour leurs consommateurs, qui doivent se déplacer afin de bénéficier du service offert. Elles nécessitent une plus grande quantité de capital humain et se doivent de faciliter les déplacements de leurs travailleurs (Credit, 2018; 2019). En contrepartie, les intrants de ce type d'entreprise sont essentiellement des informations : qui représentent de très faibles coûts, mais qui nécessitent

néanmoins une certaine proximité physique des concurrents. En ce sens, les infrastructures de transport en commun, les aménités urbaines et les localisations centrales sont préférées par les entreprises de ce secteur.

Finalement, les entreprises extractrices de ressources ont surtout avantage à se localiser près des gisements. La transformation de la matière première nécessite souvent une perte de poids qui favorise une localisation près de la ressource afin de diminuer les coûts de transport.

### **Économies d'agglomération**

La présence d'économies d'agglomération, engendrée par la concentration spatiale d'activités économiques, peut aussi influencer la fonction de profit de l'entreprise. La proximité spatiale facilite le partage des connaissances et stimule l'innovation. Les innovations ciblent généralement une optimisation du processus de production, ainsi produire une quantité donnée d'extrants nécessite maintenant une quantité d'intrants inférieure. Par dualité, une production supérieure d'extrait est attendue pour une quantité d'intrants donnée (Chatman et Noland, 2011).

Le rôle des agglomérations (*clusters*) dans le développement économique découle des relations de compétition et de collaboration entre les entreprises à l'intérieur de ces regroupements (Porter, 1998; 1990). Ce sont majoritairement les interactions interpersonnelles et le contact humain qui sont à la base des avantages observés pour les entreprises localisées dans ces agglomérations. L'attrait de ces regroupements peut aussi amplifier le bassin de main-d'œuvre accessible et ultimement diminuer les coûts de recrutement (coûts du capital humain) assumés par les entreprises. Par ailleurs, l'agglomération spatiale d'entreprises peut être définie par une concentration d'activités économiques similaires ou diversifiées.

Un marché spécialisé génère des économies de spécialisation (Marshall, 1891). Ce type d'économies est spécifique à une industrie et repose sur l'agglomération d'activités économiques similaires. La spécialisation d'une zone géographique a aussi pour effet de faciliter l'accès à un plus grand bassin de main-d'œuvre qualifiée. Selon le modèle de Marshall-Arrow-Romer (Marshall, 1891; Arrow, 1962; Romer, 1986), les connaissances sont

spécifiques à certaines industries et le partage des connaissances bénéficie uniquement aux entreprises de cette industrie.

Inversement, Jacobs (1969) soutient que les connaissances acquises et les technologies développées ne sont pas exclusives à une seule industrie. En effet, les économies d'agglomération peuvent émerger d'une diversification d'activités, soit une des caractéristiques premières de la plupart des grandes villes (voir notamment la théorie des places centrales - Christaller, 1933). Ainsi, des entreprises de secteurs distincts peuvent s'agglomérer afin de bénéficier des avantages de la proximité. Les économies d'agglomération dans un marché diversifié, aussi appelées économies d'urbanisation, favorisent les innovations et découlent du partage des connaissances entre les industries.

Enfin, l'amalgame des activités économiques dans une zone spécifique peut également influencer la distribution des aménités publiques, telles que des infrastructures de transport, qui sont implantées de façon endogène, en fonction d'objectifs prédéfinis. La localisation d'une entreprise peut lui offrir un avantage compétitif (Porter, 1998). Les entreprises ont donc intérêt à se localiser dans un environnement dense en activités commerciales, afin de faciliter le partage des connaissances (*knowledge spillovers*), de profiter d'un plus grand bassin de main-d'œuvre et de consommateurs, et de bénéficier des aménités publiques disponibles.

En somme, la quantité d'extrants produite et d'intrants nécessaire, les coûts des intrants, ainsi que les coûts de transport sont donc des fonctions variant selon la taille de l'agglomération dans laquelle une entreprise opère ( $A_{sl}$ ). En ce sens, l'impact des économies d'agglomération dépend du secteur,  $s$ , et de la localisation,  $l$ , de l'entreprise et affecte à la fois les variables liées aux recettes et aux dépenses des entreprises (équation 2.3).

$$\pi_{sl} = q_s(A_{sl})p_s - \left[ q_s(A_{sl})c_{qsl}(A_{sl})d_{qsl} + \sum_{i=1}^M x_{is}(A_{sl}) (w_{is}(A_{sl}) + c_{isl}(A_{sl})d_{isl}) \right] \quad (2.3)$$

L'intégration de chacun de ces concepts dans l'équation de profit procure donc un cadre d'analyse plus complexe que celui de la théorie économique a-spatiale. Évidemment, ces fonctions de profits peuvent évoluer au fil du temps. Une modification importante des combinaisons d'intrants liées à la production d'un bien, par un effet de substitution, peut

modifier la localisation optimale. De la même manière, l'implantation de nouvelles infrastructures de transport ou l'augmentation de la congestion routière, peuvent être des éléments importants qui influencent la structure des coûts de transport et, par conséquent, la localisation optimale. La mise en place de certaines politiques publiques, telle que la tarification de l'utilisation des infrastructures, risque aussi d'avoir un impact sur la détermination de la localisation d'une entreprise.

Bien qu'intuitive théoriquement, la fonction de profit est généralement latente, donc non-observée empiriquement. Elle peut cependant être révélée par les décisions observables et effectuées par les entreprises (Dubé et al., 2016; Cissé et al., 2020; Aïkous et al., 2023). En effet, un déménagement indique généralement qu'une nouvelle localisation maximise davantage la fonction de profit de l'entité commerciale, par rapport à sa localisation actuelle. Autrement, l'entreprise n'aurait aucun avantage à changer de localisation.

### **Conclusion partielle**

En résumé, l'environnement dans lequel l'entreprise opère a un impact indéniable sur sa capacité de production, et donc sur sa fonction de profit. Les coûts assumés et les changements dans les caractéristiques spatiales ou intrinsèques peuvent entraîner des mouvements d'entreprises dans l'espace, au fil du temps. Dans cette optique, les composantes spatiales environnantes doivent être considérées. Ces diverses caractéristiques peuvent également s'avérer importantes afin d'expliquer la performance globale des entreprises.

Ainsi, ce projet de thèse a pour objectif d'évaluer plus en détail la relation statistique, et potentiellement causale, observée entre la performance commerciale, dans une perspective de survie, et les infrastructures de transport en commun.

Différentes approches méthodologiques peuvent être mobilisées afin d'évaluer empiriquement les impacts des infrastructures de transport sur les entreprises. Parmi ces approches, les analyses causales cherchent à identifier les relations de cause à effet et permettent de dépasser le simple cadre corrélationnel offert par la plupart des outils statistiques connus. L'identification causale repose sur des hypothèses à respecter et des stratégies particulières, qui sont discutées plus en profondeur dans le prochain chapitre. Plus précisément, le troisième chapitre de cette thèse introduit et présente les approches

corrélationnelles et causales, en plus d'insister sur les différences entre ces types d'analyse et leur interprétation.



## CHAPITRE 3 : Cadre méthodologique

L'analyse des politiques publiques et de leur impact sur des variables d'intérêt pose un défi empirique important. En effet, les approches économétriques mobilisées dans les études empiriques n'identifient pas nécessairement une relation de cause à effet entre les variables.

Les modèles de régression permettent de contrôler statistiquement les effets de variables confondantes, soit les variables associées à la fois à la variable explicative et à la variable d'intérêt. Malgré tout, même si ce type de modèle est basé sur un cadre conceptuel spécifique, identifiant les variables explicatives et d'intérêt, rien n'indique pour autant que la relation estimée permet une interprétation causale. La relation observée a essentiellement une interprétation corrélacionnelle, qui peut s'avérer utile pour identifier la relation statistique entre variables. Toutefois, elle n'est pas suffisante pour identifier si, et dans quelle ampleur, une certaine variable influence la variable d'intérêt.

L'analyse causale permet d'identifier si une variable donnée engendre (cause) une réaction, et donc si elle a un impact sur la variable d'intérêt. Ce chapitre se penche sur la mise en place de stratégies appropriées afin, ultimement, d'isoler l'effet causal d'une infrastructure de transport sur la performance des entreprises.

### Relations statistiques

Les modèles de régression linéaire sont à la base des modèles économétriques (équation 3.1). Ils expriment généralement une variable d'intérêt, notée  $y_{it}$ , en fonction d'un ensemble de variables explicatives,  $x_{kit}$ , que l'on suppose linéairement liée dans les paramètres à la variable d'intérêt. Les paramètres du modèle,  $\beta_k$ , permettent de vérifier s'il existe une relation significative entre les variables explicatives, jugées exogènes, et la variable d'intérêt.

$$y_{it} = \alpha + \sum_{k=1}^K x_{kit}\beta_k + \varepsilon_{it} \quad (3.1)$$

Dans ce type d'analyse, le terme d'erreur,  $\varepsilon_{it}$ , est supposé strictement aléatoire, de moyenne nulle, indépendant et identiquement distribué (iid). Si l'une de ces hypothèses n'est pas

respectée, un biais peut être présent dans les résultats obtenus ou dans l'application des tests statistiques standards.<sup>6</sup>

Le terme d'erreur  $\varepsilon_{it}$  contient les variables qui ne sont pas incluses dans le modèle, potentiellement des caractéristiques non-observées. Ces variables sont susceptibles de violer l'hypothèse d'indépendance, mais aussi, et surtout, d'entraîner un problème d'endogénéité.

#### *Problème d'endogénéité des variables explicatives*

L'analyse de causalité repose sur différentes hypothèses, mais la plus importante est l'exogénéité des variables explicatives. La variable explicative ne doit pas dépendre de facteurs non-observables et ne doit pas être corrélée avec le terme d'erreur (Kennedy, 2003). La présence d'endogénéité empêche l'interprétation causale des impacts de  $x$  sur  $y$ , puisque le coefficient estimé est biaisé. Par ailleurs, plus la corrélation entre la variable explicative et le terme d'erreur est grande, plus le biais sera important.

Le problème d'endogénéité, causé par l'omission d'une variable explicative pertinente, peut émerger si certaines variables explicatives sont corrélées entre elles et que l'une de ces variables n'est pas incluse dans le modèle.

D'une part, il est possible que les variables explicatives soient corrélées entre elles. Cependant, cette corrélation ne doit pas être parfaite ou très forte. Autrement, l'effet isolé d'une variable sur la variable dépendante reste indéterminé, puisque l'effet observé est intrinsèquement lié à celui des autres variables. Ce problème de colinéarité (ou multicolinéarité) affecte également la variance de l'échantillon.<sup>7</sup> En effet, plus la colinéarité entre les variables est importante, plus les intervalles de confiance estimés seront larges, ce qui affecte ultimement la significativité des coefficients (Alin, 2010).

---

<sup>6</sup> Si les tests de statistiques standards ne peuvent pas être correctement estimés (par exemple l'hypothèse iid n'est pas respectée), les tests de randomisation sont des alternatives viables. Basée sur l'inférence randomisée, cette approche non-paramétrique repose sur la permutation aléatoire de la variable d'impact (traitement). Ainsi, la distribution de la statistique de test est produite en fonction de la redistribution aléatoire du traitement. Pour plus d'informations sur les tests de randomisation, le lecteur peut consulter : Dubé et al., 2023b.

<sup>7</sup> Dans cette optique, un indicateur VIF (*variance inflated factor*) peut être estimé afin d'identifier la présence de multicolinéarité dans le modèle. Pour plus d'informations, le lecteur peut consulter : Chatterjee et Hadi, 2012.

D'autre part, l'omission d'une variable pertinente génère un problème d'endogénéité lorsque la variable est corrélée avec le terme d'erreur. Par exemple, l'équation 3.2 suppose que la variable d'intérêt ( $y_{it}$ ) dépend de deux variables explicatives, soit  $x_{it}$  et  $z_{it}$  (Antonakis et al., 2010).

$$y_{it} = \alpha + \beta_1 x_{it} + \beta_2 z_{it} + \varepsilon_{it} \quad (3.2)$$

Si les variables  $x_{it}$  et  $z_{it}$  sont corrélés entre elles, on peut décrire cette relation à partir d'une expression mathématique (équation 3.3). Pour le bien de la démonstration, cette relation est supposée linéaire. Évidemment, elle peut prendre différentes formes selon la nature de la relation entre ces facteurs.

$$z_{it} = \gamma_1 x_{it} + u_{it} \quad (3.3)$$

Si la variable  $z_{it}$  n'est pas observable, elle est alors incluse dans le terme d'erreur. Cette omission mène inévitablement à un biais dans l'estimation du coefficient de  $x_{it}$  (voir équation 3.4 : obtenue en substituant l'équation 3.3 dans l'équation originale 3.2).

$$y_{it} = \delta_0 + \beta_1 x_{it} + [\beta_2(\gamma_1 x_{it} + u_{it}) + \varepsilon_{it}] \quad (3.4)$$

La réécriture du modèle permet de constater que le coefficient estimé de la variable  $x_{it}$  est biaisé par le terme  $\beta_2 \gamma_1$  (équation 3.5).

$$y_{it} = \delta_0 + (\beta_1 + \beta_2 \gamma_1) x_{it} + (\beta_2 u_{it} + \varepsilon_{it}) \quad (3.5)$$

Ce biais découle de l'omission de variables explicatives pertinentes et de l'endogénéité des variables explicatives. Empiriquement, identifier tous les facteurs qui pourraient avoir un impact sur l'une ou l'autre des variables explicatives s'avère quasi-impossible, puisque certains de ces facteurs ne sont pas observables.

Une solution à ce problème consiste à inclure des effets fixes (spatial et temporel) afin de contrôler pour ces variables omises et non-observables. Les effets fixes temporels permettent de contrôler pour des caractéristiques non-observées qui ne varient pas entre les individus. De façon complémentaire, les effets fixes spatiaux (ou individuels) sont des composantes propres à une localisation, une observation, ou un groupe d'observations. L'ajout de ces effets dans les modèles de régression permet de contrôler pour des facteurs hétérogènes non-

observables, mais supposés constants dans le temps. Le modèle à deux effets fixes (*Two-Way Fixed Effect model - TWFE*) permet généralement une interprétation causale des coefficients obtenus.

Enfin, tout comme pour l'omission de variables explicatives pertinentes, l'omission d'effets fixes pertinents influence la valeur des coefficients estimés. Un coefficient biaisé est estimé si la variable explicative est corrélée aux effets fixes et que ces derniers ne sont pas inclus dans le modèle. Par ailleurs, l'ajout d'effets fixes parfaitement corrélés aux variables explicatives n'ajoute rien au modèle. En pratique, les variables redondantes sont retirées du modèle (ignorées) par les logiciels statistiques dans le processus d'estimation, autrement le modèle ne peut être estimé.

### **Analyses causales**

Pour isoler un effet causal, il faut exclure la possibilité qu'une autre source d'impact pourrait expliquer que la cause (x) est corrélée avec l'effet (y). La cause doit généralement précéder temporellement l'effet et une corrélation entre la cause et l'effet doit être observable.

#### *Le contrefactuel*

Pour isoler l'effet de la cause, plus communément appelé « traitement », une même observation devrait idéalement être observée simultanément sous les deux statuts (traité et non-traité). Cette situation permettrait d'isoler aisément l'impact d'un traitement, en comparant un résultat spécifique (y) pour l'observation lorsqu'elle est traitée ( $D_i = 1$ ) et lorsqu'elle ne l'est pas ( $D_i = 0$ ) (équation 3.6).

$$E(y_{i,D_i=1} - y_{i,D_i=0}) \quad (3.6)$$

Or, en pratique, il est impossible pour une observation d'être à la fois traitée et non-traitée. C'est le problème fondamental de l'analyse causale (Holland, 1976).

Les analyses causales nécessitent l'identification de contrefactuels, afin d'éliminer toutes autres explications potentielles quant à la corrélation statistique entre une (ou des) variable(s) explicative(s) et la variable d'intérêt. Le contrefactuel est une observation similaire, mais de statut inverse à l'observation sélectionnée. Par exemple, une observation traitée a pour contrefactuel une observation non-traitée, et vice-versa. Le calcul de la différence entre le

résultat observé dans un statut et le résultat attendu dans le statut inverse permet d'identifier l'effet de traitement moyen. Lorsque le calcul s'effectue avec les observations des deux groupes, soient les traités et les contrôles, on parle alors d'ATE (*average treatment effect*). En contrepartie, si le calcul se fait uniquement sur la base des traités et de leurs contrefactuels, on parle alors d'effet de traitement moyen sur les traités (*average treatment effect on treated* – ATT).

Empiriquement, la population à l'étude est mobilisée afin de créer des groupes de traitement et de contrôle. Ces groupes sont formés en fonction de la distribution du traitement. Les observations « traitées » sont classées dans le groupe traitement et à l'inverse, les observations qui ne sont pas touchées par le traitement sont potentiellement utilisées afin de former le groupe de contrôle.<sup>8</sup>

### *Expériences randomisées*

Dans l'analyse causale, les conditions de traitement idéales sont celles d'une expérience randomisée. Les expériences randomisées reposent sur l'attribution aléatoire du traitement. Cette distribution aléatoire implique des probabilités équivalentes, pour une observation donnée, d'être assigné au groupe traitement ou de contrôle. En moyenne, cela suppose que les caractéristiques des observations sont très similaires, voire identiques (Shadish et al., 2002). Ces conditions assurent la qualité des contrefactuels, l'interprétation causale et la validité des conclusions.

En pratique, il est difficile de mener une expérience randomisée. Les conditions préalables à la réception du traitement et les caractéristiques des observations qui composent la population étudiée (ou l'échantillon) influencent généralement la distribution du traitement. L'assignation endogène du traitement peut provenir de restrictions physiques, topologiques, éthiques, médicales, etc., selon le contexte de l'étude. Ainsi, certains traitements ne peuvent simplement pas être répartis de manière aléatoire. Les politiques publiques urbaines sont un

---

<sup>8</sup> Toutes les observations non-traitées ne sont pas nécessairement incluses dans le groupe de contrôle. Le contexte de l'étude, l'échantillon mobilisé et les observations disponibles peuvent influencer la composition des groupes et ultimement des observations dites « de contrôle ».

bon exemple de traitements qui ne peuvent être attribués de manière strictement aléatoire, puisque l'accès au traitement (et donc sa distribution) est généralement endogène.

### *Endogénéité du traitement*

La variable de traitement est dite endogène si une caractéristique non-observée, et donc comprise dans le terme d'erreur, a un impact sur la probabilité de recevoir le traitement. Tel que discuté précédemment, l'endogénéité du traitement découle de l'omission de variables corrélées à l'attribution du traitement, et donc à la variable d'intérêt. Les variables omises sont incluses dans le terme d'erreur et engendrent une estimation biaisée du coefficient associé au traitement (Kennedy, 2003).

Par ailleurs, un biais de sélection émerge de cette endogénéité, puisque les probabilités de traitement ne sont plus les mêmes pour toutes les observations, contrairement à une expérience randomisée. Les observations traitées disposent de caractéristiques (observées ou non) qui influencent la distribution du traitement.

### **Inférence causale**

Face à la distribution non-aléatoire du traitement, des stratégies d'identification causales sont nécessaires. Des méthodes alternatives aux expériences randomisées permettent d'isoler l'impact causal d'une politique publique sur une variable d'intérêt, dans un contexte de répartition non-aléatoire de la politique.

Les plus populaires prennent la forme (1) d'ajustement statistique, possible uniquement lorsque toutes les sources de variations de la variable d'intérêt sont connues et observables, et (2) de quasi-expérience, valide uniquement sous certaines hypothèses. L'inférence causale est possible si l'expérience peut être répétée et qu'elle est correctement modélisée (fiable).

### *Ajustement statistique : Méthode d'appariement*<sup>9</sup>

---

<sup>9</sup> La méthode d'appariement n'est pas mobilisée empiriquement dans ce projet de thèse. Elle est brièvement présentée dans cette sous-section, puisqu'il s'agit d'une approche méthodologique intéressante pour l'évaluation des impacts causaux des politiques publiques. Pour plus d'informations, le lecteur peut consulter : Rosenbaum et Rubin, 1983; 1984; D'Agostino, 1998.

La méthode d'appariement, ou *matching*, est un exemple d'ajustement statistique dans le cas d'une assignation non-aléatoire du traitement. Cette méthode a pour objectif d'identifier un comparatif crédible pour une observation et d'isoler l'effet causal du traitement étudié.

Le principe de la méthode d'appariement est simple : une observation de statut  $D_i$  (traité ou contrôle) est appariée, en fonction de ses caractéristiques observables,  $x_{kit}$ , à une observation du statut inverse. Cette méthode d'analyse causale opère sous l'hypothèse que seules les caractéristiques observables ont un impact sur la variable d'intérêt et la distribution du traitement. Il s'agit d'une hypothèse forte, puisqu'elle suppose une absence totale de variables latentes (caractéristiques non-observables).

L'appariement des observations peut s'effectuer à l'aide de différentes approches. D'une part, l'appariement peut reposer sur une seule variable qui renferme l'ensemble des informations. Cette approche est habituellement qualifiée de « score de propension » (*propensity score*), soit la probabilité qu'une observation soit traitée, en fonction de ses caractéristiques observables. Cette probabilité est estimée à l'aide d'une régression binaire (*logit* ou *probit*) sur la base des caractéristiques observables (Rosenbaum et Rubin, 1983). D'autre part, l'appariement peut être complété en fonction d'un ensemble de variables binaires, modales ou même continues (divisées en catégories) qui caractérisent les observations comprises dans l'échantillon. Enfin, l'appariement peut être basé sur la proximité (ou distance) entre les caractéristiques observables. Cette proximité peut être estimée en fonction des différences entre les variables explicatives.

Les observations similaires, avec une probabilité de traitement semblable, mais de statut inverse, sont appariées entre elles afin d'estimer l'effet causal moyen d'une politique publique (ATE ou ATT).<sup>10</sup> L'appariement d'observations avec des probabilités de traitement similaires contrôle pour un potentiel biais de sélection au sein des groupes de l'échantillon et assure l'obtention de contrefactuels solides (Abadie et Imbens, 2006; Caliendo et Kopeinig, 2008).

---

<sup>10</sup> Le choix entre l'ATE et l'ATT dépend de la taille de l'échantillon dans chacun des statuts respectifs. Une analyse reposant sur l'ATT est préférable lorsqu'un grand nombre d'observations de contrôle et un faible nombre d'observations de traitement sont disponibles.

### *Quasi-Expérience : Approche de Différence-en-Différences*

Les quasi-expériences tentent d'estimer l'effet causal d'un traitement assigné de façon non-aléatoire. Plus précisément, la spécification de type différence-en-différences (DiD) cherche à isoler l'effet d'un traitement (politique publique) sur une variable donnée. Les informations disponibles avant et après la mise en place de la politique et le statut d'exposition (contrôle ou traitement) sont mobilisés afin d'identifier l'effet du traitement.

Pour les quasi-expériences, le traitement, bien qu'il soit distribué de manière non-aléatoire, doit être exogène à l'observation. En d'autres termes, l'observation n'a pas d'influence sur sa probabilité de recevoir le traitement, ni sur le moment de la réception (dans le cas échéant).

Comme pour les expériences randomisées, l'approche DiD se base sur un échantillon divisé en deux groupes : un premier affecté par la politique publique (traitement) et un second qui n'est pas touché par cette même politique (contrôle). La différence moyenne entre les deux groupes est estimée avant et après la mise en place de la politique publique (Roth et al., 2023). Cette double différence issue de la dimension temporelle (avant et après) de l'étude du traitement caractérise la spécification DiD. L'écart observé entre les groupes, avant et après le traitement, devrait être attribuable à la politique étudiée, d'où le terme différence-en-différences (Équation 3.7).

$$ATE = E[(y_{i,P=1} - y_{i,P=0} | D_i = 1) - (y_{i,P=1} - y_{i,P=0} | D_i = 0)] \quad (3.7)$$

La variable binaire  $P_t$  identifie la période temporelle, soit avant ( $P_t = 0$ ) ou après ( $P_t = 1$ ) la mise en place de la politique étudiée. Les effets fixes temporels spécifiques à la période précédant le traitement sont captés par la variable de périodes. La variable binaire,  $D_i$ , comme précédemment, sépare les observations traitées des observations de contrôle et  $y_{it}$  est la variable d'intérêt.

Dans un modèle de régression, la spécification DiD s'insère dans les variables explicatives (Équation 3.8).<sup>11</sup> L'impact de la politique publique est reflété dans le coefficient  $\delta$ , associé

---

<sup>11</sup> Il est à noter que l'équation 3.8 présente uniquement la spécification DiD, et que des variables explicatives supplémentaires ou des effets fixes spatiaux (ou individuels) doivent minimalement être ajoutés au modèle afin de contrôler pour les caractéristiques (non-)observables et d'augmenter son pouvoir explicatif (Roth et al., 2023).



au croisement des variables binaires de temps/période ( $P_t$ ) et de traitement ( $D_i$ ). Ainsi, la variable croisée est égale à 1 si et seulement si l'observation est dans le groupe traité après l'introduction du traitement (Roth et al., 2023; Clark et Tapia-Schythe, 2021).

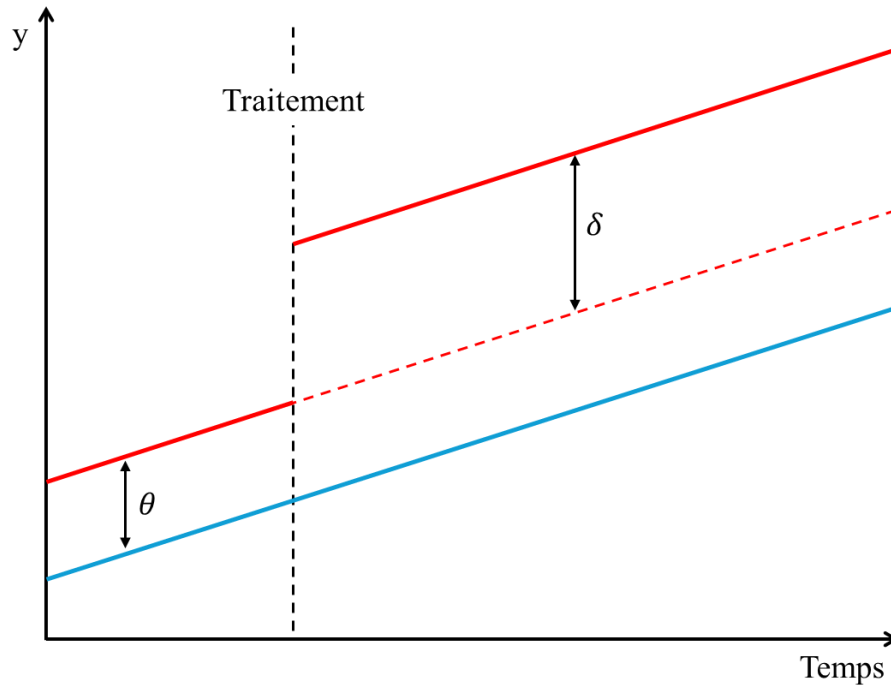
$$y_{it} = \alpha + \theta D_i + \gamma P_t + \delta (D_i * P_t) + \varepsilon_{it} \quad (3.8)$$

Le paramètre  $\theta$  indique l'écart entre le groupe de contrôle et de traitement, relativement à la variable d'intérêt. Cet écart moyen (et donc constant) représente la différence entre les deux groupes en l'absence de traitement.

Graphiquement, les droites du groupe de contrôle (en bleu – Figure 3.1) et du groupe traitement (en rouge – Figure 3.1) doivent être parallèles, avant la mise en place du traitement. Une différence initiale peut être observée entre les groupes ( $\theta$  dans la Figure 3.1 et l'équation 3.8). En pratique, sans nécessairement être identiques, des tendances similaires devraient être observées entre les groupes. Après la mise en place du traitement, l'écart entre les deux groupes peut changer et cette différence d'écart ( $\delta$  dans la Figure 3.1 et dans l'équation 3.8) est potentiellement attribuable à l'effet du traitement. Néanmoins, des tests supplémentaires doivent être effectués avant d'attribuer cette divergence entre les groupes uniquement à la mise en place du traitement.

Des hypothèses de base sont nécessaires afin d'isoler les effets associés à une politique publique par la méthode DiD. La plus importante, afin d'établir une relation de causalité, est l'hypothèse de tendances communes (Clark et Tapia-Schythe, 2021). Cette hypothèse repose sur la composition et la comparabilité des groupes de traitement et de contrôle. La similitude des groupes permet d'isoler l'effet du traitement et assure que le traitement est responsable de l'impact estimé. La différence entre les groupes engendre la soustraction des variables non-observées/facteurs confondants qui sont partagés par les deux groupes, laissant uniquement l'effet du traitement comme résultat de la différence.

**Figure 3.1 - Représentation graphique de l'approche de différence-en-différences**



Si les deux groupes ne sont pas comparables, alors l'ATE est biaisé par la présence de facteurs non-observés spécifiques à la composition d'un seul des deux groupes. L'hypothèse des tendances communes stipule également que les deux groupes se seraient comportés de manière similaire sans la mise en place de politique, toutes choses étant égales par ailleurs. Il est donc primordial de tester cette hypothèse, afin d'isoler correctement l'impact causal du traitement.

Pour tester l'hypothèse des tendances communes, un test placebo peut être effectué. Ce test consiste à utiliser un sous-échantillon comprenant les observations recensées avant la date du traitement. Une date fictive est ensuite fixée afin de générer deux nouvelles périodes temporelles, soit avant et après cette date fictive/placebo (Roth et al., 2023). À l'aide des groupes traitement et de contrôle mobilisés dans le modèle, il est possible d'estimer l'impact du traitement fictif (placebo) sur le sous-échantillon. Puisqu'aucun réel changement n'est survenu, les paramètres estimés devraient être non-significatifs. Sous l'hypothèse nulle que les deux groupes sont comparables, l'hypothèse des tendances communes ne peut pas être rejetée si l'impact estimé du traitement fictif est non-significatif (Slusky, 2017).

La spécification DiD a certaines limites. En effet, sans extension l'approche DiD évalue l'impact moyen du traitement, en supposant que l'effet est similaire sur l'ensemble de la période post-traitement. Cependant, lorsque la mise en place d'une politique s'effectue en plusieurs étapes, il est important d'évaluer chacune de ces périodes (Dubé et al., 2018). Par exemple, la mise en place d'une infrastructure de transport s'étend sur plusieurs années. L'annonce du projet, la construction et l'inauguration peuvent s'étirer dans le temps et des effets de spéculation peuvent être observés dès l'annonce du projet. Avec des périodes distinctes, isoler des effets spécifiques de spéculation et d'anticipation est possible (Devaux et al., 2017; Dubé et al., 2018; Champagne et al., 2022).

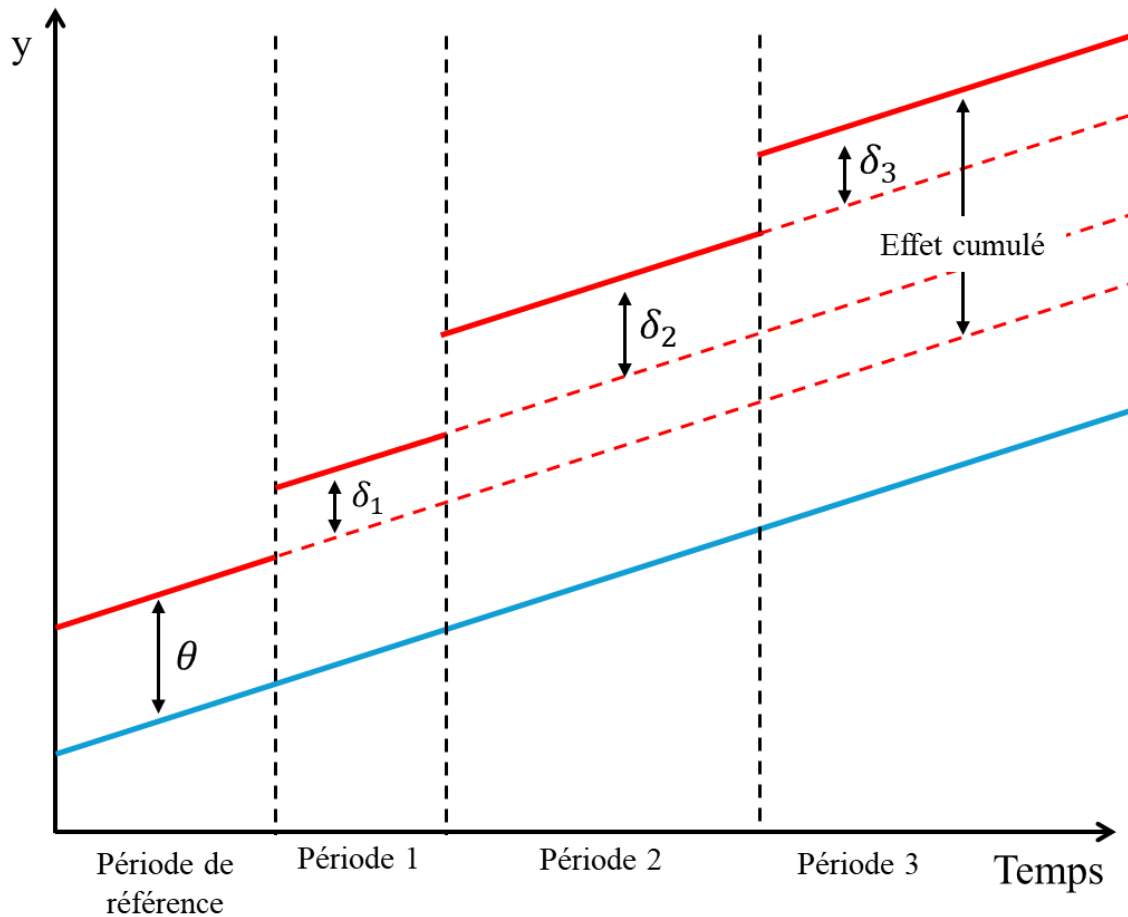
Mathématiquement, intégrer un vecteur de périodes (ou encore un ensemble de variables binaires), plutôt qu'une seule variable binaire temporelle (avant/après) peut suffire afin d'identifier et d'isoler les impacts spécifiques par période (Dubé et al., 2018). Un effet de traitement différencié par période s'obtient par la somme du produit croisé entre les variables de traitement ( $D_i$ ) et de périodes ( $P_r$ ). Le produit croisé de ces variables est égal à 1 si et seulement si l'observation  $i$  est dans le groupe traité durant la période  $r$  (Roth et al., 2023). Le paramètre d'intérêt ( $\delta_{ir}$ ) permet d'isoler l'impact causal de la politique publique étudiée sur la variable d'intérêt ( $y_{ir}$ ) pour la période  $r$  (Équation 3.9).

$$y_{ir} = \alpha + \theta D_i + \sum_{r=1}^R \gamma_r P_r + \sum_{r=1}^R \delta_{ir} (D_i * P_r) + \varepsilon_{ir} \quad (3.9)$$

Comme dans l'équation 3.8, la variable  $D_i$ , pour l'observation  $i$ , prend la valeur de 1 ou 0 en fonction de son statut, traité ou de contrôle. L'ensemble de variables  $P_r$  représente des variables binaires permettant d'identifier chacune des  $R$  périodes étudiées, où  $r = \{1, \dots, R\}$ . La variable  $P_r$  est égale à 1 si l'unité est observée durant la période  $r$ , et 0 autrement. La première période ( $P_{r=0}$ ) est utilisée comme période de référence, afin d'éviter une colinéarité parfaite entre les observations (Figure 3.2).

La spécification DiD est très utile dans un contexte d'analyse de politique publique, sous le respect des hypothèses de l'approche. Cette modélisation peut aisément s'intégrer aux variables explicatives de différents modèles, afin d'évaluer un large éventail d'impacts causaux.

Figure 3.2 - Impact par période et cumulé dans la spécification DiD



### Conclusion

En somme, les analyses causales reposent sur l'identification d'un contrefactuel crédible et le respect de certaines hypothèses dans un contexte de quasi-expérience. Dans une perspective d'évaluation de politiques publiques, il serait erroné d'assumer une assignation aléatoire de la politique sur le territoire. La mobilisation d'analyses causales, sous la contrainte d'un traitement assigné de manière non-aléatoire, est donc nécessaire afin d'évaluer les impacts des infrastructures de transport sur des variables d'intérêt, telles que la localisation des entreprises et leur performance.

En complément à la littérature inhérente aux impacts des infrastructures de transport sur les décisions de localisation des entreprises (voir Chapitre 1), l'analyse de la performance commerciale s'avère pertinente, surtout dans une optique de longévité commerciale.

Simplement, une localisation optimale devrait faciliter la survie de l'entreprise dès son entrée sur le marché. De plus, la probabilité de survie souligne la capacité d'une entreprise à s'adapter à son marché et à acquérir un bassin de consommateurs.

Pour ce faire, les deux prochains chapitres de cette thèse s'intéressent aux impacts causaux des infrastructures de transport sur la performance commerciale en milieu urbain. Plus précisément, chaque chapitre évalue empiriquement les effets d'une infrastructure de transport en commun sur la longévité des entreprises de service. Dans les deux cas étudiés, des contrefactuels crédibles sont identifiés et une spécification DiD est intégrée au modèle, dans le but d'isoler les effets causaux du transport en commun sur la probabilité de survie des entreprises.

## **CHAPITRE 4 : Standing strong? The causal impact of metro stations on service firms' survival**

### **Résumé**

Cet article s'intéresse à l'impact d'un métro sur la performance des entreprises. Plus spécifiquement, l'analyse empirique se concentre sur l'impact de l'extension d'une ligne de métro sur la probabilité de survie des entreprises. Pour ce faire, un modèle de risques proportionnels de Cox est estimé autour du métro de Montréal entre 1996 et 2016. L'annonce simultanée de deux projets d'extension, suivi de l'abandon d'un de ces projets, propose un scénario intéressant afin de définir le groupe de traitement (nouvelles stations) et de contrôle (stations spéculatives). Le modèle estime des effets par période, afin d'identifier de potentiels effets d'anticipation et de spéculation. Les stations de métro semblent avoir un impact positif sur la probabilité de survie des entreprises pendant la période de construction, mais surtout après l'ouverture des nouvelles stations. Les résultats suggèrent que l'effet est concentré de 250 à 1 250 mètres des stations, plus particulièrement pour les activités locales.

**Mots-clés :** Transport en commun, Analyse de survie, Entreprises, Modèle de risques proportionnels de Cox.

## **Abstract**

The aim of the paper is to investigate the impact of the extension of a metro line on the survival of individual firms. An empirical analysis focuses on the relation between proximity to new stations and firm survival following the announcement of the extension of the orange line of the metro in Montréal (Canada) between 1996 and 2016. To do so, a Cox Proportional Hazards model is estimated using the simultaneous announcement of two potential extensions to define the treatment (new stations) and control groups (speculative stations) based on the distance to the closest stations. The model explicitly controls for anticipation and speculation effect by introducing three distinct treatment effects by period. The impact of the new metro stations appears to be mostly positive on firms' survival probability during the construction period and after the opening of the service. The results suggest that the metro extension does have a positive influence on the survival rates of individual firms within 250m to 1,250m of the closest station, especially for local activities.

**Keywords :** Public transit; Survival analysis; Firms; Cox Proportional Hazards model

## **Introduction**

Public transport infrastructure is assumed to be an important piece of the adaptation of cities to climate change and sustainable development (Akin and Kara, 2020). It can act as an important solution to reduce car dependence and greenhouse gas emissions within large urban areas. By affecting the accessibility and attractiveness of some specific nodes, and therefore impacting factors inherent to the enhancement of commercial activities, public transit services influence the location decision of economic agents. While most of the potential impacts of urban transit systems have been largely studied based on individual firms' performances such as creation (Credit 2018; 2019), agglomeration economies (Chatman et al., 2016; Alañón-Pardo and Arauzo-Carod, 2013; Takeda et al., 2008), and location (Bazin et al., 2009; Beckerich et al., 2016; Champagne and Dubé, 2023), not much has been revealed about how it could affect firms' survival, i.e., firms' longevity. Survival probability, or the time expended by avoiding closure, is a way to study how a firm's individual and spatial characteristics influence the number of days the firm operated before closing (if applicable - Belleau-Arsenault and Dubé, 2019). This is an indirect way of investigating how public policy, such as a change in transport infrastructure, might influence the performance of individual firms.

This paper aims to evaluate the impact of public transport infrastructure on firms' longevity, and more specifically the impact of metro stations on firms' survival. To do so, the analysis is based on two distinct public transport infrastructure extensions to define treatment and control areas. The empirical investigation is based on a Cox Proportional Hazards model. The main assumption to be tested in this article is whether proximity to new public transit stations influences the survival probability of individual firms. As new stations should attract potential consumers and facilitate workers' accessibility, we investigate whether the infrastructure reduces the risk of closure, or inversely increases the survival rates. We also test to verify whether the effect differs according to specific economic activities in a given service industry, i.e., retail, finance, insurance and real estate (FIRE), knowledge-intensive business services (KIBS), and local activities, since public transit facilitates human encounter and increases contact with potential customers.



The analysis uses distance-based measure variables for individual firms located within a 1,500-meter radius around the new stations of the orange metro line extension in Montréal (Canada) and the fictive stations of the extension project of the blue line between 1996 and 2016. Both projects were simultaneously announced in 1998. While one of the projects was completed, the other (blue line) has not materialized yet. The situation provides an interesting quasi-experimental approach to investigate the causal effect of transport infrastructure on firms' survival.

The originality of the paper is twofold. First, it aims to fill the gap in the literature regarding the causal impact of transport infrastructure on the performance of individual firms from a survival perspective. Second, the choice of the “control” area based on a similar extension project that was never realized provides an interesting quasi-natural counterfactual.

The remainder of this paper is divided into five sections. Section 1 presents the literature inherent to this article. Section 2 describes the methodological approach, while Section 3 introduces the project under study and the data used for the analysis. Section 4 presents the results and discusses the most relevant ones. Finally, Section 5 summarizes the obtained results and the policy implication and provides some discussion for implications on future research.

### **Related literature**

The literature regarding the determinants of firm survival is quite recent. Survival analysis adapted to social sciences dates back to the early 1990s (Dunne et al., 1988; Audretsch and Mahmood, 1995).<sup>12</sup> While the literature focusing on the impact of transport infrastructure on firms' performance has been studied under various angles recently (Beckerich et al., 2016; Beaudoin and Lawell, 2018; Beaudoin et al., 2015; Chatman et al., 2016), little attention has been given to the survival perspective.

Literature suggests that firm-level characteristics, such as age, size, and sector, highly affect firms' survival rates (Tveterås and Eide, 2000; Evans, 1987; Bates, 1990; Headd, 2003;

---

<sup>12</sup> See also Audretsch (1995), Doms et al. (1995), Mata et al. (1995), Geroski (1995) and Caves (1998), among others.

Phillips and Kirchhoff, 1989; Bates and Nucci, 1989). New firms have a smaller chance to survive compared to mature firms (Pérez et al., 2004; Caves, 1998). Geroski (1995) demonstrates that it is harder for a firm to survive for a certain time in a competitive industry (*barrier to survival*) than to enter this same industry (*barrier to entry*). In fact, the barrier to entry can indicate a business's likelihood to survive, since lower cost of entry makes it easier for firms to enter the market, making it more competitive. The entering firms also face a greater risk of failure because they must quickly achieve efficiency in order to reach a minimum competitive level, which is unknown until the firm enters the market (Pérez et al., 2004; Stinchcombe, 1965).

Furthermore, smaller firms have a lower survival chance compared to bigger firms (Geroski et al., 2010; Evans, 1987; Mata and Portugal, 1994; Audretsch and Mahmood, 1995; Sterns et al., 1995; Storey and Wyncarczyk, 1996). Start-ups are usually smaller than well established and mature firms, meaning that smaller firms have a greater chance of experiencing the *liability of newness* (Stinchcombe, 1965). Smaller firms may also face a cost disadvantage and have fewer opportunities in terms of labor and capital, compared to bigger firms (Pérez et al., 2004).

The economic sector of the firms affects their specific needs and behavior, regarding their consumers and employees, which can in turn affect firms' risk of failure. Therefore, the economic sector may influence business decisions, such as location, as they want to operate at the center of their market to take advantage of spatial attributes and local amenities (Credit, 2019; 2018; Chatman and Noland, 2011).

For example, manufacturing firms prefer to be located close to transport infrastructures such as highways, ports, or even airports, to reduce the transport costs of their inputs and outputs (Champagne and Dubé, 2023; Verhetsel et al., 2015). They also want to be far enough from a city to benefit from a reduction in land costs (Holl, 2004a; 2004b). Inversely, service firms would prefer to be located close to the central business district (CBD) or at the center of their market to take advantage of agglomeration economies and public infrastructure, such as urban transit systems (Credit, 2019; 2018; Chatman and Noland, 2011).

Beyond individuals' characteristics, firms' performance and survival may be influenced by other factors such as spatial attributes and local amenities (i.e., transport infrastructure).

Spatial attributes, such as agglomeration economies, are also discussed in the literature on firms' survival (Chatman et al., 2016; Tveterås and Eide, 2000; Manzato et al., 2011; 2010). The agglomeration economies are defined as the decrease in average production costs (associated with an increase in the level of production) within a specific region and arise from the spatial concentration of commercial activities (Chatman and Noland, 2011). The advantage of spatial concentration can be either related to specialization or diversification.

On the one hand, specialization of the market implies the concentration of similar economic activities and causes an increase in local competitiveness (Marshall, 1891). On the other hand, firms may be attracted by a more diverse environment and prefer to locate close to complementary economic activities, rather than similar ones (Jacobs, 1969). In other words, agglomeration economies, such as sharing of the labor pool and consumers, and knowledge externalities increase the firms' incentives to spatially concentrate (Chatman and Noland, 2011; Arzaghi and Henderson, 2008). In the literature, agglomeration economies are usually addressed using a proxy, location quotient and/or Herfindahl index, as explanatory variables to evaluate the statistical links between economies of agglomeration and firms' performance (Yi and Nam, 2019; Low and Brown, 2017; Renski, 2011; Pe'er et al., 2008; Pe'er and Keil, 2013).

Transport infrastructure can also influence firms' performance by decreasing transportation costs (Glaeser et al., 2008) and therefore influencing the spatial distribution of economic agents (Fujita, 1989). While public transit stations tend to increase rents in their direct vicinity, they also appear to increase the pedestrian flow and ensure a closer proximity to consumers (Ganning and Miller, 2020). Moreover, the reduction in travel time could also influence firms' global performance, such as their growth and survival (Manzato et al., 2011; 2010), since firms aim to minimize their costs (including transportation costs).

Locations close to the public transport stations greatly facilitate the movement of workers and business recruitment and increase the attraction of consumers (Mejia-Dorantes et al., 2012; Chatman et al., 2016). Thus, the convergence of service firms might be influenced by public transit infrastructure, since their benefits are identified over short distances (Hurst and

West, 2014; Arzaghi and Henderson, 2008). In fact, the impact of public transit on commercial activities is rarely observed beyond an 800-meter radius (Hurst and West, 2014; Cervero and Landis, 1997; Goetz et al., 2010). Overall, the studies on urban transit systems show that public transit mostly benefits service firms, since their economic activities are based on human contacts and interpersonal relationships (Credit, 2019; 2018; Mejia-Dorantes et al., 2012; Chatman et al., 2016).

Empirical evidence on the impact of transport infrastructure on firm survival is scarce, and even scarcer for public transit. Most of the studies have been conducted from a firm's birth perspective, but the results are still ambiguous. The impact of public transit on firms' birth (or location decision) has led to either positive (Credit, 2018; 2019; Iseki and Eom, 2019; Song et al., 2012) or non-significant impact (Schuetz, 2015; Leitham et al., 2000; Castillo-Manzano and López-Valpuesta, 2009), suggesting that the impact of public transit on a firm's location is place dependent (Chatman et al., 2016; Champagne and Dubé, 2023).

Few qualitative studies have investigated the impact of light rail transit (LRT) on firms' performance, by questioning managers and owners of firms located within the direct vicinity of the newly constructed infrastructure (Tornabene and Nilsson, 2021; Sukaryavichute et al., 2021). The results suggest the importance of temporal specifications, since the construction period appeared to have a massive negative impact, especially on small local businesses. The construction period usually leads to reduced visibility and accessibility, and lower foot traffic (i.e., loss of consumers and decrease of sales volume) (Tornabene and Nilsson, 2021; Sukaryavichute et al., 2021). Results also suggest that firms are most at risk if there are substitutes outside the construction corridor that can absorb their consumers (De Solminihac and Harrison, 1993).

Using a quantitative approach, Ray (2017) evaluates the impact of proximity to an LRT on firms' probability of failure. The results suggest that firms operating within a 400-meter radius of the LRT have a 46% higher chance of closing, compared to firms operating outside the LRT corridor. However, the study mainly focuses on small firms, which are mostly at risk of closure, and the author could not identify by which mechanism business failure occurs.

Finally, Manzato et al. (2011; 2010) study the impact of railways and highways on firms' survival. To the author's knowledge, those studies are the only ones using survival models to investigate the impact of transport infrastructure on firms' risk of failure. Their results indicate that proximity to a highway or a railway may increase the survival rate of the firms operating mainly in retail, finance, real estate, social security, business services, computer and information technology, and research and development.

## **Methodology**

The most common survival model in the literature is the proportional hazards (PH) model with explanatory variables, first introduced by Cox (1972). This model was mainly used in the medical domain to analyze the time between the recurrence of a disease, the probability of survival of an ill patient, and the potential occurrence of death (Andersson and Lambert, 2012; Hinchliffe et al., 2013). It has recently been adapted and applied to analyze the lifetime of the firms and their probability of closure, rather than a patient, and their probability of death. Using the Cox PH model, we evaluate the probability that a firm  $i$  is still operating at a time  $t$ , conditional to its survival until  $t-1$  (Kleinbaum and Klein, 2005; Fotopoulos and Louri, 2000; Kiefer 1988).

Survival analyses must consider the censoring of their distribution, in order to obtain unbiased results (Kiefer, 1988; Geroski et al., 2010; Rich et al., 2010). One of the main advantages of the Cox PH model is the possibility to correctly estimate the model by explicitly incorporating the concept of censored data. The censoring comes from the firm surviving beyond the study period (right-censoring), or from the firm opening before the study period (left-censoring).

To study the impact of a treatment (metro station) on the firms' survival probability, the hazard rate is estimated. The hazard rate for a firm  $i$ , at the time  $t$  is assumed to be related to its characteristics, but also to its local environment, resumed in a matrix  $X$  (Equation 4.1).

$$h(t, X) = h_0(t) e^{X\beta} \quad (4.1)$$

Where  $h_0$  is the baseline hazard function, which represents the hazard function when all the regressors are set to 0, or  $X = 0$  (Kiefer, 1988). The Cox PH model is a semi-parametric

model, because the baseline function is not explicitly defined, while an exponential function is assumed for the survival time (effects of the explanatory variables) (Manzato et al., 2011; Kleinbaum and Klein, 2005). The undefined distribution of the baseline hazard function allows the model to overcome problems linked to unobserved heterogeneity, which could emerge if the baseline distribution was incorrectly specified (Kleinbaum and Klein, 2005; Pérez et al., 2004).

Therefore, another advantage of the Cox PH model stems from the possibility of estimating the model without specifying the form of the baseline hazard function. With a partial likelihood method, the vector of coefficients,  $\beta$ , can be estimated for any baseline hazard function,  $h_0(t)$  (Kleinbaum and Klein, 2005; Tveterås and Eide, 2000).<sup>13</sup>

To compare the probability of survival of different firms, the hazard ratio (HR) is used (Equation 4.2). The HR establishes the relative impact of the independent variables on the risk of failure and, inversely, the probability of survival. The relative probability of survival is the ratio of the hazard rates (Equation 4.1), where the rate of firm  $i$  (numerator) is compared to the rate of firm  $j$  (denominator), which is used as a reference (Manzato et al., 2010; Tveterås and Eide, 2000). The exponential form of the hazard functions ensures positive coefficients.

$$\widehat{\text{HR}} = \frac{\widehat{h}(t|X_i)}{\widehat{h}(t|X_j)} = e^{\sum_{k=1}^K (x_{i,k} - x_{j,k}) \widehat{\beta}_k} \quad (4.2)$$

Where  $x_{i,k}$  and  $x_{j,k}$  are respectively the  $k^{th}$  characteristic/explanatory variable of the firms  $i$  and  $j$ , with  $i \neq j$ , and  $j$  being in the reference group. The interpretation of the HR is relative to the reference group, so if the ratio equals one, there is no effect of the treatment on firms' survival; if the ratio is greater than 1, the probability of survival is lower (risk of failure is higher) for the treated area as compared to the reference group, while the probability of survival is higher when the ratio is between 0 and 1 (Anavatan and Karaoz, 2013; Spruance et al., 2004).

---

<sup>13</sup> For more information on the estimation of coefficients with the partial likelihood method, refer to Cox and Oakes (1984).

Contrary to the hazard rates (Equation 4.1), the factor of time ( $t$ ) is absent from the HR (Equation 4.2), meaning that characteristics must be proportional over time (Anavatan and Karaoz, 2013). The Proportional Hazard Assumption (PHA) can be formally tested using the Schoenfeld residuals.<sup>14</sup>

The Cox-Snell residuals can be used to investigate the goodness-of-fit of the model, as a postestimation test. Cox-Snell residuals estimate the difference between the predicted and real (observed) survival function. To do so, it applies a negative logarithmic transformation to the survival function of each observation (Cox and Snell, 1968). The Cox-Snell plots are characterized by the cumulative hazard on the y-axis, and the Cox-Snell residuals on the x-axis. To ensure the goodness-of-fit, the plotted cumulative hazard should follow a straight diagonal line, which assumes that the observed and estimated survival function are the same (model perfectly fitted). Those residuals are standardized and commonly applied in reliability and survival analysis (Ansin, 2015; Dauda et al., 2015; Stepanova and Thomas, 2002).

## **Case study and data**

### *Montréal Metro*

The metro of Montréal was inaugurated in 1966. Nowadays, the metro comprises 4 lines and 68 stations on the Island of Montréal, as well as one station on Montréal's South Shore (Longueuil) and three stations on the North Shore (Laval). The stations on the North Shore are the last ones added to the total network of stations in 2007 (Société des transports de Montréal, 2020).

With the last extension, the orange line became the first metro line to cross the *Rivière des Prairies*, which physically separates the Island of Montréal from the City (and island) of Laval. Located in the northern suburbs of Montréal, Laval had a population of 423,000 in

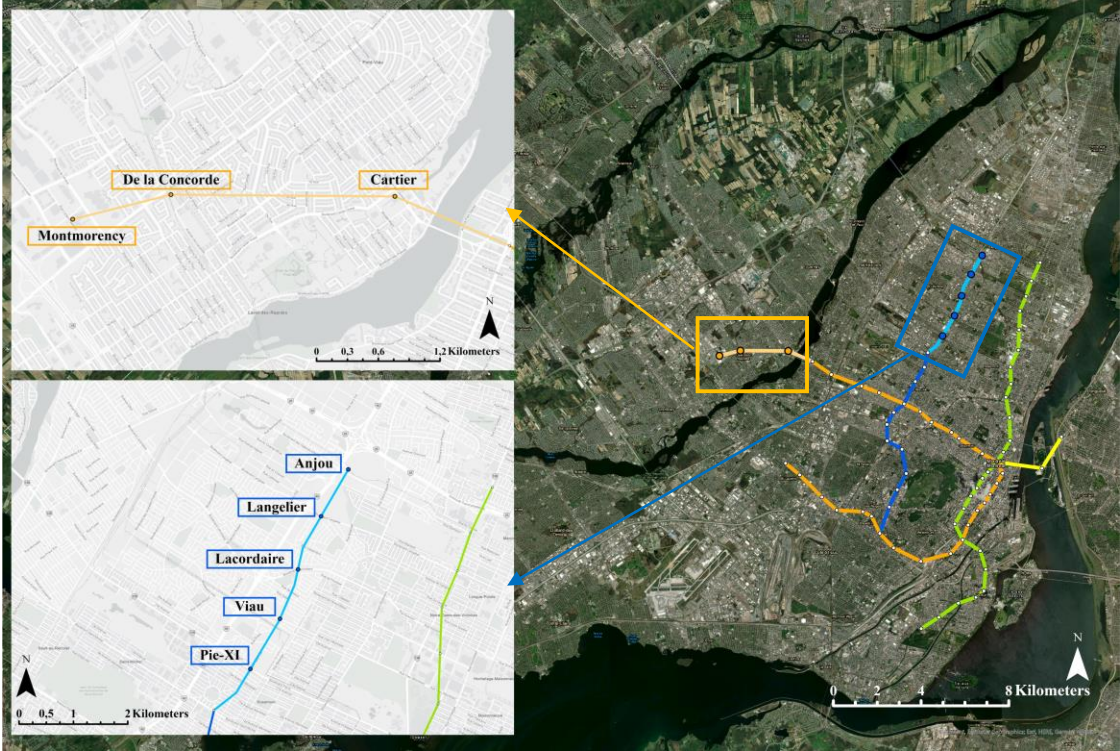
---

<sup>14</sup> Schoenfeld residuals individually test each explanatory variable to ensure that all variables respect the PHA. Under the null assumption, the logarithmic function of the hazard ratio should be constant (Kuitunen et al., 2021; Grambsch and Therneau, 1994; Stata, 2021). Using the Schoenfeld plots, the slopes should be close to zero, indicating that the explanatory variables are independent of the time (Kuitunen et al., 2021; Grambsch and Therneau, 1994). If the null hypothesis can be rejected the PHA does not hold, suggesting that the variable violates the null assumption and should not be used for the Cox PH regression, since it could produce bias results (Stata, 2021).

2016, which represents approximately 11% of the urban population of the Greater Montréal area (Statistics Canada, 2021). Laval represents a significant labor and consumer pool. More than half (53%) of the population works on the Island of Montréal and 82% of the workforce operates in the service sector (Devaux et al., 2017), reinforcing the aim to facilitate accessibility to the downtown area.

On October 7, 1998, the extensions of the orange and blue lines were announced. The submitted extensions included 3 stations towards the City of Laval (west direction – orange line) and 5 stations towards the east of the island (blue line) (Figure 4.1). While the orange extension was finished in 2007, the blue line did not materialize due to a lack of funding (Gilbert and Baker, 2020). The simultaneous announcement and the extension of the metro represent an interesting quasi-natural experiment for analyzing the impact of public transportation on firms’ survival.

**Figure 4.1 - The submitted extensions of the orange and blue lines in October 1998**



Reference: STM (2020)

While some empirical experiences have tried to focus on the surroundings of the extension as a counterfactual (Diao et al., 2017; Dubé et al., 2018; Credit, 2018), the situation of Laval is particular. The extension occurs in a part of the city where residential and commercial



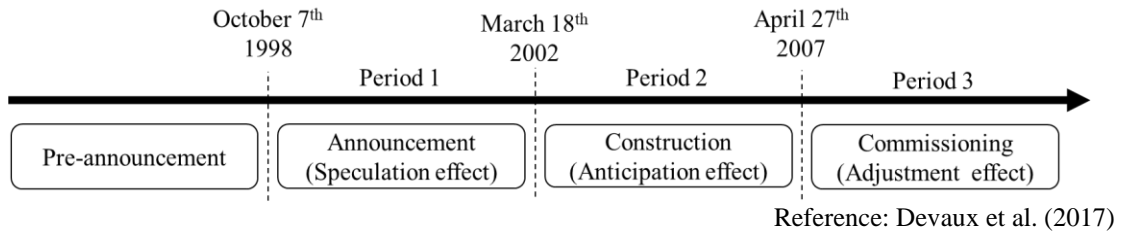
densities are relatively high as compared to other parts of the city. Also, the speculation effect, which might be important when the announcement of a major infrastructure is made, does not necessarily materialize before the start of construction.

This is especially true regarding transport infrastructure in the region of Montréal. Many adjustments only happen after the work begins, since the announcement might never materialize. For example, the pink line, an additional line to the Montréal metro, was officially announced in 2011, but abandoned ten years later in 2021 (Radio-Canada, 2018; Tiranti, 2021). Regarding the blue line, after being abandoned in 2001, announced again in 2013, and a third and final time in 2017, the *Société de Transport de Montréal* (STM) began construction in June 2022, 24 years after its first official announcement (Le Devoir, 2013; Secrétariat du Conseil du Trésor, 2023).

For those reasons, the northern part of Montréal and the project of extension of the blue line appear as a more interesting counterfactual. First, the neighborhood can be compared to where the extension of the Metro occurs in Laval regarding the density of commercial and residential activities. Second, the simultaneous announcement of the extension projects also implies that the two locations had similar anticipation regarding the adjustment in public transit supply. The basic scenario also suggests that the firms would have no way of knowing that the project would be canceled, much less which extension would materialize over the other, which ensures some comparison between the two groups before 2001. Third, for the specific case of Laval and its metro extension, speculation effect is weak, and businesses tend to respond to concrete action, such as construction work, rather than announcement (Champagne et al., 2022).

On March 18, 2002, the construction of the orange line's extension began, and the 3 new stations opened on April 27, 2007. Using those dates, the analysis is separated into three distinct periods that could influence the perception of the market (Figure 4.2). Those periods are used to account for potential speculation, anticipation, and adjustment effect.

**Figure 4.2 - Temporal evolution of the extension of the orange metro line in Laval**



### *Data and Descriptive Statistics*

The data was obtained from the *Registraire des entreprises du Québec* (REQ), which contains all firms that are (or have been) active in the province. Each firm is associated with a *Numéro d'Entreprise du Québec* (NEQ). The NEQ is a unique number and allows firms to be tracked on an individual scale. Moreover, the addresses are known and geocoded. The REQ database contains all the opening (and closing) dates of individual firms operating in the province of Québec. Firms opening between 1996 and 2016 are used in this study. To be included in the analysis, the firm must stay active for at least 30 days.

For all firms, the data set contains information about their main economic sectors. The economic activities are separated into 15 aggregated categories based on their economic activities (defined by the *Economic Activity Code – EAC*), which are summarized in *Annexe du Chapitre 4* (Table A4.1). Those categories are used to evaluate how the firms' survival rates might vary according to different economic sectors by studying the impact on a specific service industry (Retail, FIRE and KIBS, or Local Activities). To avoid false longevity, public administration, education, health, and social services activities are excluded from the sample. Agriculture, manufacturing, construction, and logistics activities were also excluded since they usually operate outside of urban areas and are not really attracted or influenced by public transport infrastructure (Verhetsel et al., 2015; Holl, 2006).

The data set also includes information about the size of the firms. This variable is often used in the literature on firms' survival, as it historically appears to be one of the most significant variables linked to a firm's survival (Evans, 1987; Mata and Portugal, 1994; Audretsch and Mahmood, 1995; Sterns et al., 1995; Storey and Wynarczyk, 1996). The number of employees served as a proxy to rank the firms into one of the six categories: (1) Self-employed worker (0 to 1 employee), (2) between 2 and 10 employees, (3) between 11 and 25 employees, (4) between 26 and 49 employees, (5) between 50 and 99 employees and (6)

more than 100 employees. The first category (0 to 1 employee) is used as a reference. The current size (last available) is used instead of the initial size, because it captures the firms' ability to adapt to their environment (Mata et al., 1995).

The opening (and closing) dates are useful to analyze firms' lifetime and to define the censoring. Only firms that opened during the study period are retained for further analysis. Therefore, left-censoring does not apply to the sample used, but right-censoring is clearly an issue that must be explicitly addressed. A firm  $i$  is right-censored if it is still open in 2017 (Geroski et al., 2010; Rich et al., 2010). The dummy "Open" is an event indicator, and thus of censoring. If the event, in this case the closure of the firm  $i$ , happens before the end of 2016, the dummy "Open" will be equal to 0, and 1 otherwise (Equation 4.3). To be considered right-censored, a firm would have to still be open after the end of the study period (Open = 1).

$$Open_i = \begin{cases} 0 & \text{if } t_i \leq T \\ 1 & \text{if } t_i > T \end{cases} \quad (4.3)$$

Where  $t_i$  represents the closing date of the firm  $i$  and  $T$  is the censoring point of the model, which is the end of the study period (December 31, 2016). Approximately 62% of the observations in the sample close during the period studied, on either the treatment or control group (Table 4.1).

**Table 4.1 - Firms' opening and closure (descriptive statistics)**

Line	Censored Still open after 2016	Closed During the study period	Opening (Total) Firms sample	Closing Rate
Orange	3,003	4,948	7,951	62.2%
Blue	7,335	12,209	19,544	62.5%
Total	10,338	17,157	27,495	62.4%

The explanatory variables used to evaluate the impact of public transit on firms' survival are common in the literature on either firms' survival or location decision (Credit 2018; 2019; Devaux et al., 2017; Manzato et al., 2010; 2011; Yi and Nam, 2019).

The distance to the closest station is calculated using the individual geographical coordinates of the firms and of the stations. The Euclidean distance is a time-unvarying measure of distance and ensures a constant geographic area to delimit treatment and control areas for the

entire study period. While imperfect, the Euclidean distance offers the advantage of relying on a single layer of information as compared to an accessibility index, where the construction of isochrones needs to take into account changes in public transport routes and changes in road designs. However, the correlation in both variables is usually high (Apparicio et al., 2008), especially in an urban context where no major highways or rail lines pass by (Higgins and Kanaroglou, 2017), which is the case for our case study.

Using the distance to the closest stations, firms within a 1,500-meter radius of both lines are identified with their coordinates and retained for further analysis. However, firms located on the Island of Montréal even within 1,500 meters of the first station (Cartier) are discarded from the sample, since the impact of the selected stations should not extend over the river which physically separates the islands of Laval and Montréal.

Although the literature mainly identifies the impact of public transit within 800 meters of its infrastructure, the radius is extended to 1,500 meters to capture potential spillover effects ranging over a typical range. Moreover, during the studied time period (1996-2016) in Laval, real estate construction of commercial buildings around the orange extension has been identified to locate within a 750-meter radius. Residential constructions, for their part, were mostly concentrated in the 1,000- to 1,250-meter area (Champagne et al., 2022). The construction of commercial buildings mainly took place after 2007, while new residential entities were built over the entire period studied. Those trends in local real estate may impact the survival dynamics of firms. Using the observations located within 1,500 meters of the stations, two groups are created, i.e., treatment (orange line) and control (blue line). To avoid capturing spillover effects from the green line on the control group, which is parallel to the blue line (Figure 4.1), the studied area is limited to a 1,500-meter radius. To study local spatial variation, the zone is also subdivided into exclusive 250-meter zones around each station (real and projected).

A set of dummy variables are also created to generate interactive variables and isolate the causal effect based on the predetermined groups (treated or control group), as well as on temporal dummy variables to identify during which period firms have opened or closed. The results of the cross-product variables represent the variables of interest in the interpretation of the model to analyze the impact of public transit on firms' survival, by time period and

zone (Table 4.2). The observations of firms within the two groups before and after specific periods allow an interesting comparison and the use of a difference-in-differences (DiD).

**Table 4.2 - Distribution among the interaction variables (descriptive statistics)**

Observation	0m - 250m	250m - 500m	500m - 750m	750m - 1000m	1000m - 1250m	1250m - 1500m	
Ref.	Orange Line	23	95	152	155	154	105
	Blue Line	103	251	319	335	311	285
	Total	126	346	471	490	465	390
Ann.	Orange Line	34	179	224	262	275	212
	Blue Line	212	448	554	475	582	517
	Total	246	627	778	737	857	729
Constr.	Orange Line	59	201	253	349	291	286
	Blue Line	300	574	766	653	777	646
	Total	359	775	1019	1002	1068	932
Service	Orange Line	102	492	528	692	539	476
	Blue Line	587	1078	1381	1276	1364	1159
	Total	689	1570	1909	1968	1903	1635
Total	Orange Line	218	967	1157	1458	1259	1079
	Blue Line	1202	2351	3020	2739	3034	2607
	Total	1420	3318	4177	4197	4293	3686

Note: The number of firms that operated during the period is presented. Therefore, firms that closed during the period or opened before the beginning of the period are still included in the time period, since they did operate during the time covered by the period.

The DiD approach is quite appropriate and known in the context of linear regression models but is rarely used in other specifications. Nevertheless, a DiD approach can be specified in non-linear models (Wooldridge, 2021). More precisely, for non-linear models using exponential function, the common trends assumption can be tested with the conditional mean function, to which a log-odds transformation is applied. For the covariates, ratios are reported, instead of differences for linear models.<sup>15</sup>

To control for other local characteristics, a set of local indexes, i.e., Herfindahl Index (HI) and Location Quotients (LQ), are calculated in order to characterize the spatial attributes surrounding the firms. First, the HI characterizes the relative diversification of the surroundings of the firm  $i$  (Dubé and Brunelle, 2014). It is defined as the sum of the squared proportion of each economic sector  $s$ , in a 500-meter radius around the firm  $i$  (Equation 4.4).

<sup>15</sup> For more information on non-linear regression and DiD specification, refer to Wooldridge (2021).

$$HI_i = \sum_{s=1}^{15} \left( \frac{nbFirm_s}{TotalFirm} \right)^2 \quad (4.4)$$

On the one hand, the HI is unique to each firm unless firms have the same location (coordinates). Higher values of HI imply a concentration of economic activities, while lower values correspond to a more diverse environment (Dubé and Brunelle, 2014; Santarelli, 2000).

On the other hand, the LQ studies the relative specialization of an economic sector around the firm  $i$ . It is defined as the proportion of firms operating in a sector  $s$ , in a 500-meter radius around the firm  $i$ , divided by the proportion of firms operating in the sector  $s$ , for the entire study area (Equation 4.5).

$$LQ_{i,s} = \frac{\frac{Firm_{i,s}}{Firm_i}}{\frac{Firm_s}{Firm}} \quad (4.5)$$

The LQ calculates the intensity of an activity sector  $s$ , in a 500-meter radius around a firm  $i$ , meaning that each LQ is specific to a sector  $s$  and a firm  $i$  (Dubé and Brunelle, 2014), unless the firms have the same coordinates. If the LQ is greater than 1, the surrounding of the firm  $i$  is specialized in the sector  $s$ . The specialization of an area may indicate the presence of agglomeration economies, which may arise from the concentration of similar economic activities. The LQ is calculated for each sector.

All spatial indicators are calculated for the opening year of the firm  $i$ , and only include active firms for that specific year. This spatio-temporal specification of the local indicators avoids a large part of the endogeneity relation, as the firms that emerge are usually not big enough to have a global influence on the local structure. Moreover, the local indexes are used at time  $t$  in order to evaluate the impact of firms' surroundings and market composition on the *liability of newness*, that may be experienced by entering firms (Stinchcombe, 1965).

Finally, the Kaplan-Meier survival curves can be estimated using the (censored) data. The plots illustrate the probability that a specific firm will survive until time  $t$ , conditional to its survival until time  $t-1$ , with on the x-axis, the number of days since the firms entered the

market and on the y-axis their survival probability (Goel et al., 2010). From a temporal perspective, the survival function can be compared statistically by using the log-rank test and testing its null hypothesis, i.e., no significant difference between the survival functions (Marubini et al., 1995).<sup>16</sup> Based on the distance to the closest station and the economic sector (all economic sectors, retail, FIRE-KIBS, and local activities), the survival functions are plotted by time period and their significance is evaluated using a log-rank test.

## Results

The Cox PH model has been estimated for two distinct specifications to ensure no economic sector is driving a potential trend. The first one includes all the selected firms and assumed a homogenous (and single) average treatment effect for all economic sectors, by period and zone (model 1 – Table 4.3). While model 1 returns interesting results generalized to all economic sectors, no conclusion regarding firms' survival in specific economic sector can be drawn. The second model evaluates the impact for three (3) specific economic sectors and explores the possibility of having heterogenous impact within economic activities by period and zones (model 2 – Table 4.4). Using a sample limited to Retail, FIRE and KIBS, and Local Activities, model 2 presents sector specific results and ensures the obtained results are not unique to the aggregated economic sectors. It is also used to verify whether a certain sector is driving any potential trend observed in model 1.

For both models, the results for the first radius (0- to 250-meter area) must be interpreted with caution, because of the small number of observations. Moreover, the small number of observations in the direct vicinity of the stations can explain the large number of non-significant results obtained for both models (Table 4.3 – Column 1 and Table 4.4 – Column 1).

As underlined in the literature, the results imply a positive influence of size on firms' survival, when compared to firms with 0 to 1 employee. In both models (tables 4.3 and 4.4), the estimation suggests higher survival probability for all firms' size and varies depending on the number of employees and distance to the closest station. However, very large firms

---

<sup>16</sup> The Kaplan-Meier plots are available in *Annexe du Chapitre 4* (Figure A4.2.1 to Figure A4.2.4)

(more than 100 employees) appear to have a lower probability of survival, in comparison with smaller firms (0 to 1 employee).

The agglomeration of economic activities (location quotients i.e., specialization indexes), also appears as an important determinant on firms' survival. The spatial concentration of retail activities, referred to as LQ(Retail) in tables 4.3 and 4.4, increases firms' probability of survival from 500 meters to 1,000 meters from the closest station. In the 750- to 1,000-meter area, a high concentration of FIRE activities seems to increase the probability of survival, while spatial agglomeration of food and accommodation firms appears to reduce the survival probability (Table 4.3 – Column 4 and Table 4.4 – Column 4). The results suggest that the spatial agglomeration of KIBS, arts and entertainment, and personal services have no significant effects on firms' survival probability.

Concentrating on the effect of proximity to metro stations, both specifications (models 1 and 2) produce similar results. The impact appears to be concentrated between 250 meters and 1,250 meters from the stations. With no regard to temporal specifications, the results indicate that firms located around the new stations have lower survival probabilities when compared to firms operating around the speculative stations (tables 4.3 and 4.4). However, the introduction of the time periods reverses the trend (for both models), especially during construction and after the opening of the line.

#### *Model 1 – All Economic Sectors*

Focusing on the impact of the implementation of a new public transport service, the results vary depending on the period (Table 4.3). For the announcement period, the results are mostly non-significant, confirming, to some extent, that the market might not react at the announcement, since the project may never materialize. The only positive impact observed during this period is located within the 250- to 500-meter range (Table 4.3 – Column 2).

During the construction period, the HR ranges from 0.687 to 0.790, suggesting higher survival probabilities, ranging from 21.0% to 31.3%, for firms operating around the metro stations in Laval, as compared to the reference (blue line) stations. The impact increases from the 250- to 500-meter area, peaks in the 750- to 1,000-meter area and decreases until the 1,250- to 1,500-meter area (Table 4.3 – Column 2 to 6).



After the opening of the service, the results indicate that firms operating around the orange extension have a 40.9% higher survival probability in the 500- to 750-meter area (Table 4.3 - Column 3). For all the other studied areas (excluding the 0- to 250-meter area) the firms' probability of survival increases from 25.6% to 29.4%. Thus, the opening of the metro service appears to have a global significant impact on the survival of all selected economic activities.

**Table 4.3 - Cox regression results for all economic sectors (Model 1)**

Column		1	2	3	4	5	6
		0-250m	250-500m	500-750m	750-1,000m	1,000-1,250m	1,250-1,500m
		H. Ratio (Std Err.)	H. Ratio (Std Err.)	H. Ratio (Std Err.)	H. Ratio (Std Err.)	H. Ratio (Std Err.)	H. Ratio (Std Err.)
Reference (0 to 1 employee)							
Size (employees)	2 - 10	0.786 (0.064) ***	0.926 (0.048)	0.915 (0.043) *	0.891 (0.042) **	0.863 (0.041) ***	0.872 (0.044) ***
	11 - 25	0.770 (0.195)	0.590 (0.145) **	0.747 (0.184)	0.455 (0.119) ***	0.385 (0.101) ***	0.400 (0.105) ***
	26 - 49	0.455 (0.207) *	0.482 (0.217)	1.422 (0.416)	0.812 (0.260)	0.564 (0.190) *	0.351 (0.203) *
	50 - 99	0.378 (0.172) **	0.661 (0.386)	0.660 (0.271)	0.530 (0.376)	0.421 (0.244)	0.375 (0.217) *
	≥ 100	1.872 (0.639) *	2.886 (0.545) ***	1.237 (0.208)	1.937 (0.295) ***	1.848 (0.298) ***	1.904 (0.353) ***
	Reference (Wholesale)						
Sector	Retail	1.244 (0.182)	1.077 (0.089)	0.996 (0.078)	1.067 (0.085)	0.921 (0.075)	0.947 (0.077)
	FIREKIBS	0.980 (0.130)	0.730 (0.055) ***	0.704 (0.051) ***	0.794 (0.058) ***	0.737 (0.054) ***	0.750 (0.056) ***
	Loc. Act.	1.236 (0.175)	1.018 (0.080)	0.797 (0.061) ***	1.033 (0.080)	0.853 (0.067) **	0.901 (0.070)
	Local Indexes						
Local Indexes	LQ (Retail)	0.503 (0.497)	0.575 (0.424)	0.207 (0.140) **	0.192 (0.132) **	0.444 (0.294)	0.848 (0.642)
	LQ (FIRE)	0.845 (1.209)	0.672 (0.695)	0.365 (0.365)	0.129 (0.123) **	1.674 (1.641)	8.038 (9.503) *
	LQ (KIBS)	0.584 (0.413)	0.571 (0.406)	0.721 (0.343)	0.666 (0.336)	2.396 (1.138) *	1.425 (0.772)
	LQ (Food & Acc.)	4.243 (5.115)	2.491 (2.159)	5.022 (4.164) *	8.348 (6.562) ***	0.504 (0.403)	0.193 (0.184) *
	LQ (Arts & Ent.)	0.304 (0.193) *	0.721 (0.341)	0.649 (0.290)	0.682 (0.313)	0.855 (0.375)	0.806 (0.407)
	LQ (Pers. Serv.)	0.416 (0.282)	0.454 (0.227)	0.756 (0.345)	0.921 (0.444)	1.021 (0.459)	0.523 (0.269)

	HI	2186 (41520)	1283 (17813)	6277649 (8.29e7)	8.26e12 (1.06e14 )	**	0.011 (0.136)	0.000 (0.000)	
	Metro (Orange line)	1.193 (0.286)	1.374 (0.159)	*** 1.199 (0.111)	** 1.136 (0.108)		1.278 (0.124)	** 1.064 (0.127)	*
Metro	Orange x Ann.	1.024 (0.313)	0.727 (0.103)	** 1.151 (0.134)	0.926 (0.110)		0.818 (0.099)	* 1.052 (0.151)	
	Orange x Const.	1.121 (0.313)	0.743 (0.104)	** 0.721 (0.085)	*** 0.687 (0.081)	***	0.790 (0.094)	** 0.763 (0.107)	*
	Orange x Serv.	0.884 (0.256)	0.706 (0.096)	*** 0.591 (0.069)	*** 0.725 (0.083)	***	0.722 (0.087)	*** 0.744 (0.109)	**

Note: The Cox-Snell's residuals were estimated for both models to ensure that the data used were well suited for the Cox model. The plots are available in *Annexe du Chapitre 4* (Figure A4.3 and Figure A4.4). The Schoenfeld's residuals were also computed for all explanatory variables, using the "estat phtest" command in Stata. Based on those tests, the data are well suited and respect the PH assumption.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

**Table 4.4 - Cox regression results for all specific economic sectors (Model 2) part 1**

Column		1	2	3
		0 - 250 m.	250 - 500 m.	500 - 750 m.
		Haz. Ratio (Std Err.)	Haz. Ratio (Std Err.)	Haz. Ratio (Std Err.)
		Reference (0 to 1 employee)		
Size (employees)	2 - 10	0.781 (0.064) ***	0.928 (0.048)	0.914 (0.043) *
	11 - 25	0.773 (0.196)	0.596 (0.147) **	0.751 (0.185)
	26 - 49	0.452 (0.205) *	0.473 (0.213) *	1.406 (0.412)
	50 - 99	0.321 (0.149) **	0.693 (0.404)	0.662 (0.273)
	More than 100	1.833 (0.626) *	2.959 (0.556) ***	1.252 (0.211)
		Reference (Wholesale)		
Sector	Retail	1.314 (0.198) *	1.146 (0.104)	1.062 (0.089)
	FIRE & KIBS	0.940 (0.129)	0.765 (0.064) ***	0.726 (0.056) ***
	Act Loc	1.230 (0.181)	1.090 (0.095)	0.819 (0.068) **
Local Indexes	LQ(Retail)	0.553 (0.547)	0.549 (0.406)	0.205 (0.139) **
	LQ(FIRE)	1.040 (1.493)	0.689 (0.715)	0.398 (0.398)
	LQ(KIBS)	0.571 (0.406)	0.722 (0.373)	0.701 (0.334)
	LQ(Food & Acc.)	3.706 (4.473)	2.455 (2.130)	4.564 (3.785) *
	LQ(Arts & Ent.)	0.305 (0.194) *	0.684 (0.324)	0.640 (0.287)
	LQ(Pers. Serv.)	0.400 (0.271)	0.492 (0.247)	0.806 (0.368)
	HI	164.2 (3122)	1914 (26652)	5287964 (6.99e7)
Metro (Orange line)		1.154 (0.238)	1.208 (0.114) **	1.067 (0.088)
Retail	Orange x Ann.	0.578 (0.241)	0.759 (0.154)	1.170 (0.197)
	Orange x Const.	1.192 (0.598)	0.830 (0.197)	0.861 (0.156)
	Orange x Serv.	0.587 (0.257)	0.850 (0.167)	0.520 (0.099) ***
FIRE & KIBS	Orange x Ann.	2.072 (0.832) *	0.972 (0.156)	1.428 (0.215) **
	Orange x Const.	1.291 (0.373)	0.868 (0.135)	0.729 (0.108) **
	Orange x Serv.	1.196 (0.383)	0.769 (0.116) *	0.757 (0.106) **
Local Act.	Orange x Ann.	1.085 (0.421)	0.750 (0.148)	1.409 (0.225) **
	Orange x Const.	0.987 (0.378)	0.827 (0.161)	0.935 (0.175)
	Orange x Serv.	0.984 (0.341)	0.756 (0.127) *	0.634 (0.108) ***

**Table 4.4 - Cox regression results for all specific economic sectors (Model 2) part 2**

Column		4		5		6	
		750 - 1,000 m.		1,000 - 1,250 m.		1,250 - 1,500 m.	
		Haz. Ratio (Std Err.)		Haz. Ratio (Std Err.)		Haz. Ratio (Std Err.)	
Reference (0 to 1 employee)							
Size (employees)	2 - 10	0.887 (0.042)	**	0.857 (0.041)	***	0.876 (0.044)	***
	11 - 25	0.462 (0.121)	***	0.397 (0.104)	***	0.406 (0.107)	***
	26 - 49	0.835 (0.267)		0.557 (0.187)	*	0.342 (0.198)	*
	50 - 99	0.534 (0.379)		0.403 (0.234)		0.372 (0.216)	*
	More than 100	1.987 (0.303)	***	1.796 (0.290)	***	1.955 (0.363)	***
Reference (Wholesale)							
Sector	Retail	1.089 (0.096)		1.003 (0.087)		0.928 (0.084)	
	FIRE & KIBS	0.826 (0.066)	**	0.749 (0.059)	***	0.744 (0.060)	***
	Act Loc	1.090 (0.092)		0.912 (0.076)		0.883 (0.075)	
Local Indexes	LQ(Retail)	0.183 (0.127)	**	0.410 (0.271)		0.751 (0.571)	
	LQ(FIRE)	0.156 (0.150)	*	1.611 (1.583)		8.935 (10.60)	*
	LQ(KIBS)	0.636 (0.321)		2.272 (1.080)	*	1.326 (0.717)	
	LQ(Food & Acc.)	7.478 (5.896)	**	0.531 (0.426)		0.188 (0.180)	*
	LQ(Arts & Ent.)	0.654 (0.301)		0.846 (0.373)		0.736 (0.372)	
	LQ(Pers. Serv.)	1.003 (0.481)		1.076 (0.480)		0.586 (0.300)	
	HI	2.08e12 (2.68e13)	**	0.034 (0.443)		0.000 (0.000)	
Metro (Orange line)		0.991 (0.082)		1.221 (0.103)	**	0.873 (0.086)	
Retail	Orange x Ann.	1.170 (0.197)		0.822 (0.164)		1.498 (0.276)	**
	Orange x Const.	1.035 (0.183)		0.757 (0.159)		0.899 (0.186)	
	Orange x Serv.	0.784 (0.134)		0.520 (0.103)	***	0.891 (0.219)	
FIRE & KIBS	Orange x Ann.	1.044 (0.146)		0.922 (0.120)		1.279 (0.203)	
	Orange x Const.	0.712 (0.101)	**	0.904 (0.117)		0.938 (0.146)	
	Orange x Serv.	0.919 (0.113)		0.854 (0.115)		0.943 (0.148)	
Local Act.	Orange x Ann.	1.165 (0.197)		0.562 (0.128)	**	1.255 (0.267)	
	Orange x Const.	0.754 (0.121)	*	0.789 (0.140)		1.074 (0.189)	
	Orange x Serv.	0.704 (0.115)	**	0.759 (0.122)	*	0.912 (0.173)	

Note: The Cox-Snell's residuals were estimated for both models to ensure that the data used were well suited for the Cox model. The plots are available in *Annexe du Chapitre 4* (Figure A4.3 and Figure A4.4). The Schoenfeld's residuals were also computed for all explanatory variables, using the "estat phtest" command in Stata. Based on those tests, the data are well suited and respect the PH assumption.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

### *Model 2 – Specific Economic Sectors*

In the second model, the results suggest that early response to the new stations appears to induce few negative impacts on survival probability (Table 4.4). For the announcement period, the estimated HR in the high commercial density areas in Laval (500- to 750-meters) are 1.428 for FIRE and KIBS firms and 1.409 for local activities, suggesting those firms have a 42.8% and 40.9% lower probability of survival, respectively (Table 4.4 – Column 3). Once again, this partly confirms that anticipation regarding major project in Montréal’s regions might not start at the announcement.

Conversely, high residential density areas in Laval (1,000- to 1,250-meters) appear to have a positive impact on local firms, with a 43.8% higher probability of survival for those located in this area during the announcement period (Table 4.4 – Column 5). Here, again, survival appears to be more related to current clientele, and not much on anticipation. For retail activities around the orange extension, non-significant effect is observed for all zones, except for the 1,250- to 1,500-meter area in which retail firms have a 49.8% lower probability of survival, as compared to retail firms around the blue line (fictive) stations (Table 4.4 – Column 6).

During the construction period, being located around new metro stations of the orange line has no significant impact on the survival of retail firms. For this specific period, FIRE and KIBS firms located around the metro stations have 27.1% and 28.8% higher probability of survival, respectively, within the 500- to 750-meter area (Table 4.4 – Column 3) and the 750- to 1,000-meter area (Table 4.4 – Column 4). For their part, local activities operating in the 750- to 1,000-meter area around the new stations have a 24.6% higher probability of survival (Table 4.4 – Column 4).

After the transit service opens, a positive impact on retail activities is observed, with a 48.0% higher probability of survival for firms located in both the 500- to 750-meter and 1,000- to 1,250-meter areas around the orange extension (Table 4.4 – columns 3 and 5). Those two areas have been identified as important locations of commercial and residential real estate development (Champagne et al., 2022), which may influence local dynamics (Dubé et al., 2017), and therefore appear to positively impact firms’ survival rates.

For FIRE and KIBS activities, the results imply that firms have, respectively, a 23.1% and 24.3% higher probability of survival if they are located in the 250- to 500-meter and 500- to 750-meter areas, during the last time period (Table 4.4 – columns 2 and 3). The impact is concentrated within a short distance to the stations, meaning that commercial areas seem beneficial for FIRE and KIBS firms' survival. No significant impacts are observed over 750 meters.

The effects of the stations on local activities' survival are noted between 250 and 1,250 meters, after the opening of the transit service (Table 4.4 – Column 2 to 5). The effects peak in the 500- to 750-meter area and decrease until the 1,000- to 1,250-meter area. Results indicate that commercially dominant areas (500- to 750-meters) have a positive impact on the survival of local activities, with a 36.6% higher survival probability (Table 4.4 – Column 3). No significant effects are noted in the furthest area.

Finally, robustness tests have been conducted on restricted sample size. As the last period, i.e., the period when the service is in operation, extends over a 10-year period and some events could affect the survival of the treatment and control groups in different ways, the results were estimated using a 5-year and 7-year operation period. The models are adjusted to take into account the censored data according to the temporal specification. Overall, the obtained results for both models are robust, except for the FIRE and KIBS activities in the 500- to 750-meter area after the opening of the stations. Therefore, results regarding FIRE and KIBS activities must be interpreted with caution. The results of the robustness tests are available in *Annexe du Chapitre 4* (Figure A4.5 and Figure A4.6).

## **Conclusion**

The goal of this article is to study the causal impact of public transit on firms' survival, in order to fill a gap in the literature. The main assumption is that public transit increases the survival probability for the firms located close to the new stations, especially for service firms, since public transit facilitates human contact and increases the opportunity to meet potential customers. To test this assumption, a Cox Proportional Hazards model is developed using the simultaneous announcement of two extensions to define the treatment (new stations) and control groups (speculative stations), based on a maximal radius of 1,500

meters. The quasi-natural framework developed opens the door to a causal interpretation of the impact of a new metro station on firms' survival.

The results underline the mostly non-significant impact after the project was announced by the authorities. For all selected economic sectors, the results suggest that proximity to metro stations introduces a stronger survival probability for firms located within 250 meters to 1,250 meters from the closest station. For the service industry, the impact mainly materialized in the commercial area (500- to 750-meters), after the service comes into operation. More precisely, after the transit service opens, retail firms within 500- to 750-meters and 1,000- to 1,250-meters from the nearest station have greater survival probabilities than similar firms in the control group. The survival probability is also higher for local activities within 250 meters to 1,250 meters of the new stations during the last period. For the same time period, FIRE and KIBS activities appear to have higher survival probabilities within 750 meters from the nearest stations, however those results are not robust to shorter time periods and must be interpreted with caution.

From a public policy perspective, the results imply that the development of a public transit service offers more than just a reduction in transportation costs for households and better accessibility for workers. The results underline the fact that public transit enhances firms' performance from a survival perspective. Public transport infrastructure creates powerful nodes where potential clients and buyers circulate on a daily basis. This attractiveness appears to influence the individual firms' performance by providing a stronger survival probability. Thus, by designing new lines and services, it may be possible to shape and influence how economic activities evolve over time. This has important implications for policy-makers who aim to influence the structure of a city, at least around the new stations. This article highlights the importance of public transit and presents inherent information for future decision-making related to planning and urban growth policies.

The analysis is faced with some limitations as it was impossible to distinguish between firms' relocation and creation. Therefore, no conclusions can be drawn based on whether the metro stations attract new or already operating firms that choose to relocate. Moreover, the analysis excludes the branch offices and focuses exclusively on the location of the firms' head office. While including only headquarters may decrease the number of observations, franchises are



usually more volatile and their opening/closure largely depends on the headquarters' decision, not their own. Therefore, the results obtained here are valid for the location of the headquarters, and not for all individual establishments, branches, and franchises.

Ultimately, extensive research is still needed on the impact of public transit on the performance (survival) of firms. The approach presented in this article could be a starting point for future survival analysis on urban transit systems. For future research, introducing a nuance between firms' relocation and creation would be an interesting addition to a relatively sparse literature on the impact of public transit on firms' survival.

# **CHAPITRE 5 : Revitalizing Commute: The Impact of Bus-to-Tramway Shift on Small Businesses' Survival**

## **Résumé**

Cet article évalue l'impact de la substitution d'un service de transport en commun sur la probabilité de survie des petites entreprises. L'approche empirique repose sur le cas de la ville de Dijon (France), qui a partiellement remplacé son réseau d'autobus à haut niveau de service (*Bus Rapid Transit - BRT*) par deux lignes de tramway. Afin d'identifier un effet causal, un modèle de survie de Cox est mobilisé, et un contrefactuel est construit à l'aide des lignes d'autobus à haut niveau de service inchangées et antérieures à l'introduction du tramway. Les résultats suggèrent que la probabilité de survie des petites entreprises localisées autour des stations de tramway augmente après l'ouverture du service. L'impact observé est spatialement concentré dans les zones de 400 et 500 mètres autour des stations.

**Mots-clés :** Tramway; Autobus à haut niveau de service; Analyse de survie; Petites entreprises.

## **Abstract**

This paper aims to investigate the impact of a substitution in public transit services on small businesses' economic performance, from a survival perspective. The empirical approach relies on the case of Dijon's tramway (France), which has partially replaced its bus rapid transit (BRT) network. Using a Cox proportional hazards model, the causal impact of the tramway on small businesses' survival probability along the lines and stations is estimated using a quasi-experimental framework, based on the former (pre-tramway) bus rapid transit network. The results suggest that the survival probability of small businesses located around the tramway stations increases after the opening of the service. The impact is spatially concentrated within the 400m and 500m areas around the stations.

**Keywords:** Tramway; Bus Rapid Transit; Survival analysis; Small Businesses.

## **Introduction**

Transport infrastructures have undeniable impacts on the spatial distribution of population and economic activities within cities. From a theoretical perspective, the Alonso-Muth-Mills (AMM) model (Alonso, 1964; Muth, 1969; Mills, 1972) clearly underlines how transport infrastructure can affect location decision. By increasing accessibility and reducing travel time, transport infrastructures influence economic agent location decisions. They also generate externalities, positive and/or negative, which might have wider economic effect, affecting the performance of individual firms.

The development of cities after World War II has been mainly oriented towards the use and extension of highways. This revolution has highly shaped and influenced the development of suburban areas (Baum-Snow, 2007, 2010), and contributed to generate important negative externalities such as traffic congestion, deterioration of air quality, and production of fine particles, emission of greenhouse gases and noise along specific corridors. From a sustainable perspective, and with respect to climate change, reducing negative externalities caused by car use is important. To this end, the development of public transport appears to be a good way of achieving this goal, since public transit represents an interesting alternative to car usage. However, public transport systems might also introduce additional effects that could affect firms' performance.

The literature inherent to public transit infrastructures and their impacts is well documented. From a commercial perspective, research has dealt with firms' creation (Credit 2018), presence of agglomeration economies (Chatman et al., 2016; Alañón-Pardo and Arauzo-Carod, 2013; Takeda et al., 2008), new real estate development and constructions (Champagne et al., 2022), firms' location or relocation (Iseki and Jones, 2018; Beckerich et al., 2016; Champagne and Dubé, 2023), and uplift value (Debrezion et al., 2007; Seo et al., 2019; Zhang and Xu, 2017; Xu et al., 2016). Nevertheless, there is still work to be done regarding the impact of transport infrastructure on firms' performance, especially from a survival standpoint. Moreover, impacts regarding firms' longevity are complementary to the results already presented in the literature.

Literature revolving around the impact of transport infrastructure on firms' survival is relatively sparse. So far, it has mainly been studied for highways and heavy railways from a quantitative perspective (Manzato et al., 2010; 2011). Other studies based on qualitative approaches have also investigated how public transit affects firms' survival (Tornabene and Nilsson, 2021; Sukaryavichute et al., 2021).

This paper aims to evaluate, from a quantitative and causal perspective, the impact of the substitution of a bus rapid transit system (BRT) by a tramway, a specific case of light rail transit (LRT), on firms' longevity. To do so, a survival analysis with a difference-in-differences (DiD) framework based on small businesses operating in the service industry is undertaken. A Cox proportional hazards (PH) model focuses on the impact of public transit on firms' survival rates from a causal perspective by comparing the effect of the former BRT network to the new tramway lines. The main assumption is that public transit decreases the risk of failure (and inversely increases the survival rates) for the firms located around the new LRT stations, by increasing the attractiveness of the nodes created around the tramway stations, therefore having a positive influence on firms' performance. The paper fills a gap inherent to public transit and the wider economic effect of transport infrastructure from a firm's survival perspective.

The remainder of this paper is divided into five sections. Section 1 presents the literature related to survival analysis and firm performances. Section 2 describes the methodological framework for the Cox PH model. Section 3 introduces the case study, Dijon's tramway and BRT network, and presents the data mobilized for the analysis. Section 4 presents the results and discusses the most relevant. Section 5 summarizes the obtained results and discusses implications for future research.

### **Related literature**

Public transit infrastructure can stimulate economic recovery (Government of Canada, 2020), increase accessibility (Glaeser et al., 2008), influence commercial activities (Credit, 2019; Iseki and Jones, 2018), and prompt real estate development (Champagne et al., 2022). While the literature on transport systems and their impact on firms' performances has been studied

under various scopes, the survival perspective, in relation to transport infrastructure, has barely been addressed.

Survival analyses were historically used in the medical field, in order to determine the recurrence of an illness, efficiency of a treatment, and/or occurrence of death. Those specific analyses were adapted to social sciences in the early 1990s (Dunne et al., 1988; Audretsch and Mahmood, 1995; Box-Steffensmeier and Jones, 2004). Literature inherent to the topic indicates that firms' survival rates are largely affected by the firm's intrinsic characteristics (age, size, and economic sector), as well as its surroundings (spatial environment, and local amenities - Tveterås and Eide, 2000; Evans, 1987; Bates, 1990; Headd, 2003; Phillips and Kirchoff, 1989; Bates and Nucci, 1989).

First, literature suggests that mature firms have a lower risk of closure or, inversely, a higher survival rate, compared to newer firms (Caves, 1998). The concept of *liability of newness*, introduced by Stinchcombe (1965), suggests that surviving as a new firm in a market is harder than entering this same market. Indeed, mature firms have already made their mark and adapted to the market in which they evolve, compared to newer firms that must quickly adjust to the market in order to be competitive and avoid closure (Pérez et al., 2004). Moreover, by operating for a long time, mature firms are likely to have acquired a loyal consumer base. They have also reached a certain level of competitiveness, by acquiring knowledge of their own market (Pérez et al., 2004).

Complementary, smaller firms appear to have lower survival rates (Evans, 1987; Audretsch and Mahmood, 1995). Start-ups are usually smaller than well-established firms and are therefore more at risk of experiencing *liability of newness* (Stinchcombe, 1965). Locally-owned firms, which are generally smaller than large chain businesses, are more at risk of closing, as they have fewer opportunities (labor, capital) and can hardly benefit from scale economies, facing cost disadvantages (Pérez et al., 2004).

The Alonso-Muth-Mills (AMM) model illustrates the trade-off for economic agents, between land prices (location) and transportation costs (Alonso, 1964; Muth, 1969; Mills, 1972). By locating in the city center, a firm reduces its transportation costs, while being more accessible and visible for its customers. However, to secure that location, they must assume a higher

rent (or real estate price), as the demand for its location is higher than in the suburbs (remote area). Thus, the city center location is attractive for firms operating in the service industry, since their economic activities mainly rely on interpersonal relationships and face-to-face contacts (Mejia-Dorantes et al., 2012; Chatman et al., 2016). Simply, firms' behaviors are largely related by their economic sector, as it influences their needs, consumer base, type of worker, and location decision (Fujita, 1989; Champagne and Dubé, 2023).

The spatial concentration of the service industry in city centers affects the local environment in which firms operate. Spatial concentration of commercial activities often induces agglomeration economies.<sup>17</sup> Firms tend to spatially concentrate to reduce transportation cost, benefit from a shared labor pool, knowledge spillovers and the merging of their consumer bases (Chatman and Noland, 2011; Arzaghi and Henderson, 2008). Spatial concentration can be related to similar economic activities (Marshall, 1891) or diverse activities (Jacobs, 1969). Empirically, local indexes, such as location quotients and the Herfindahl index, are used as proxy to assess the presence of agglomeration economies and evaluate their role on firms' performance (Yi and Nam, 2019; Low and Brown, 2017; Renski, 2011; Pe'er et al., 2008; Pe'er and Keil, 2013).

Firms' location decision is also influenced by local amenities, including transport infrastructure. Focusing on public transit, stations tend to create powerful commercial nodes, around which an increase in pedestrian traffic may be observed (Ganning and Miller, 2020).<sup>18</sup> Public transport infrastructure mostly benefits the service industry, since their performance revolves around their capacity to be accessible and facilitate human contact (Mejia-Dorantes et al., 2012; Chatman et al., 2016). Nonetheless, the new attractiveness of the area may simultaneously lead to an increase in real estate prices (and rents), raising the firms' expenses and potentially affecting their performance (Nelson, 1999; Cordera et al., 2019). Thus, the

---

<sup>17</sup> Chatman and Noland (2011) defined agglomeration economies as the increase in production (and/or decrease in production cost) arising from the spatial concentration of (similar or diverse) economic activities in a given region.

<sup>18</sup> It has influenced the creation of a transit-oriented development (TOD) concept developed by Calthorpe (1993).

trade-off between positive and negative externalities around the stations reflect the net impact of public transit.

Based on the literature, the impact of public transit on firms' birth (or location decision) seems to be mostly place dependent (Chatman et al., 2016; Champagne and Dubé, 2023), with either positive (Credit, 2018; 2019; Iseki and Eom, 2019; Song et al., 2012) or non-significant results (Schuetz, 2015; Leitham et al., 2000; Castillo-Manzano and López-Valpuesta, 2009).

Qualitative research has been conducted on the impact of public transport infrastructure on firm survival. Based on surveys conducted on managers and owners of small local businesses, the results suggest a negative impact on firms' survival during the construction of the public transit infrastructure (Tornabene and Nilsson, 2021; Sukaryavichute et al., 2021). A reduction of survival rates for small local firms, during the construction period, may be explained by lower visibility and accessibility, as well as a loss of consumer and complementary sales (Tornabene and Nilsson, 2021; Sukaryavichute et al., 2021).

From a quantitative perspective, Manzanto et al. (2010; 2011) evaluate the impact of proximity to highways and railways on firms' survival by using a survival model. Their results suggest a positive impact of both infrastructures on service firms' survival, especially for retail, finance, real estate, business services, social security, computer and information technology, and research and development activities. Finally, Champagne et al. (2024) study the impact of metro stations on firms' survival in Montréal's suburbs (Canada). The authors identify a positive impact of the metro extension, mainly after the opening of the stations, for retail and local activities.

## **Methodology**

The most common survival model in the literature is the Cox Proportional Hazards (PH) model (Cox, 1972). This survival model estimates the survival rate of an observation at a given time  $t$ , conditional to its survival until the period prior, at  $t-1$  (Kleinbaum and Klein, 2005).



For survival analysis, the Cox model is preferred over other models that also estimate the probability of an event occurring (e.g., logit and probit models), because it controls for censored data (Kleinbaum and Klein, 2005; Lee and Wang, 2003). An observation is considered as censored if it has opened before the study begins (left-censoring), or if it remains active after the studied period (right-censoring). Survival models, such as the Cox PH model, do explicitly account for this type of data and ensure that the model correctly estimates the survival probability, returning unbiased results (Kiefer, 1988).

The Cox PH model estimates the hazard rate of an observation  $i$ , which represents its survival (or, inversely, death) rates. The hazard rates are calculated using inherent characteristics of the observations and depending on the sample and aim of the study, the selected characteristics (covariates) may vary (Kleinbaum and Klein, 2005). The hazard rate function is divided into two parts. The baseline hazard rate  $h_0(t)$ , and the set of independent variables,  $x_{ik}$  (Equation 5.1).

$$h(t, X_i) = h_0(t) e^{\sum_{k=1}^K x_{ik}\beta_k} \quad (5.1)$$

The baseline hazard rate is an undefined function and represents the hazard when all the explanatory variables are equal to 0 at a given time  $t$  (Kiefer, 1988). The observation's characteristics are introduced using an exponential form (Kleinbaum and Klein, 2005). The explanatory variables are not time dependent (time factor is only included in the baseline hazard rate). The same assumption holds for the coefficients, which are assumed to be independent of the time factor.

The first part of the function is non-parametric, while the second one is parametric. For this reason, the Cox PH model is referred to as a semi-parametric model, which represents an advantage of the model that addresses and overcomes issues related to unobserved heterogeneity (Kleinbaum and Klein, 2005). The Cox PH model is estimated without defining the baseline hazard rate, using a partial likelihood method (Lee and Wang, 2003).

To evaluate the impact of a specific treatment (or event) on survival rates, a difference-in-differences (DiD) specification is added to the parametric components of the model. Additional covariates aim to distinguish between the treatment group ( $D_i = 1$ ), which includes all the observations that are affected by a specific change, and the control group

( $D_i = 0$ ), which contains the observations that are not impacted by the change (Roth et al., 2023). The DiD specification also includes variables identifying specific time periods ( $P_r$ , where  $r = \{0, 1, 2, 3\}$ ), related to important dates regarding the treatment (reference, announcement, construction and operation periods). The interaction of these two new variables isolates the impact of the treatment by period ( $D_i * P_r$ ).

To be valid, the DiD specification should respect the parallel trend assumption, meaning the two groups must be comparable and evolve the same way before the change occurs. This assumption can be tested using a placebo test<sup>19</sup> on observations collected before any change occurs.

The Cox PH model with the DiD specification is used to evaluate the impact of a treatment on survival rates through a hazard ratio (Equation 5.2). This ratio estimates the impact of each covariate on relative survival rate, meaning the impact on the survival rate of an observation  $i$  is relative to the survival rate of its counterfactual  $j$ , with  $i \neq j$  (Kleinbaum and Klein, 2005; Lee and Wang, 2003).

$$HR = \frac{h(t, X_i)}{h(t, X_j)} = e^{\sum_{k=1}^K (X_{ik} - X_{jk})\beta_k} \quad (5.2)$$

The baseline hazard rates cancel out, meaning that the hazard ratio is independent of time and the exponential form of the second term ensures a positive hazard ratio. When the ratio is greater (lower) than 1, the survival rate is lower (higher) in comparison to its counterfactual.

The Cox PH model is estimated under the Proportional Hazard Assumption (PHA) stating that all characteristics contained in the list of independent variables must be proportional over time (Lee and Wang, 2003). The Schoenfeld residuals test can be used to ensure that variables respect the PHA (Schoenfeld, 1982). The residuals are retrieved and used to test the zero-slope assumption, i.e., the absence of relation between the covariate and the time, or that the

---

<sup>19</sup> The placebo test uses the time periods before the event of interest (treatment) takes place. With this subsample, a “fictive” event is implemented, generating a new variable of treatment. Since that (fictive) treatment was never implemented, the impact of the (fictive) treatment variable should not be statistically significant. Under the null hypothesis that both groups are comparable, the parallel trend assumption cannot be rejected if the difference between the impact of this (fictive) treatment is nonsignificant (Slutsky, 2017).

log hazard ratio function is constant over time (Cleves et al., 2010). The Schoenfeld residuals individually tests each explanatory variable against time, while a global test of the Schoenfeld residuals combines the effect of all variables included in the model and studies its relation in regard to time (Grambsch and Therneau 1994). Overall, residuals should be proportional over time, the opposite indicating that the PHA is violated.

## **Case study and data**

### *Dijon's public transit network*

Dijon is located in northeastern France, in the Bourgogne-Franche-Comté region. The (medium-sized) city counts about 160,000 inhabitants within its core, while the agglomeration has more than 250,000 inhabitants (INSEE, 2024). Dijon's public transit network was rebranded and restructured in 2004, with the inauguration of *Divia Mobilité*. This change has led to numerous modifications, in terms of quantity and quality of services. Since its introduction, *Divia Mobilité* has connected suburbs that were not previously accessible by public transit, thereby increasing its number of stations. It also implemented a bus rapid transit (BRT) network on main streets, also known as *Lianes* (*Lignes à niveau élevé de service*<sup>20</sup> – Figure 5.1 Before).

The high volume of passengers on the *Lianes* has rapidly led to the saturation of certain routes, such as *Lianes 1* and *Lianes 5*, which could account for, respectively, 33,000 passengers and 20,000 passengers per day (Dijon Métropole, 2009). Traffic congestion caused by the important number of buses (BRT and buses) on the main roads led the city of Dijon to think about alternative solutions for its public transit network.

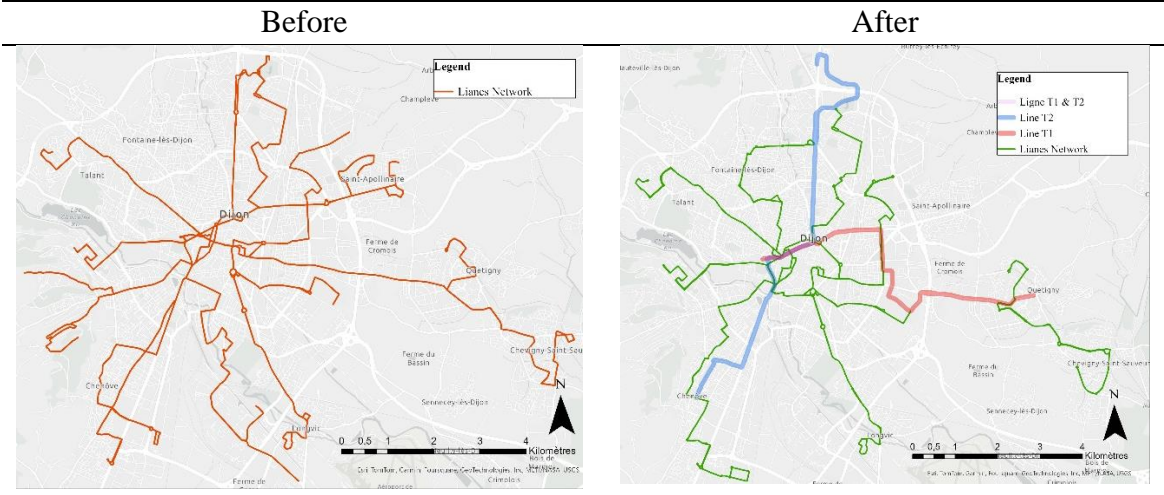
The tramway was officially announced in May 2008 as a sustainable solution to the congestion problem caused by the saturation of the BRT (and buses) network. The tramway project also aimed to facilitate access to education and employment centers, and limit urban sprawl and congestion, by offering a competitive substitute to the private car. By implementing such an infrastructure, the city wished to promote urban renewal in certain districts, while increasing opportunities for people located in more sensitive urban areas.

---

<sup>20</sup> Which can be translated by “*High service level lines.*”

The project included two lines. On the one hand, Line T1 connects the train station (west) to the eastern suburbs (Quétigny) via the university campus. It mainly replaces the routes of *Lianes 3*, *Lianes 4* and *Lianes 5*, in the city center, and *Lianes 1*, from the university campus to the eastern suburbs. On the other hand, Line T2 connects the northern and southern suburbs via Dijon’s city center. This second line partly replaces *Lianes 2* on the north of the city center and *Lianes 4* on the south of the city center (Figure 5.1 After).

**Figure 5.1 - Dijon’s BRT network before and after the tramway**



The construction of the two tramway lines began in October 2010. Faced with the imminent arrival of additional public transport services, *Divia Mobilité* restructured most of the existing network. The actual routes of *Lianes 3*, *Lianes 4* and *Lianes 6* were last changed in 2010. After the opening of the tramway stations, at the end of 2012 (September 2012 for Line 1, and December 2012 for Line 2), final changes to the BRT network were made and *Lianes 7* was added to the network. This same year, *Lianes 1* and *Lianes 2* were discarded, while part of the *Lianes 5* was replaced by the new tramway lines.

*Data*

The data used for the analysis was extracted from the *Sirene* database. It is accessible online and records all firms operating in France. The database contains information on the opening and, if applicable, closing date, the main economic activity, full address (used to geocode the data), and the number of employees (categorical variable). The study aims to evaluate the

impact on small businesses, therefore, only firms with 2 employees or less were kept for analysis.

The *Sirene* database also includes the *Siret* number, which is a unique number assigned to each firm, allowing an individual-level study. However, no information on the financial state of the firms is disclosed.

Individual survival time (in days) is calculated using the opening and closing date, if the firm closed within the period studied. All firms active for less than 30 days were excluded from the sample. Moreover, if the firm did not close before January 1, 2020 (end of the study period), it is considered as censored. A failure event indicator dummy is generated to address the right-censoring in the model. This dummy is equal to 1 if the firm closed during the period studied, or 0 otherwise. The dataset covers 14 years, from 2006 to 2019 (inclusively) and firms that opened before 2006 are discarded (no left-censoring).

The selected covariates are commonly used in the literature on firms' survival or on the impact of public transit on firms' location decision. Since this study focuses solely on firms with 2 employees or less, no specific covariate is added to the model in regard to a firm's size, since the sample is limited to small businesses.

The economic sector influences firms' longevity, as it may react differently to certain shocks or public policies. This study focuses on economic sectors that can easily operate as a small business (2 employees or less), and that are most likely to operate in an urban context and be affected by public transit. It includes firms active in Finance, Insurance and Real Estate (FIRE), Knowledge Intensive Business Services (KIBS) activities, other business services (excluding KIBS), and personal services.

Many sectors were discarded from the sample, such as Wholesale, Retail, Arts and Entertainment, and Food and Accommodation activities, since they are less likely to operate with 2 employees or less. Agriculture, manufacturing, construction, and logistics sectors were also excluded from the sample, since those firms are more likely to operate in peripheral areas, and therefore less likely to be affected by public transit. Finally, to avoid estimating false longevity, public administration, education, health, and social services activities were removed from the sample.

Besides the intrinsic impact of the economic sector on firms' survival, the spatial distribution of economic activities and market composition are also factors influencing firms' survival. Distance-based local indexes, such as location quotients and the Herfindahl Index, are calculated to define the firm's local environment, within 500 meters (Dubé and Brunelle, 2014). The indexes are measured at the firm's opening date, since the spatial distribution and local structure of the market are likely to impact the firm's probability of survival (*liability of newness* - Stinchcombe, 1965). The specifications used (spatial and temporal) to create the local indexes prevent the potential presence of endogeneity, since the individual weight of entering firms should not be enough to influence the local composition or structure of the market.

Those indices are built using a distance-based method (DBM) with the individual geographical coordinates of individual firms (Dubé and Brunelle, 2014). Throughout this study, the Euclidean distance is preferred. In an urban context, the Euclidean distance has the advantage of being time-unvarying and relying on a single layer of information. It is also mobilized to calculate distance to certain location.

The specialization of the market, or the agglomeration of similar activities, is measured using location quotients (LQ). Those local indicators calculate the relative proportion of a specific economic sector, in relation to the rest of the study area (Equation 5.3). For this index, the number of firms ( $F$ ) is calculated for a specific sector  $s$  around the firm  $i$  at its opening date  $t$  ( $F_{i,t,s}$ ), as well as for the whole sample ( $F_{t,s}$ ). The total number of firms located around the firm  $i$  ( $F_{i,t}$ ), and within the studied area ( $F_t$ ), for all economic activities, is also counted. The selected radius of influence is relatively small (500 meters), because the impact appears to be mainly local in the literature (Andersson and Larsson, 2016).

$$LQ_{i,t,s} = \frac{\frac{F_{i,t,s}}{F_{i,t}}}{\frac{F_{t,s}}{F_t}} \quad (5.3)$$

A location quotient greater (lower) than one indicates a more (less) specialized environment, in relation to the rest of the sample. Finally, location quotients are sector-specific, location-specific, and time-specific, meaning they are measured for each sector and each firm (based

on their coordinates), according to the composition of the environment at the firm's opening date.

The Herfindahl Index (HI) is used to study the diversification of the local market. As for the location quotients, the Herfindahl Index is measured at the firm's opening date  $t$ , within 500 meters. The squared proportion of economic activities, i.e., their respective market share, are summed to obtain the diversification index (Equation 5.4).

$$HI_{i,t} = \sum_{s=1}^S \left( \frac{F_{i,t,s}}{F_{i,t}} \right)^2 \quad (5.4)$$

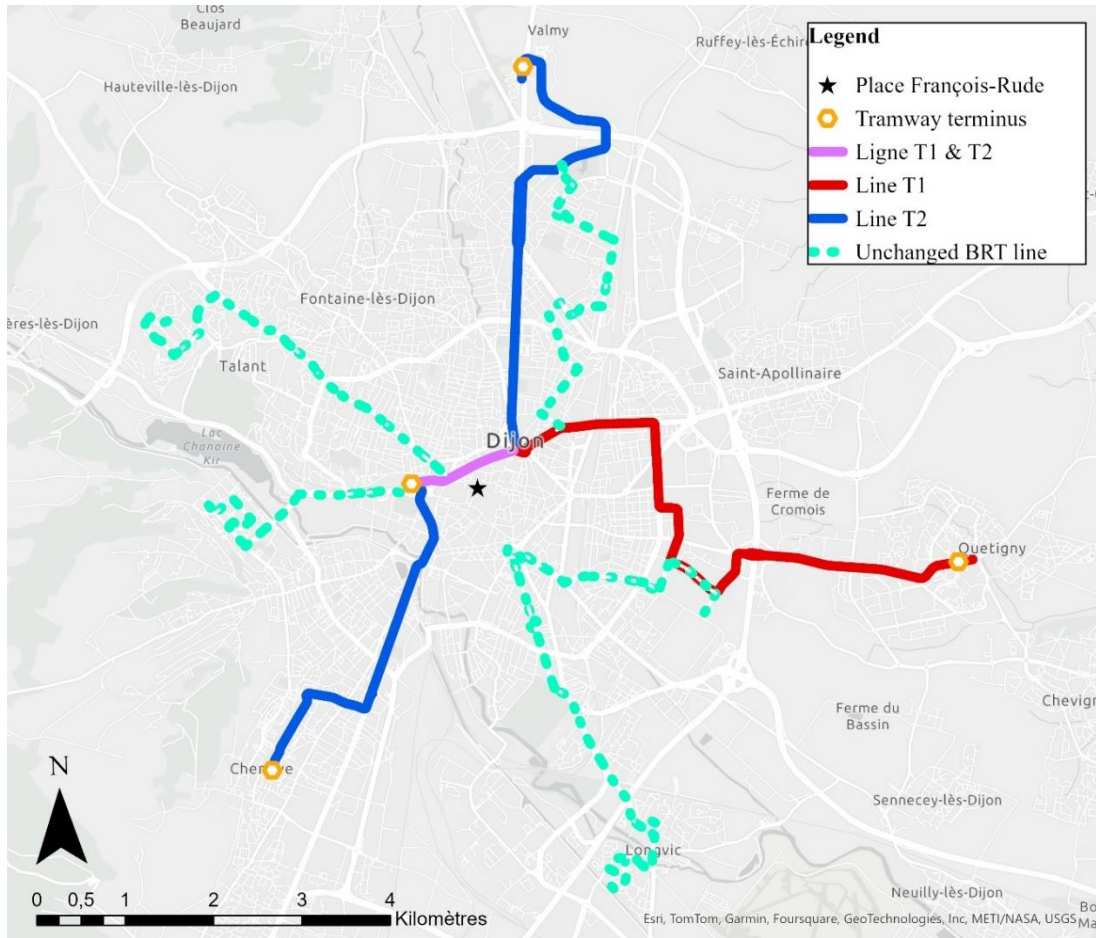
The agglomeration of diverse economic activities implies a lower value of the index, while a high value indicates that the environment is less diversified.

The distance to the city center (*Place François-Rude*) and the distance to the closest tramway's terminus station are included in the model to control for the proximity to potential commercial nodes (Figure 5.2).

To isolate the impact of the tramway on small businesses' survival, distance to the closest station is mobilized. The BRT network, established 4 years before the announcement of the tramway (in 2004), is used to create the control group, while firms located around the tramway stations are considered as treated. The control group is based on the BRT routes that have remained unchanged since their introduction in 2004. While the name/number assigned may have been altered, the *Lianes* stations and routes included in the control group remain globally unaffected by the introduction of the tramway (Figure 5.2).

The studied area is concentrated within 600 meters of stations along both services. The selected radius was based on the literature. On the one hand, the impact of public transit on commercial activities appears to be spatially concentrated around the stations, generally within 400 meters (1/4 mile) and rarely exceeding 800 meters (1/2 mile) (Chatman et al., 2016; Credit, 2018; Arzaghi and Henderson, 2008; Champagne et al., 2022). On the other hand, Dubé et al. (2018) studied Dijon's tramway and its impact on the real estate market, and their results suggest that the impact is limited to 600 meters around the stations.

**Figure 5.2 - Tramway lines and unchanged BRT lines**



To evaluate the local impact of the substitution and isolate specific effect in relation to the distance to the closest station, the studied area is divided using sets of 100-meter buffers (Table 5.1). The treatment and control groups are defined to investigate the impact of the distance to public transit on firms' survival, using non-mutually exclusive buffers, meaning that the area covers from the station to the critical distance. The treatment dummy variable is equal to 1 if the firm is located within the critical distance of the tramway station, and 0 if it is located within the critical distance of a *Lianes* station.

Furthermore, mutually served areas, i.e., firms located within 200 meters of both a tramway and a *Lianes* station, are removed from the sample. All observations within 600 meters of *Place François-Rude* (city center) were also discarded, to avoid capturing a city-center effect, which has undergone several phases of pedestrianization since 2006 (Figure 5.3).



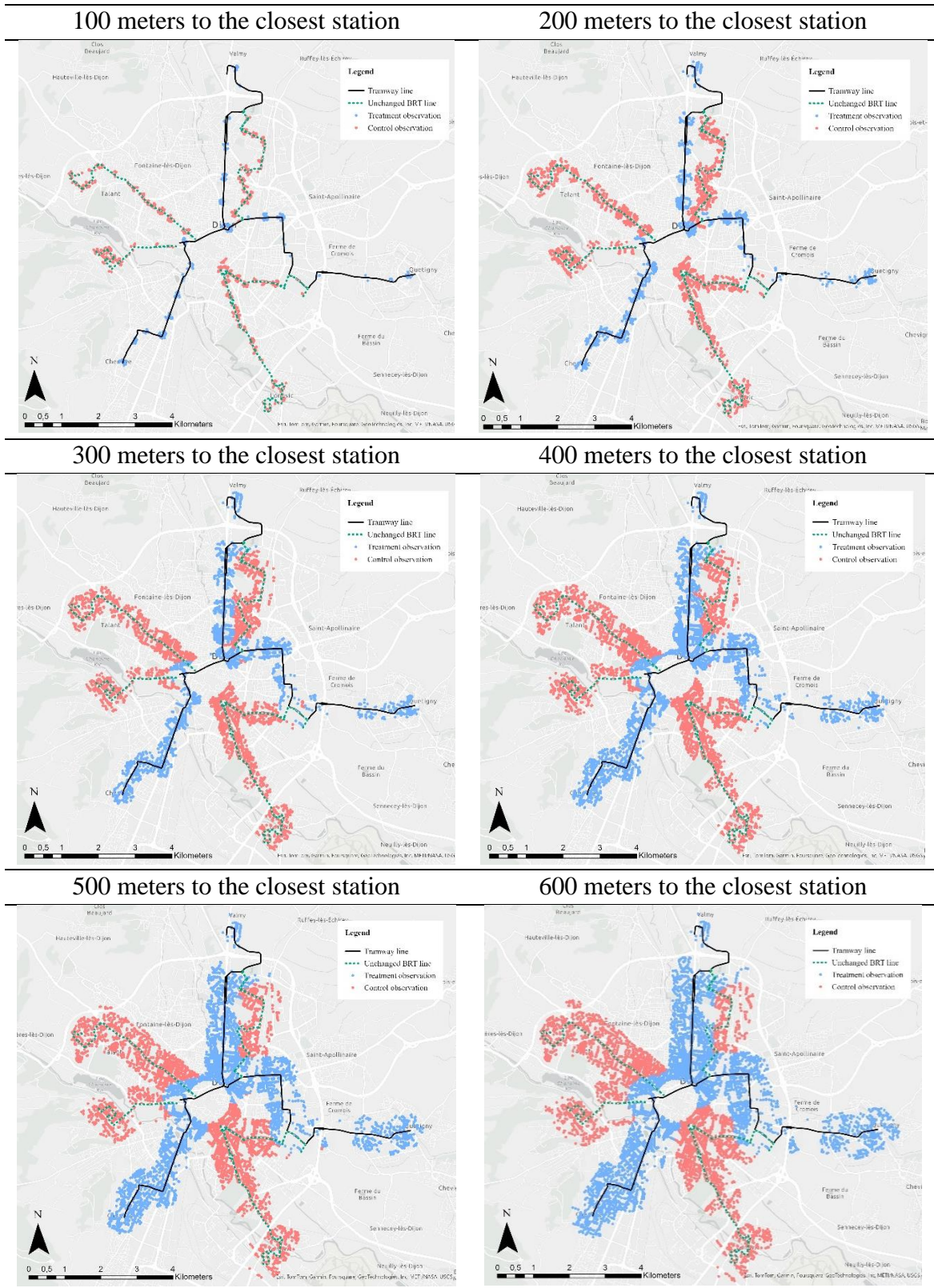
**Table 5.1 - Number of closed and censored firms in the sample**

Critical distance	Closed	Censored	Total
100m	913	684	1,597
200m	2,506	1,965	4,471
300m	3,833	3,104	6,937
400m	4,851	4,079	8,930
500m	5,563	4,743	10,306
600m	6,059	5,143	11,202

From a temporal perspective, period dummies are created. The time period is divided into 4 distinct sub-periods to identify the reference (pre-announcement), announcement, construction, and operation periods. The dummy for a given period is equal to 1 if the firm is active in the studied period, and zero otherwise.

Finally, the interaction between the treatment and period dummies enables the model to estimate the causal impact of the tramway by period. The interaction terms identify the firms located around the tramway and the temporal period in which they opened and operated, allowing for the evaluation of potential anticipation and speculation effects.

**Figure 5.3 - Spatial distribution and group division based on critical distance**



## Results

This section discusses the main results and presents alternative Cox PH specifications that were estimated to test the robustness of the model. The main results are presented in Table 5.2, while robustness tests' results are presented in Table 5.3, Table 5.4 and Table 5.5.

While the impact of being located around the tramway stations appears to have a negative effect, this impact becomes either nonsignificant or positive when it is studied by period. For the first two periods (announcement and construction periods), the impact of the proximity to tramway stations on small businesses' survival is not significant, when compared to firms located around the *Lianes* stations. Before the tramway goes into service, small businesses do not benefit from the positive impact that can emerge from their proximity to its stations. Only after the opening can firms really take advantage of the higher visibility and increased consumer flow (among others), which may reflect on their performances. Thus, the substitution of a BRT service with a tramway does not seem to induce any speculation or anticipation effects, i.e., during the announcement and construction periods, for all critical distances.

After the opening of the tramway, the 200m and 300m areas return non-significant impacts. The hazard ratios (estimated coefficients) are a relative probability of survival, meaning a non-significant impact indicates that there is no difference between the two groups (BRT vs. Tramway). Since BRT stations tend to have a smaller radius of influence, in comparison to tramway stations, the spatial influence observed in the results is not surprising. Also, negative externalities, such as an increase in rents following the arrival of the tramway, can arise and cancel out the positive externalities generally associated with public transit. The literature does point out that the impact on commercial real estate value (or rent) is highly concentrated around public transit stations, i.e., within 400m (Debrezion et al., 2007; Seo et al., 2019; Zhang and Xu, 2017; Xu et al., 2016).

Furthermore, the results suggest that close proximity, i.e., being located within 100m from a tramway station compared to a *Lianes* station, increases small businesses' survival rates by 38.0% (Table 5.2 – column 1). Positive and significant impacts of the tramway on small businesses' survival are also observed in the 400m, 500m, and 600m areas. The results imply

that survival probabilities of firms located within those critical distances increases by, respectively, 18.0%, 22.8%, and 12.7%, compared to a small business located by the BRT service. The result suggests a U-shape impact of the tramway, with the impact being significantly concentrated closely, i.e., within 100m, being non-significant within 100m to 300m, and then positive and statistically significant until 600m. Over 500m, the impact decreases with the distance to the station, as the hazard ratio increases to 0.833.

Moreover, the environment in which firms operate seems to influence small businesses' survival rates. The results suggest that the spatial concentration of FIRE activities increases firms' survival rates by 16.8% to 19.6% for all critical distances, except for the 100m area. The results also indicate that spatial concentration of wholesale activities does have a small positive and significant influence. This effect seems concentrated farther away from the stations, in the 500m and 600m areas, and slightly increases small businesses' probability of survival. This small positive effect might be explained by the large volume of consumers attracted by wholesale activities. Small businesses can therefore benefit from their consumer base, when located close by.

To ensure that both groups are comparable, a placebo test was conducted. Using June 1, 2009, as a "fictive" treatment date, the model was estimated (*Annexe du Chapitre 5 – Table A5.1*). The sample used is composed of firms that opened solely before the opening of the tramway (September 2012). *Placebo Period* is a dummy variable indicating whether the firm opened before (*Placebo Period* = 0) or after (*Placebo Period* = 1) the "fictive" treatment date. Results are non-significant for both the treatment variable (*Tramway*) and its impact before and after the placebo date (*Tramway x Placebo Period*), suggesting the parallel trend assumption is respected by the selected groups, for all critical distances.

Finally, a variety of tests were conducted to verify the validity of the model and its goodness-of-fit. First, Schoenfeld residuals tests are conducted to test the validity of the PHA (Table 5.2 – column PHA). For all distances, the results suggest that the PHA is respected, globally and individually, at a 5% confidence level, except for 3 covariates (1 in the 100m area, and 2 in the 300m area).

The Cox-Snell residuals test evaluates the goodness-of-fit of the model, according to its data and covariates (Cox and Snell, 1968; Cleves et al., 2010). The survival function is estimated based on the Cox model and its covariates. A negative logarithmic transformation is applied to the individual survival function and the difference between those two functions (estimated and observed) returns the residuals (Cox and Snell, 1968). The y-axis of a Cox-Snell plot represents the cumulative hazard, while the residuals are on the x-axis. According to this test, the cumulative hazard function (conditional on the covariates) of a well-fitted model should follow a straight 45° line, which represents the theoretical distribution of a perfectly fitted model. To evaluate the overall fit of the model estimated, the Cox-Snell residuals were plotted (*Annexe du Chapitre 5* – Figure A5.2). The Cox-Snell residuals tests suggest that the model used is well-fitted (Cleves et al., 2010).

**Table 5.2 - Cox model results (part 1)**

Critical distance	100 meters		200 meters		300 meters				
	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA			
Tramway	1.337 (0.240)	0.480	1.102 (0.125)	0.419	1.165 (0.100)	*	0.313		
Period 1 (Ann.)	0.990 (0.153)	0.860	1.207 (0.111)	**	0.098	1.196 (0.096)	**	0.740	
Period 2 (Const.)	1.251 (0.207)	0.628	1.272 (0.131)	**	0.228	1.273 (0.114)	***	0.664	
Period 3 (Oper.)	1.286 (0.189)	*	0.838	1.277 (0.116)	***	0.168	1.342 (0.105)	***	0.898
Tramway x P. 1	0.808 (0.198)	0.292	0.866 (0.128)		0.669	0.839 (0.096)		0.404	
Tramway x P. 2	0.785 (0.217)	0.554	0.774 (0.125)		0.984	0.849 (0.106)		0.480	
Tramway x P. 3	0.620 (0.136)	**	0.705	0.835 (0.113)	0.551	0.849 (0.088)		0.564	
Distance to the city center	1.042 (0.048)	0.716	1.017 (0.024)		0.129	1.034 (0.020)	*	0.197	
Distance to the closest terminus	0.950 (0.052)	0.557	1.002 (0.031)		0.548	0.987 (0.024)		0.430	
Herfindahl index	0.328 (0.416)	0.509	0.874 (0.574)		0.193	0.570 (0.343)		0.045	
LQ FIRE	0.804 (0.127)	0.194	0.832 (0.071)	**	0.344	0.821 (0.059)	***	0.145	
LQ KIBS	0.921 (0.094)	0.391	0.926 (0.056)		0.133	0.958 (0.048)		0.211	
LQ Business services	0.944 (0.082)	0.911	0.982 (0.056)		0.069	0.983 (0.046)		0.041	
LQ Personal services	0.946 (0.070)	0.015	0.971 (0.044)		0.058	0.985 (0.035)		0.545	
LQ Food and Acc.	0.943 (0.090)	0.848	0.956 (0.049)		0.808	0.930 (0.036)	*	0.712	
LQ Wholesale	0.841 (0.085)	*	0.940	0.932 (0.046)	0.383	0.948 (0.037)		0.862	
Global PHA		0.477			0.130			0.210	

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

**Table 5.2 - Cox model results (part 2)**

Critical distance	400 meters			500 meters			600 meters		
	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA
Tramway	1.142 (0.088)	*	0.266	1.194 (0.088)	**	0.776	1.200 (0.087)	**	0.686
Period 1 (Ann.)	1.234 (0.094)	***	0.623	1.285 (0.097)	***	0.451	1.317 (0.103)	***	0.134
Period 2 (Const.)	1.319 (0.111)	***	0.805	1.449 (0.119)	***	0.740	1.444 (0.122)	***	0.498
Period 3 (Oper.)	1.431 (0.106)	***	0.564	1.528 (0.112)	***	0.761	1.506 (0.114)	***	0.876
Tramway x P. 1	0.907 (0.092)		0.487	0.907 (0.087)		0.501	0.899 (0.085)		0.796
Tramway x P. 2	0.884 (0.097)		0.283	0.849 (0.088)		0.730	0.913 (0.093)		0.854
Tramway x P. 3	0.820 (0.076)	**	0.469	0.772 (0.068)	***	0.725	0.833 (0.072)	**	0.873
Distance to the city center	1.019 (0.018)		0.572	1.023 (0.016)		0.881	1.022 (0.015)		0.879
Distance to the closest terminus	0.989 (0.022)		0.155	0.985 (0.020)		0.482	0.983 (0.020)		0.688
Herfindahl index	0.892 (0.461)		0.088	0.924 (0.438)		0.270	0.889 (0.404)		0.534
LQ FIRE	0.823 (0.054)	***	0.257	0.812 (0.049)	***	0.988	0.804 (0.046)	***	0.595
LQ KIBS	0.951 (0.042)		0.077	0.995 (0.040)		0.564	0.993 (0.038)		0.211
LQ Business services	1.028 (0.040)		0.221	1.019 (0.037)		0.227	1.032 (0.034)		0.067
LQ Personal services	1.019 (0.032)		0.774	1.035 (0.030)		0.802	1.023 (0.026)		0.403
LQ Food and Acc.	0.970 (0.032)		0.817	0.987 (0.030)		0.998	0.983 (0.029)		0.814
LQ Wholesale	0.939 (0.031)	*	0.313	0.940 (0.029)	**	0.326	0.936 (0.027)	**	0.253
Global PHA			0.075			0.338			0.075

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

Overall, the results of the test suggest that the model estimated does respect the necessary assumptions and is well-defined and well-fitted, based on the data and covariates used. Those tests highlight the quality of the model in the context of this study.

#### *Robustness tests*

To check the robustness of the results, the model is evaluated under different spatial and temporal specifications. First, to ensure that the impact is really limited to the critical distance selected, the Cox model is estimated a second time with larger distances (500m to 700m).

For this application, an additional specification is applied. While a tramway might be an upgrade in comparison to a BRT service, proximity does play a role in the usage and ridership of those transit systems. To avoid wrongfully attributing an effect to the tramway, firms located within 300m from a *Lianes* station, while being located at more than 300m from a tramway station, are considered as control, even if they are located within the treatment area.<sup>21</sup> Indeed, close proximity to a *Lianes* station, in contrast to a tramway station located further away, is more likely to guide the use of public transport for local residents, and influence firms' performance.

The results suggest that the impact of the tramway on small businesses' survival is limited within 500m of the stations (Table 5.3). The impact is also similar, both in value and significance for this critical distance. Moreover, the positive and significant impact identified for the 600m area is not robust and must be interpreted with caution. For the greater distance (700m area), no significant impacts are observed, suggesting the impact of the tramway on small businesses' survival rates is located within 500m of its stations.

---

<sup>21</sup> Literature suggest that 0.3 km is approximately the distance transit-users are prepared to walk in order to reach a station (Untermann, 1984; Stringham, 1982; Cervero, 1994; Nelson, 1999).



**Table 5.3 - Cox model results using distance-based specification for larger critical distances**

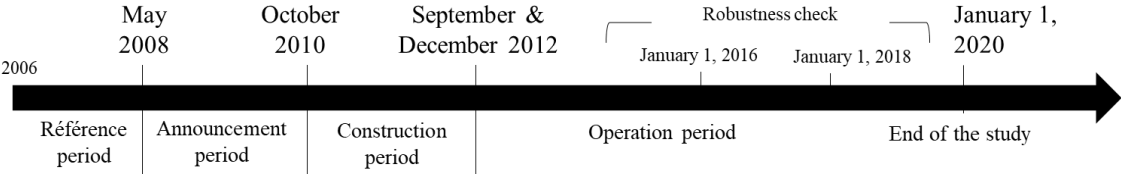
Critical distance	500 meters			600 meters			700 meters		
	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA
Tramway	1.137 (0.082)	*	0.662	1.107 (0.077)		0.783	1.037 (0.070)		0.556
Period 1 (Announcement)	1.275 (0.091)	***	0.374	1.280 (0.090)	***	0.165	1.226 (0.084)	***	0.073
Period 2 (Construction)	1.447 (0.113)	***	0.516	1.406 (0.109)	***	0.430	1.353 (0.101)	***	0.167
Period 3 (Operation)	1.454 (0.101)	***	0.730	1.432 (0.098)	***	0.864	1.399 (0.093)	***	0.544
Tramway x Period 1	0.915 (0.087)		0.509	0.934 (0.085)		0.904	1.008 (0.089)		0.717
Tramway x Period 2	0.840 (0.087)	*	0.980	0.944 (0.093)		0.869	0.963 (0.092)		0.718
Tramway x Period 3	0.832 (0.072)	**	0.635	0.889 (0.074)		0.728	0.885 (0.071)		0.743
Distance to the city center	1.023 (0.017)		0.892	1.020 (0.015)		0.882	1.011 (0.014)		0.685
Distance to the closest terminus	0.980 (0.020)		0.293	0.974 (0.019)		0.405	0.962 (0.018)	**	0.961
Herfindahl index	0.938 (0.443)		0.259	0.908 (0.412)		0.542	0.948 (0.383)		0.112
LQ FIRE	0.808 (0.048)	***	0.886	0.793 (0.045)	***	0.683	0.772 (0.041)	***	0.677
LQ KIBS	0.997 (0.040)		0.505	1.000 (0.038)		0.177	1.017 (0.036)		0.189
LQ Business services	1.027 (0.037)		0.161	1.040 (0.034)		0.046	1.029 (0.032)		0.105
LQ Personal services	1.034 (0.030)		0.787	1.020 (0.026)		0.394	1.011 (0.025)		0.440
LQ Food and Accommodation	0.987 (0.030)		0.950	0.981 (0.029)		0.839	0.983 (0.027)		0.860
LQ Wholesale	0.940 (0.029)	**	0.327	0.934 (0.027)	**	0.261	0.958 (0.025)	*	0.135
Global PHA			0.285			0.077			0.021

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

Regarding the Schoenfeld residuals tests, the PHA is globally and individually respected for the 500m area, as well as for the 600m area, except for one covariate (LQ Business Services). For the 700m area, the PHA is respected for all covariates individually, while it is not globally respected, at a 5% level. Overall, the results of this robustness test are considered acceptable, since the PHA is respected (individually) from each covariate, and thus may be interpreted (with caution).

Furthermore, to verify that the results do not depend on a temporal specification, the model was also estimated using (January) 2016, and 2018 as a cut-off for the final period (Figure 5.4), i.e., the study period covers until December 2015 and December 2017, respectively. Initially, the operation period covers 7 years, as compared to 2 to 3 years for the predeceasing period. The test with shorter periods aims to verify if the impact appears to be effective years after the introduction of the tramway, or whether the mean treatment effect is influenced by the time span. The results are presented in Table 5.4 and Table 5.5.

**Figure 5.4 - Complete studied timeline in relation to Dijon’s tramway**



All coefficients obtained using different cut offs dates (January 1, 2016, 2018 and 2020) are relatively similar, except for the operation period (*Period 3*) coefficient, which is not surprising, since its composition/length was changed. While they are similar, a decreasing trend appears in the hazard ratios for the impact of the tramway in the last period (*Tramway x Period 3*), suggesting that the impact appears shortly after the introduction of the service and slightly decreases afterwards. Nevertheless, the difference in the hazard ratios is relatively small between the periods, reaching at most 0.107. The observed differences are greater for shorter distances (100m and 200m areas).

However, the results of the Schoenfeld residuals tests suggest that both temporal specifications used to test the robustness of the model do not respect the PHA. This should not be considered as an invalid robustness check as “these tests [for the PHA] are really just

model specification tests that verify you have adequately parametrized the model and you have chosen a good specification for  $x\beta_x$ .” (Cleves et al., 2010, p. 203). As a robustness check analysis, the results reflect the adequacy, the consistency and the overall fit of the model.

Overall, the model from which the conclusions are drawn respect the PHA, both individually and globally. It is also well-fitted, based on the sample and the covariates used (see Cox-Snell plots in *Annexe du Chapitre 5* - Figure A5.2). Finally, even if the temporal specifications do not respect the PHA, the coefficients, as well as their significance, are relatively robust to temporal changes. Therefore, the results presented in this section can be confidently discussed, based on the models estimated and the tests conducted.

**Table 5.4 - Cox model results for a shorter period, January 2006 – December 2015 inclusively (part 1)**

Critical distance	100 meters		200 meters		300 meters				
	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA			
Tramway	1.377 (0.277)	0.180	1.099 (0.138)	0.084	1.191 (0.114)	*	0.877		
Period 1 (Ann.)	0.962 (0.167)	0.310	1.051 (0.108)	0.320	1.069 (0.096)		0.013		
Period 2 (Const.)	1.164 (0.214)	0.640	1.115 (0.128)	0.084	1.114 (0.113)		0.004		
Period 3 (Oper.)	0.712 (0.131)	*	0.011	0.669 (0.077)	***	0.000	0.703 (0.071)	***	0.000
Tramway x P. 1	0.872 (0.238)		0.447	0.864 (0.143)		0.236	0.844 (0.108)		0.569
Tramway x P. 2	0.845 (0.259)		0.421	0.758 (0.138)		0.703	0.819 (0.117)		0.776
Tramway x P. 3	0.515 (0.161)	**	0.253	0.729 (0.134)	*	0.336	0.774 (0.109)	*	0.993
Distance to the city center	1.006 (0.059)		0.249	1.012 (0.031)		0.630	1.043 (0.025)	*	0.440
Distance to the closest terminus	0.998 (0.069)		0.306	1.004 (0.040)		0.792	0.992 (0.031)		0.263
Herfindahl index	0.378 (0.552)		0.076	0.750 (0.591)		0.266	0.399 (0.297)		0.103
LQ FIRE	0.811 (0.162)		0.038	0.815 (0.086)	*	0.051	0.853 (0.079)	*	0.063
LQ KIBS	0.932 (0.114)		0.430	0.921 (0.066)		0.671	0.962 (0.058)		0.957
LQ Business services	0.913 (0.100)		0.700	0.924 (0.067)		0.432	0.920 (0.055)		0.215
LQ Personal services	0.850 (0.079)	*	0.979	0.905 (0.051)	*	0.289	0.953 (0.042)		0.519
LQ Food and Acc.	0.978 (0.117)		0.217	0.978 (0.062)		0.571	0.989 (0.048)		0.786
LQ Wholesale	0.813 (0.103)		0.269	0.867 (0.061)	**	0.161	0.935 (0.046)		0.729
Global PHA		0.004			0.000				0.000

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

**Table 5.4 - Cox model results for a shorter period, January 2006 – December 2015 inclusively (part 2)**

Critical distance	400 meters			500 meters			600 meters		
	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA
Tramway	1.160 (0.100)	*	0.277	1.245(0.103)	***	0.547	1.207 (0.099)	**	0.392
Period 1 (Announcement)	1.095 (0.093)		0.006	1.163 (0.099)	*	0.058	1.189 (0.105)	**	0.344
Period 2 (Construction)	1.161 (0.111)		0.000	1.261 (0.118)	**	0.000	1.257 (0.121)	**	0.001
Period 3 (Operation)	0.769 (0.073)	***	0.000	0.843 (0.079)	*	0.000	0.825 (0.080)	**	0.000
Tramway x Period 1	0.918 (0.104)		0.313	0.887 (0.096)		0.803	0.894 (0.095)		0.555
Tramway x Period 2	0.860 (0.108)		0.224	0.846 (0.100)		0.692	0.889 (0.104)		0.899
Tramway x Period 3	0.765 (0.096)	**	0.864	0.706 (0.084)	***	0.352	0.779 (0.091)	**	0.275
Distance to the city center	1.017 (0.022)		0.960	1.016 (0.021)		0.726	1.024 (0.020)		0.594
Distance to the closest terminus	1.001 (0.028)		0.088	1.013 (0.027)		0.227	0.993 (0.026)		0.193
Herfindahl index	0.665 (0.423)		0.562	0.879 (0.496)		0.522	1.076 (0.567)		0.433
LQ FIRE	0.855 (0.072)	*	0.197	0.825 (0.062)	**	0.971	0.788 (0.056)	***	0.871
LQ KIBS	0.940 (0.050)		0.410	0.975 (0.048)		0.409	0.961 (0.044)		0.168
LQ Business services	0.972 (0.048)		0.141	0.968 (0.044)		0.064	1.001 (0.040)		0.007
LQ Personal services	1.008 (0.038)		0.623	1.031 (0.035)		0.762	1.015 (0.031)		0.576
LQ Food and Accommodation	0.997 (0.041)		0.993	1.016 (0.039)		0.701	1.011 (0.037)		0.639
LQ Wholesale	0.913 (0.039)	**	0.926	0.906 (0.037)	**	0.986	0.904 (0.034)	***	0.731
Global PHA			0.000			0.000			0.000

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

**Table 5.5 - Cox model results for a shorter period, January 2006 – December 2017 inclusively (part 1)**

Critical distance	100 meters			200 meters			300 meters		
	Haz. Ratio (Std. Err.)	PHA		Haz. Ratio (Std. Err.)	PHA		Haz. Ratio (Std. Err.)	PHA	
Tramway	1.444 (0.271)	**	0.066	1.113 (0.131)		0.053	1.193 (0.107)	**	0.747
Period 1 (Announcement)	0.937 (0.154)		0.319	1.115 (0.108)		0.930	1.100 (0.093)		0.080
Period 2 (Construction)	1.268 (0.219)		0.474	1.230 (0.132)	*	0.828	1.187 (0.112)	*	0.229
Period 3 (Operation)	0.998 (0.162)		0.284	0.981 (0.098)		0.049	1.053 (0.091)		0.000
Tramway x Period 1	0.805 (0.207)		0.173	0.833 (0.129)		0.148	0.834 (0.100)		0.683
Tramway x Period 2	0.750 (0.215)		0.154	0.737 (0.125)	*	0.351	0.824 (0.109)		0.896
Tramway x Period 3	0.513 (0.131)	***	0.154	0.799 (0.122)		0.093	0.801 (0.094)	*	0.887
Distance to the city center	1.056 (0.054)		0.542	1.010 (0.027)		0.119	1.031 (0.022)		0.111
Distance to the closest terminus	1.002 (0.061)		0.753	0.987 (0.035)		0.276	0.989 (0.027)		0.130
Herfindahl index	0.300 (0.410)		0.065	1.156 (0.793)		0.041	0.566 (0.368)		0.003
LQ FIRE	0.824 (0.145)		0.016	0.750 (0.068)	***	0.004	0.800 (0.064)	***	0.000
LQ KIBS	0.930 (0.104)		0.824	0.942 (0.061)		0.137	0.976 (0.053)		0.271
LQ Business services	0.971 (0.088)		0.832	0.975 (0.061)		0.292	0.956 (0.050)		0.269
LQ Personal services	0.913 (0.072)		0.013	0.934 (0.046)		0.143	0.968 (0.038)		0.757
LQ Food and Accommodation	0.984 (0.105)		0.432	0.975 (0.055)		0.818	0.952 (0.041)		0.265
LQ Wholesale	0.828 (0.092)	*	0.924	0.898 (0.053)	*	0.904	0.947 (0.040)	**	0.964
Global PHA			0.000			0.000			0.000

**Table 5.5 - Cox model results for a shorter period, January 2006 – December 2017 inclusively (part 2)**

Critical distance	400 meters			500 meters			600 meters		
	Haz. Ratio (Std. Err)		PHA	Haz. Ratio (Std. Err.)		PHA	Haz. Ratio (Std. Err.)		PHA
Tramway	1.165 (0.094)	*	0.413	1.231 (0.096)	***	0.799	1.198 (0.092)	**	0.713
Period 1 (Announcement)	1.119 (0.090)		0.036	1.178 (0.094)	**	0.118	1.182 (0.098)	**	0.240
Period 2 (Construction)	1.213 (0.108)	**	0.008	1.297 (0.113)	***	0.011	1.292 (0.116)	***	0.034
Period 3 (Operation)	1.150 (0.093)	*	0.000	1.213 (0.098)	**	0.000	1.173 (0.098)	*	0.000
Tramway x Period 1	0.922 (0.098)		0.388	0.901 (0.091)		0.679	0.923 (0.093)		0.884
Tramway x Period 2	0.878 (0.102)		0.286	0.866 (0.096)		0.592	0.913 (0.099)		0.845
Tramway x Period 3	0.784 (0.081)	**	0.886	0.749 (0.074)	***	0.641	0.827 (0.081)	*	0.564
Distance to the city center	1.014 (0.020)		0.459	1.014 (0.018)		0.677	1.019 (0.017)		0.880
Distance to the closest terminus	0.992 (0.024)		0.059	0.998 (0.023)		0.196	0.989 (0.022)		0.267
Herfindahl index	0.723 (0.415)		0.129	0.879 (0.452)		0.409	0.925 (0.454)		0.590
LQ FIRE	0.826 (0.060)	***	0.007	0.801 (0.053)	***	0.088	0.778 (0.049)	***	0.122
LQ KIBS	0.957 (0.046)		0.048	1.004 (0.044)		0.146	1.000 (0.042)		0.030
LQ Business services	0.997 (0.043)		0.548	0.986 (0.040)		0.359	1.013 (0.037)		0.068
LQ Personal services	1.015 (0.034)		0.905	1.033 (0.032)		0.753	1.018 (0.028)		0.488
LQ Food and Accommodation	0.981 (0.036)		0.478	1.001 (0.034)		0.746	1.005 (0.033)		0.955
LQ Wholesale	0.932 (0.035)	*	0.683	0.931 (0.033)	**	0.479	0.922 (0.030)	**	0.441
Global PHA			0.000			0.000			0.000

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

## **Conclusion**

This study aims to investigate the impact of the implementation of a tramway service on small businesses', i.e., having 2 employees or less, survival rates. To do so, a quasi-natural experiment based on the substitution of a bus rapid transit (BRT) service by a better public transport service is mobilized. The main assumption is that an upgrade of public transit supply quality, usually associated with an increase in the use of public transport, might be internalized into firms' longevity by further facilitating human contact, increasing consumption and employment opportunities, and ensuring greater accessibility. Literature suggests that such impact, when detected, is spatially concentrated around the new stations, mainly for firms active in tertiary (service) activities. To test those assumptions, a Cox proportional hazards (PH) model is applied to estimate the relative survival rates of firms.

The results suggest that the presence of the tramway station has an impact that is spatially concentrated around the station, i.e., within 500m. The study highlights a positive and significant impact of the upgrade of public transit on small businesses' survival after the opening of the service. No anticipation or speculation impact is identified on small businesses' survival. Moreover, the impact appears to have a U-shape, with important survival effect noticed in the direct vicinity, i.e., within less than 100m, as well as a little further away. The results suggest that small businesses' survival probabilities increase by, respectively, 18.0%, and 22.8% when located within 400m to 500m of a tramway station, instead of a BRT station. For shorter distances (200m and 300m areas), non-significant results arise, suggesting the impact on small businesses' survival is similar for the BRT network and the tramway lines.

From a public policy perspective, the results imply that developing a better public transport service within a city can influence individual small firms' performance. The results suggest that direct exposure to a new tramway station impacts firms' survival, after the service is in operation, while no statistical impact is noted during the announcement and the construction periods. Thus, even if public transport improvement does not necessarily reflect an improvement in travel time (Dubé et al., 2018), a better-quality service increases the use of the service and, indirectly, affects small firms' survival. The impact seems, in the worst case, non-significant between 100m to 300m, and appears to be significant within 300m to 500m,



but lower than direct exposure, i.e., within 100m. Consequently, indirect and wider economic effect should be taken into account when trying to evaluate the full impact of implementing new public transport services in a city.

This study faces certain limitations. First, no distinction is made between firms' creation and relocation. Thus, longevity may be underestimated for small businesses that relocates in the studied area, since the number of days in operation is calculated starting from its location date (which is assumed to be its opening date). Also, no conclusion can be drawn regarding the attraction of new or already operating firms around the tramway stations. Moreover, the study aggregates 4 economic sectors, meaning one sector could be driving the trend upward. While the global impact appears to be positive and significant, there is no way to identify from which sector this impact mostly comes from. There is a possibility that the significant impact might be driven by firms active in specific economic sectors.

To conclude, firms' longevity represents one of the aspects that defines firms' performance. Transport infrastructures, and change within the nature of those infrastructures, may affect firms' accessibility, visibility, consumer traffic and more, which in turn may impact their performances. More research on the matter is sorely needed in order to complement the current, relatively sparse literature. For future research, the use of financial data, or other intrinsic characteristics, to create a more complete profile of the firms studied could be an interesting addition to the investigation of firms' longevity and their probability of survival.

## Conclusion générale

Le rôle structurant des transports dans le développement et l'aménagement des milieux urbains est indéniable. L'accessibilité, et donc des déplacements plus efficaces, influence les décisions de localisation des ménages et des entreprises. Bien que les transports affectent les agents économiques, au sens large, cette thèse est réalisée dans une visée strictement commerciale.

En premier lieu, une analyse exhaustive de la littérature sur les impacts des infrastructures de transport sur les décisions de localisation des entreprises est effectuée. L'objectif de ce premier article est de brosser un portrait de la littérature inhérente à ce sujet. Basée sur une revue de littérature systématique, une méta-analyse est réalisée. Les résultats de la méta-analyse suggèrent que les décisions de localisation commerciales sont affectées positivement, dans le meilleur des cas, ou non affectés, dans le pire des cas, par la mise en place de ces infrastructures. Les dimensions spatiales et temporelles de la modélisation semblent également avoir un impact sur les résultats identifiés.

En théorie, les fondements économiques néo-classiques présentent les décisions de localisation comme un problème d'optimisation, basé sur une fonction de profit et une fonction de coût. L'optimisation des recettes repose sur de nombreux facteurs spécifiques à l'entreprise, tels que ses caractéristiques intrinsèques, sa consommation de capital (physique et humain), la composante physique de ses extrants, la localisation de son marché, et ses besoins en déplacement (coûts de transport). Les économies d'agglomération, associées à la concentration spatiale des activités économiques, représentent également un facteur d'influence dans les décisions de localisation commerciale, et ces dernières s'insèrent dans le problème d'optimisation. Le résultat de ce problème résulte d'un arbitrage effectué par l'entreprise afin de sélectionner une localisation jugée optimale.

La théorie souligne que, outre les caractéristiques intrinsèques, l'environnement spatial et les aménités locales influencent la localisation des entreprises. Afin d'identifier et d'attribuer un effet observé sur le développement commercial à une infrastructure de transport, une analyse statistique n'est souvent pas suffisante. Dans ces cas, une analyse causale est nécessaire. Ce type d'analyse va au-delà de la simple corrélation et cherche à isoler une cause et son effet.

L'évaluation causale, contrairement à l'analyse corrélacionnelle, repose sur l'identification de contrefactuels crédibles dans le but d'identifier un impact attribuable à l'infrastructure de transport.

Le concept d'infrastructure de transport est relativement large. En Amérique du Nord, particulièrement, ce concept est dominé par la présence d'infrastructures autoroutières. Cette offre démesurée de structure routière s'accompagne d'une dépendance automobile qui n'est pas sans conséquences (Brugge et al., 2007; Beaudoin et al., 2015; Zhong et al., 2017). Au Canada, en 2021, le secteur des transports était responsable de 22% des émissions de gaz à effet de serre (GES), soit le deuxième secteur le plus polluant en termes de GES, derrière l'industrie pétrolière et gazière (Gouvernement du Canada, 2023). De cette proportion, une grande majorité est associée aux déplacements automobiles individuels, impliquant une croissance des externalités négatives inhérentes à ce mode de transport, telle que des problèmes de congestion et de pollution (de l'air, visuelle et sonore) dans les métropoles.

En réponse aux changements climatiques et dans une démarche environnementale, les réseaux structurants de transport en commun représentent une alternative intéressante à la voiture. L'objectif du transport en commun est, entre autres, la réduction du nombre de voitures sur les routes et une diminution des externalités négatives qui en découlent. Cependant, l'instauration d'infrastructures de transport en commun nécessite des investissements d'envergure. Leur mise en place doit donc être réfléchiée et informée. Quantifier les impacts, au sens large, du transport collectif s'avère pertinent pour les décideurs publics, encore plus dans une perspective de développement durable.

Dans cette optique, les analyses empiriques présentées dans la thèse s'intéressent spécifiquement aux infrastructures de transport en commun. Les impacts du transport collectif sur les entreprises sont évalués sous différents angles dans la littérature. Tandis que les décisions de localisation sont largement abordées (Champagne et Dubé, 2023), l'analyse de la performance commerciale, en termes de survie, est peu développée. En réponse à une localisation théoriquement optimale, l'étude de la longévité commerciale est un angle d'approche complémentaire aux résultats présentés dans la littérature.

Les deux applications proposées souhaitent identifier l'impact causal du transport en commun sur la performance des entreprises, dans une perspective de survie. Un premier article étudie l'impact du prolongement du métro de Montréal. Le second article s'intéresse plus particulièrement aux petites entreprises et à leur probabilité de survie à la suite d'une substitution d'offre de transport en commun. Dans les deux cas, un modèle de risques proportionnels de Cox est appliqué afin d'étudier la probabilité de survie des entreprises. Ce modèle est primordial dans l'analyse de survie, puisqu'il permet une évaluation adéquate des observations censurées, soit les entreprises qui ne ferment pas avant la fin de l'étude.

Tout d'abord, la dernière extension du métro de Montréal (Canada) est étudiée. Cette extension de 3 stations, dans la ville de Laval, a été complétée en 2007. Un intérêt particulier du cas d'étude est l'annonce simultanée de deux projets d'extension du métro en 1998, soit environ 3 ans avant l'abandon d'un des projets, et près de 10 ans avant l'ouverture des 3 stations lavalloises. Le projet non-réalisé représente un contrefactuel intéressant dans l'optique où la demande anticipée pour le transport collectif et les effets d'anticipation (avant le début des travaux) devraient être similaires pour les deux projets proposés. Les résultats suggèrent que l'impact sur la probabilité de survie des entreprises se matérialise majoritairement après l'ouverture des stations, dans un rayon d'au plus 1 250 mètres des stations. Ces résultats sont spécifiques au contexte de Laval, où une vague de constructions immobilières (résidentielles et commerciales) a été observée jusqu'à 1 250 mètres des stations de l'extension (Champagne et al., 2022).

La seconde application empirique évalue l'impact de l'arrivée du tramway de Dijon (France) sur la survie des petites entreprises (2 employés et moins). Ce système structurant de transport en commun a été implanté en 2012, afin de remplacer, partiellement, le réseau d'autobus rapides (*Bus Rapid Transit – BRT*) déjà en place (Dubé et al., 2018). L'achalandage important du service d'autobus a incité Dijon à se tourner vers le tramway. Afin d'isoler l'impact causal du tramway sur la probabilité de survie des petites entreprises, les lignes d'autobus rapides antérieures à l'introduction du tramway sont utilisées comme contrefactuel. Les effets estimés sont positifs et localisés dans un rayon d'au plus 500 mètres des stations de tramway, après sa mise en service. Enfin, aucun impact n'est recensé avant l'ouverture des stations de tramway.

En somme, l'étude du transport en commun sur la longévité des entreprises identifie un impact temporel similaire pour les deux cas d'étude. En effet, les impacts des stations de transport en commun semblent être internalisés uniquement après l'ouverture du réseau de transport collectif. Par ailleurs, les impacts négatifs généralement associés à la période de construction, surtout dans le cas d'un tramway, ne se sont pas avérés statistiquement significatifs.

Toutefois, l'étendue spatiale de l'impact varie en fonction du contexte urbain et de l'environnement (composition du marché) dans lesquels évoluent les entreprises, mais aussi en fonction du type d'infrastructure considéré. Malgré tout, l'impact observé reste positif. Enfin, des analyses supplémentaires, mais surtout spécifiques à d'autres milieux urbains, seraient nécessaires pour consolider et potentiellement généraliser les tendances observées entre les deux cas étudiés.

En termes de politiques publiques, les résultats de cette thèse montrent que l'implantation d'un réseau de transport en commun structurant génère des effets plus larges que ceux associés uniquement aux utilisateurs. Dans un tel contexte, les analyses de types coûts-bénéfices mobilisées pour analyser les retombées de tels projets devraient également inclure les bénéfices liés aux entreprises situées à proximité des stations. Si l'effet sur les prix est largement documenté dans la littérature, cette thèse a comme retombée d'identifier des effets encore plus larges que ceux actuellement documentés. L'impact sur la survie des entreprises fait partie de ce que la littérature identifie comme les *widers economic impacts* (ou impacts économiques plus larges – Rothengatter, 2017), qui sont souvent négligés dans les processus de prises de décision.

## Bibliographie

- Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *econometrica*, 74(1), 235-267.
- Aïkous, M., Dubé, J., Brunelle, C., & Champagne, M. P. (2023). Is highway expansion and impacts on land use changes: An event study approach. *Transportation Research Part D: Transport and Environment*, 119, 103730.
- Akin, D., & Kara, D. (2020) Multicriteria Analysis of Planned Intercity Bus Terminals in the Metropolitan City of Istanbul, Turkey. *Transportation Research: Part A: Policy and Practice*, 132, 465-489. <https://doi.org/10.1016/j.tra.2019.12.003>.
- Alañón-Pardo, Á., & Arauzo-Carod, J.-M. (2013) Agglomeration, accessibility and industrial location: Evidence from Spain. *Entrepreneurship & Regional Development* 25(3-4), 135-173.
- Alin, A. (2010). Multicollinearity. *Wiley interdisciplinary reviews: computational statistics*, 2(3), 370-374.
- Alonso, W. (1964). *Location and land use: toward a general theory of land rent*. Harvard university press.
- Anavatan, A., & Karaoz, M. (2013) Cox Regression Models with Time-Varying Covariates Applied to Survival Success of Young Firms. *Journal of Economic and Social Studies*, 3(2), 53-68.
- Andersson, M., & Larsson, J. P. (2016). Local entrepreneurship clusters in cities. *Journal of Economic Geography*, 16(1), 39-66.
- Andersson, T. M. L., & Lambert, P. C. (2012) Fitting and modeling cure in population-based cancer studies within the framework of flexible parametric survival models. *The Stata Journal*, 12(4), 623-638.
- Ansin, E. (2015). An evaluation of the Cox-Snell residuals (Dissertation). Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-256665>
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The leadership quarterly*, 21(6), 1086-1120.
- Apparicio, P., Abdelmajid, M., Riva, M., & Shearmur, R. (2008) Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *Int J Health Geogr*. 18(7), 7.
- Arrow, K. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3), 155-173.
- Arzaghi, M., & Henderson, J. V. (2008) Networking off Madison Avenue. *Review of Economic Studies*, 75(4), 1011-1038.
- Audretsch, D. B. (1995) *Innovation and Industry Evolution*. MIT Press, Cambridge.
- Audretsch, D. B., & Mahmood, T. (1995) New Firm Survival: New Results Using a Hazard Function. *The Review of Economics and Statistics*, 77(1), 97-103.
- Baptista, R., & Swann, P. (1998). Do firms in clusters innovate more? *Research Policy*, 27, 525-40.

- Bates, T. (1990). Entrepreneur human capital inputs and small business longevity. *The Review of Economics and Statistics*, 72, 551–559. doi:10.2307/2109594
- Bates, T., & Nucci, A. (1989). An analysis of small business size and rate of discontinuance. *Journal of Small Business Management*, 27, 1.
- Baum-Snow, N. (2007a) Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2), 775-805.
- Baum-Snow, N. (2007b). Suburbanization and transportation in the monocentric model. *Journal of Urban Economics*, 62(3), 405-423.
- Bazin, S., Beckerich, C., & Delaplace, M. (2009) Dessertes TGV et localisation des entreprises dans les quartiers d'affaires : Nouvelle accessibilité ou nouvelle offre de bureaux. *Les Cahiers Scientifiques du Transport*, 56, 37-61.
- Beaudoin, J., & Lawell, C. Y. (2018) The Effects of Public Transit Supply on the Demand for Automobile Travel. *Journal of Environmental Economics and Management*, 88, 447-467.
- Beaudoin, J., Farzin, Y. H., & Lawell, C. Y. (2015) Public transit investment and sustainable transportation: A review of studies of transit's impact on traffic congestion and air quality. *Research in Transportation Economics*, 52, 15-22.
- Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy* 38, 318–337.
- Beckerich, C., Bazin, S., & Delaplace, M. (2016). Dessertes TGV et localisation des entreprises dans les quartiers de gare : une activation du potentiel de proximité avec Paris? Le cas du quartier Clairmarais à Reims. *Les Cahiers Scientifiques Du Transport*, 69–70, 3–35.
- Beckerich, C., Benoit, S., & Delaplace, M. (2019) Are the reasons for companies to locate around central versus peripheral high-speed rail stations different? The cases of Reims central station and Champagne-Ardenne station. *European Planning Studies* 27(3), 574–594.
- Belleau-Arsenault, C., & Dubé, J. (2019). Quels impacts de l'aide financière sur la performance des établissements manufacturiers en régions non-métropolitaines? Une réponse par méthode d'appariement spatial. *Revue canadienne de science régionale*, 42(1), 11-23.
- Bhat, C. R., Paleti, R., & Singh, P. (2014) A Spatial Multivariate Count Model for Firm Location Decisions: Firm Location Decisions. *Journal of Regional Science* 54(3), 462–502.
- Bodenmann, B. R., & Axhausen, K. W. (2012) Effects and Side Effects of Measures to Attract Firms—A Micro-Simulation Study of Firm Location Choice. *disP—The Planning Review* 48(4), 14–28.
- Bogart, W. T., & Ferry, W. C. (1999) Employment Centres in Greater Cleveland: Evidence of Evolution in a Formerly Monocentric City. *Urban Studies*, 36 (12), 2099-110.
- Bordeleau, S. (Déc. 2022) Prolongement de la ligne bleue : l'appel d'offres est lancé. *Radio-Canada*. <https://ici.radio-canada.ca/nouvelle/1938730/prolongement-ligne-bleue-appel-offres>
- Box-Steffensmeier, J. M., & Jones, B. S. (2004) *Event history modeling: A guide for social scientists*. Cambridge University Press.

- Brugge, D., Durant, J. L., & Rioux, C. (2007). Near-highway pollutants in motor vehicle exhaust: a review of epidemiologic evidence of cardiac and pulmonary health risks. *Environmental health*, 6, 1-12.
- Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *J. Econ. Surv.*, 22 (1), 31–72.
- Calthorpe, P. (1993). *The Next American Metropolis*. New York: Princeton Architectural Press, 175 p.
- Card, D., Kluve, J., & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548), F452-77.
- Castillo-Manzano, J. I., & López-Valpuesta, L. (2009) Urban retail fabric and the metro: A complex relationship. Lessons from middle-sized Spanish cities. *Cities* 26(3), 141–147.
- Caves, R. E. (1998) Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature*, 36, 1947–1982.
- Cervero, R. (1994). Rail transit and joint development: Land market impacts in Washington, DC, and Atlanta. *Journal of the American Planning Association*, 60 (1): 83-94
- Cervero, R., & Landis, J. (1997). Twenty years of the Bay Area Rapid Transit System: land use and development impacts. *Transportation Research Part A: Policy Practice*, 31A, 309–333.
- Champagne, M. P., & Dubé, J. (2023). The impact of transport infrastructure on firms' location decision: A meta-analysis based on a systematic literature review. *Transport Policy*, 131, 139-155.
- Champagne, M. P., Dubé, J., & Barla, P. (2022). Build it and they will come: How does a new public transit station influence building construction?. *Journal of Transport Geography*, 100, 103320.
- Champagne, M. P., Dubé, J., & Legros, D. (2024) Standing strong? The causal impact of metro stations on service firms' survival. *Transportation Research Part A: Policy and Practice*, 181, 103994.
- Chang, T., Yang, D., Yang, Y., Huo, J., Wang, G., & Xiong, C. (2018) Impact of Urban Rail Transit on Business Districts Based on Time Distance: Urumqi Light Rail. *Journal of Urban Planning and Development* 144(3), 1–14.
- Chatman, D. G., & Noland, R. B. (2011) Do Public Transport Improvements Increase Agglomeration Economies? A Review of Literature and an Agenda for Research. *Transport Reviews*, 31 (6), 725-742.
- Chatman, D. G., Noland, R. B., & Klein, N. J. (2016) Firm Births, Access to Transit, and Agglomeration in Portland, Oregon, and Dallas, Texas. *Transportation Research Record: Journal of the Transportation Research Board*, 2598(1), 1-10.
- Chatterjee, S., & Hadi, A. S. (2012) *Regression Analysis by Example*. 5th ed. New York: Hoboken, NJ.
- Christaller, W. (1933). *Die zentralen Orte in Süddeutschland*. Jena: Fischer. (English translation by C.W. Baskin. Central Places in Southern Germany. Englewood Cliffs, NJ: Prentice-Hall.)
- Cissé, I., Dubé, J. & Brunelle, C. (2020). New Business Location: How Local Characteristics Influence Individual Location Decision? *Annals of Regional Science*, 64, 185–214



- Clarke, D., & Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal*, 21(4), 853-884.
- Cleves, M. A., Gould, W., & Marchenko, Y. V. (2010). *An introduction to survival analysis using Stata* (Revised third edition). Stata Press. 412 p.
- Coalition Avenir Québec (2018) Plan de Décongestion : Une réponse sensée pour améliorer la mobilité dans la région métropolitaine. CAQ <https://coalitionavenirquebec.org/wp-content/uploads/2018/08/decongestion.pdf>
- Cordera, R., Coppola, P., Dell'Olio, L., & Ibeas, Á. (2019). The impact of accessibility by public transport on real estate values: A comparison between the cities of Rome and Santander. *Transportation Research Part A: Policy and Practice*, 125, 308-319.
- Coughlin, C. C., & Segev, E. (2000). Location Determinants of New Foreign-Owned Manufacturing Plants. *Journal of Regional Science*, 40(2), 323–51.
- Cox, D. R. (1972) Regression Models and Life Tables. *Journal of the Royal Statistical Society*, 34, 187–220.
- Cox, D.R., & Oakes D. (1984) *Analysis of Survival Data*. Chapman and Hall, London.
- Cox, D. R., & Snell, E. J. (1968) A general definition of residuals. *Journal of the Royal Statistical Society, series B*, 30(2), 248–275.
- Credit, K. (2018) Transit-Oriented Economic Development: The Impact of Light Rail on New Business Starts in the Phoenix, AZ Region, USA. *Urban Studies*, 55(13), 2838-2862.
- Credit, K. (2019) Transitive properties: a spatial econometric analysis of new business creation around transit. *Spatial Economic Analysis*, 14(1), 26-52.
- Curtis, C., & Mellor, R. (2011) Anticipating a New Railway: The Relevance of Transit-Oriented Development to Businesses in a Car-Oriented City. *Urban Policy and Research* 29(2), 141–165.
- D'Agostino Jr, R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in medicine*, 17(19), 2265-2281.
- Daoud, J. I. (2017, December). Multicollinearity and regression analysis. In *Journal of Physics: Conference Series* (Vol. 949, No. 1, p. 012009). IOP Publishing.
- Dauda, K. A., Yahya, W. B., & Banjoko, A. W. (2015) Survival analysis with multivariate adaptive regression splines using cox-snell residual. *Annals. Computer Science Series*, 13(2), 25-41.
- de Bok, M. (2009) Estimation and validation of a microscopic model for spatial economic effects of transport infrastructure. *Transportation Research Part A: Policy and Practice* 43(1), 44–59.
- De Groot, H. L., Poot, J., & Smit, M. J. (2016). Which agglomeration externalities matter most and why? *Journal of Economic Surveys*, 30(4), 756-82.
- De Solminihac, H. E., & Harrison, R. (1993). Analyzing effects of highway rehabilitation on businesses. *Transportation Research Record*, 1397, 137–143.
- Debrezion, G., Pels, E., & Rietveld, P. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, 35(2), 161-80.

- Des Rosiers, F., Thériault, M., Voisin, M., & Dubé, J. (2010). Does an improved urban bus service affect house values?. *International Journal of Sustainable Transportation*, 4(6), 321-346.
- Devaux, N., Dubé, J., & Apparicio, P. (2017) Anticipation and Post-Construction Impact of a Metro Extension on Residential Values: The Case of Laval (Canada), 1995–2013. *Journal of Transport Geography*, 62, 8-19.
- Diao, M., Leonard, D., & Sing, T. F. (2017) Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, 67(C), 64-77.
- Dijon métropole (anciennement Communauté urbaine du Grand Dijon). *Réalisation de deux lignes de tramway dans l'agglomération dijonnaise : Dossier de candidature - Appel à projets « Transports urbains » Hors Île-de-France - Suites du Grenelle Environnement*, Dijon, 2009, 68p. <https://archive.wikiwix.com/cache/index2.php?url=http%3A%2F%2Fwww.grand-dijon.fr%2Fonlinemedia%2FUpload%2FAppelaprojetsTRAM.pdf%2Findex.html#federation=archive.wikiwix.com&tab=url>
- Doms, M., Dunne T., & Roberts M.J. (1995) The Role of Technology use in the Survival and Growth of Manufacturing Plants. *International Journal of Industrial Organization*, 13, 523–542.
- Downs, A. (1962). The Law of Peak-Hour Expressway Congestion. *Traffic Quarterly*, 16(3): 393–409.
- Dubé J., Brunelle C., & Legros D. (2016). Location Theories and Business Location Decision: A Micro-Spatial Investigation in Canada. *Review of Regional Studies*, 46, 143.
- Dubé, J., & Brunelle, C. (2014) Dots to Dots: A General Methodology to Build Local Indicators Using Spatial Micro-Data. *The Annals of Regional Science*, 53(1), 245-272.
- Dubé, J., Babin, C.-D., Dubé, J.-C., Lekkat, H., Potvin, A., & Ringue, O. (2017). What Comes First, Residential or Commercial Development? Measuring the Causal Links for a Canadian City over a Century, Canada, *Letters in Spatial and Resource Sciences*, 10(1), 57-74.
- Dubé, J., Dieng, O. Lévesque, M., Racine, O., Roberge, O., & Trapé, T. (2023a). How Public Policies and Other Events can Shape Spatial Distribution of Local Activities over Time? An Investigation based on Spatial Micro-data, *Land Use Policy*, 125: 106498.
- Dubé, J., Cardenas, D., & Champagne, M. P. (2023b). Tests de randomisation: une façon plus flexible de tester la significativité avec le logiciel Stata. *Canadian Journal of Regional Science*, 46(3), 50-56.
- Dubé, J., Des Rosiers, F., Thériault, M., & Dib, P. (2011). Economic impact of a supply change in mass transit in urban areas: a Canadian example. *Transportation Research Part A: Policy and Practice*, 45(1), 46-62.
- Dubé, J., Legros, D., & Devaux, N. (2018). From bus to tramway: Is there an economic impact of substituting a rapid mass transit system? An empirical investigation accounting for anticipation effect. *Transportation Research Part A: Policy and Practice*, 110, 73-87.
- Dubé, J., Legros, D., Thériault, M., & Des Rosiers, F. (2017). Measuring and interpreting urban externalities in real-estate data: a spatio-temporal difference-in-differences (STDID) estimator. *Buildings*, 7(2), 51.

- Dubé, J., Legros, D., Thériault, M., & Des Rosiers, F. (2014). A spatial difference-in-differences estimator to evaluate the effect of change in public mass transit systems on house prices. *Transportation Research Part B: Methodological*, 64, 24-40.
- Dubé, J., Thériault, M., & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice*, 54, 49-66.
- Duncan, M. (2011) The Impact of Transit-Oriented Development on Housing Prices in San Diego, CA. *Urban Studies*, 48(1), 101-27.
- Dunne, T., Roberts M. J. & Samuelson, L. (1988) Patterns of Firm Entry and Exit in US Manufacturing Industries. *RAND Journal of Economics*, 19(4), 495–515.
- Duranton, G., & Turner, M. A. (2011). The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review*, 101 (6): 2616-52.
- Duranton, G., & Turner, M. A. (2012). Urban Growth and Transportation. *The Review of Economic Studies*, 79(4): 1407–1440. <http://www.jstor.org/stable/23355076>
- Duvivier, C. (2019). Broadband and firm location: Some answers to relevant policy and research issues using a meta-analysis. *Canadian Journal of Regional Science/Revue Canadienne des Sciences Régionales*, 42(1), 24-45.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629-63.
- Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The Journal of Industrial Economics*, 35, 567–581. doi:10.2307/2098588
- Figueiredo, R. J., Neto, J. V., Quelhas, O. L. G., & Ferreira, J. J. d. M. (2017). Knowledge Intensive Business Services (KIBS): bibliometric analysis and their different behaviors in the scientific literature. *RAI Revista de Administração e Inovação*, 14(3), 216-25.
- Forkenbrock, D. J., & Foster, N., S. J. (1996) Highways and Business Location Decisions. *Economic Development Quarterly* 10(3), 239–248.
- Fotopoulos, G., & Louri, H. (2000) Location and Survival of New Entry. *Small Business Economics*, 14, 311-321.
- Fujita, M. (1989) *Urban economic theory: Land use and city size*. Cambridge University Press, UK.
- Fujita, M., Krugman, P., & Mori, T. (1999). On the evolution of hierarchical urban systems. *European Economic Review*, 43(2), 209-251.
- Ganning, J., & Miller, M. (2020). Transit oriented development and retail: Is variation in success explained by a gap between theory and practice? *Transportation Research Part D*, 85, 102357.
- Geroski P.A. (1995) What do we know about entry? *International Journal of Industrial Organization*, 13, 421–440.
- Geroski, P. A., Mata, J., & Portugal, P. (2010) Founding conditions and the survival of new firms. *Strategic Management Journal*, 31, 510-529.

Ghani, E., Goswami, A. G., & Kerr, W. R. (2016) Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *The Economic Journal* 126(591), 317–357.

Gibbons, S., Lyytikäinen, T., Overman, H. G., & Sanchis-Guarner, R. (2019). New road infrastructure: The effects on firms. *Journal of Urban Economics*, 110, 35–50.

Gilbert, D., & Baker, N. (Jan. 2020) Montreal Metro. *The Canadian Encyclopedia*  
<https://www.thecanadianencyclopedia.ca/en/article/montreal-metro>

Gingerich, K., & Maoh, H. (2019) The role of airport proximity on warehouse location and associated truck trips: Evidence from Toronto, Ontario. *Journal of Transport Geography* 74, 97–109.

Giuliano G., & Small, K. A. (1991) Subcenters in the Los Angeles region. *Regional Science and Urban Economics*, 2(2), 163-182.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of urban Economics*, 63(1), 1-24.

Glaeser, E.L., & Kohlhase, J.E. (2004). Cities, regions and the decline of transport costs. In: Florax, R.J.G.M., Plane, D.A. (eds) *Fifty Years of Regional Science. Advances in Spatial Science*. Springer, Berlin, Heidelberg.

Goel, M.K., Khanna, P., & Kishore, J. (2010) Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research*, 1(4), 274-278.

Goetz, E.G., Ko, K., Hagar, A., Ton, H., & Matson, J., (2010) The Hiawatha Line: impacts on land use and residential housing value. *Center for Transportation Studies Report* 10-09. University of Minnesota.

Gómez-Antonio, M., & Sweeney, S. (2018) Firm location, interaction, and local characteristics: A case study for Madrid's electronics sector: Firms location choice interaction and local characteristics. *Papers in Regional Science* 97(3), 663–685.

Gouvernement du Canada (2023). Émissions de gaz à effet de serre : Émissions de gaz à effet de serre par secteur économique. *Environnement et Changement climatique Canada*.  
<https://www.canada.ca/fr/environnement-changement-climatique/services/indicateurs-environnementaux/emissions-gaz-effet-serre.html#transports>

Government of Canada, (2020). Backgrounder: A Plan to Permanently Fund Public Transit and Support Economic Recovery. *Infrastructure Canada*. <https://www.canada.ca/en/office-infrastructure/news/2021/02/a-plan-to-permanently-fund-public-transit-and-support-economic-recovery.html>.

Grambsch, P., M., and Therneau. T. M. (1994) Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika*, 81(3), 515–526.

Grossman, G. M., & Helpman, E. (1994). Protection for sale. *The American Economic Review*, 84(4), 833-850.

Hailpern, S. M., & Visintainer, P. F. (2003). Odds ratios and logistic regression: further examples of their use and interpretation. *The Stata Journal*, 3 (3), 213–25.

Headd, B. (2003). Redefining business success: Distinguishing between closure and failure. *Small Business Economics*, 21, 51–61.

- Henderson, J. V., & Ono, Y. (2008). Where do manufacturing firms locate their headquarters? *Journal of Urban Economics*, 63(2), 431-50.
- Higgins, C., & Kanaroglou, P. (2018). Rapid transit, transit-oriented development, and the contextual sensitivity of land value uplift in Toronto. *Urban Studies*, 55(10), 2197-2225.
- Hinchliffe, S. R., Scott, D. A., & Lambert, P. C. (2013) Flexible parametric illness-death models. *Stata Journal*, 13, 759–775.
- Hoffman, J. I. E. (2019). Chapter 36 - *Meta-analysis*, in Basic Biostatistics for Medical and Biomedical Practitioners (Second Edition), 621-629.
- Holl, A. (2004a). Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. *Regional Science and Urban Economics*, 34(3), 341-363.
- Holl, A. (2004b). Startups and relocations: Manufacturing plant location in Portugal. *Papers in Regional Science*, 83(4), 649-668.
- Holl, A. (2004c) The role of transport in firms' spatial organization: evidence from the Spanish food processing industry. *European Planning Studies* 12(4), 537–550.
- Holl, A. (2004d). Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal. *Journal of Regional Science*, 44(4), 693–712.
- Holl, A. (2006). A Review of the Firm-Level Role of Transport Infrastructure with Implications for Transport Project Evaluation. *Journal of Planning Literature*, 21(1), 3–14.
- Holl, A., & Mariotti, I. (2018) The Geography of Logistics Firm Location: The Role of Accessibility. *Networks and Spatial Economics* 18(2), 337–361.
- Holland, S. (1976). *Capital versus the Regions*. Macmillan, London, UK.
- Hong, J. (2007a) Location Determinants and Patterns of Foreign Logistics Services in Shanghai, China. *The Service Industries Journal* 27(4), 339–354.
- Hong, J. (2007b) Transport and the location of foreign logistics firms: The Chinese experience. *Transportation Research Part A: Policy and Practice* 41(6), 597–609.
- Hong, J., & Chin, A.T.H. (2007) Modeling the location choices of foreign investments in Chinese logistics industry. *China Economic Review* 18(4), 425–437.
- Hoogstra, G.J., van Dijk, J., & Florax, R.G.M. (2017). Do jobs follow people or people follow jobs? A meta-analysis of Carlino-Mills studies. *Spat. Econ. Anal.*, 12 (4), 357–378.
- Horváthová, E. (2010). Does environmental performance affect financial performance? A meta-analysis. *Ecological economics*, 70(1), 52-59.
- Hurst, N. B., & West, S. E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics*, 46, 57-72.
- INSEE (Institut National de la Statistique et des Études Économiques), 2024. Comparateur de territoires : Dijon. <https://www.insee.fr/fr/statistiques/1405599?geo=EPCI-242100410+COM-21231>
- Iseki, H., & Eom, H. (2019) Impacts of Rail Transit Accessibility on Firm Spatial Distribution: Case Study in the Metropolitan Area of Washington, DC. *Transportation Research Record: Journal of the Transportation Research Board*, 2673(11), 220–232.

- Iseki, H., & Jones, R. P. (2018) Analysis of firm location and relocation in relation to Maryland and Washington, DC metro rail stations. *Research in Transportation Economics* 67, 29–43.
- Jacobs, J. (1969). *The Economy of Cities*. Random House, New York, USA.
- Jiang, Y., Ng, A.K.Y., Wang, Y., Wang, L., & Yu, B. (2018a) Locational Characteristics of Firms in the Business Service Industry in Airport Economic Zones: Case of Shanghai Hongqiao International Airport. *Journal of Urban Planning and Development* 144(1), 1–13.
- Jiang, H., Xu, W., & Zhang, W. (2018b) Transportation Accessibility and Location Choice of Japanese-Funded Electronic Information Manufacturing Firms in Shanghai. *Sustainability* 10(2), 1–21.
- Kang, S. (2020) Warehouse location choice: A case study in Los Angeles, CA. *Journal of Transport Geography* 88, 1–11.
- Kang, S. J., & Lee, H. S. (2007) The determinants of location choice of South Korean FDI in China. *Japan and the World Economy* 19(4), 441–460.
- Kennedy, P. (2003). *A guide to econometrics* (5th ed.). MIT Press, Cambridge, MA, USA.
- Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of economic literature*, 26(2), 646-679.
- Kim, H., Ahn, S., & Ulfarsson, G. F. (2018) Transportation infrastructure investment and the location of new manufacturing around South Korea’s West Coast Expressway. *Transport Policy* 66, 146–154.
- Kleinbaum, D. G., & Klein, M. (2005) *Survival Analysis: A self-learning text* (2nd Ed.), Springer, New York, USA.
- Kuitunen, I., Ponkilainen, V.T., Uimonen, M.M., Eskelinen, A., & Reito, A. (2021) Testing the proportional hazards assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: methodological perspectives and review. *BMC Musculoskeletal Disord*, 22(1), 1-7.
- Kulish, M., Richards, A., & Gillitzer, C. (2012) Urban Structure and Housing Prices: Some Evidence from Australian Cities. *Economic Record*, 88 (282), 303-22.
- Le Devoir. (Sept. 2013) Québec va de l’avant avec le prolongement de la ligne Bleue. *Le Devoir*. <https://www.ledevoir.com/politique/quebec/387911/metro-quebec-va-de-l-avant-avec-le-prolongement-de-la-ligne-bleue>
- Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (Vol. 476). John Wiley & Sons.
- Leitham, S., McQuaid, R.W., & Nelson, J. D. (2000) The influence of transport on industrial location choice: a stated preference experiment. *Transportation Research Part A*, 34, 515–535.
- LeSage J. P., & Pace, R. K. (2009) *Introduction to spatial econometrics*. CRC Press, New York, USA.
- Li, J., & Huang, H. (2020) Effects of transit-oriented development (TOD) on housing prices: A case study in Wuhan, China. *Research in Transportation Economics*, 80, 100813.

- López, F. A., & Páez, A. (2017) Spatial clustering of high-tech manufacturing and knowledge-intensive service firms in the Greater Toronto Area: Spatial clustering in the GTA. *The Canadian Geographer/Le Géographe canadien* 61(2), 240–252.
- Lösch, A. (1954) *The economics of location*. Yale University Press, New Haven, USA.
- Low, S. A., & Brown, J. P. (2017) Manufacturing Plant Survival in a Period of Decline: Manufacturing Plant Survival. *Growth and Change*, 48(3), 297-312.
- Manzato, G. G., Arentze, T. A., Timmermans, H. J. P., & Ettema, D. (2010) Location and Accessibility Mediated Influences on Office Firm Closure Rates: A Proportional Hazard Model. *International Journal of Urban Sciences*, 14(1), 1-15.
- Manzato, G. G., Arentze, T. A., Timmermans, H. J. P., & Ettema, D. (2011) Exploration of Location Influences on Firm Survival Rates with Parametric Duration Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2245(1), 124-130.
- Maoh, H., & Kanaroglou, P. (2009) Intrametropolitan Location of Business Establishments: Microanalytical Model for Hamilton, Ontario, Canada. *Transportation Research Record: Journal of the Transportation Research Board* 2133(1), 33–45.
- Marshall, A. (1891). *Principles of Economics*. Macmillan and Co., London, UK.
- Marubini, E., & Valsecchi, M.G. (1995) Estimation of Survival Probabilities. In: *Analysing survival data from clinical trials and observational studies* pp. 41-48. John Wiley and Sons, Chichester (UK).
- Mata, J., & Portugal, P. (1994) Life duration of new firms. *The journal of industrial economics*, 42(3), 227-245.
- Mata, J., Portugal, P., & Guimarães, P. (1995) The survival of new plants: Start-up conditions and post-entry evolution. *International Journal of Industrial Organization*, 13, 459–481.
- Matas, A., Raymond, J. L., & Roig, J. L. (2020) Evaluating the impacts of HSR stations on the creation of firms. *Transport Policy* 99, 396–404.
- Mayer, T., & Trevien, C. (2017) The impact of urban public transportation evidence from the Paris region. *Journal of Urban Economics*, 102, 1–21.
- McDonald, J. F., & McMillen, D. P. (1990) Employment subcenters and land values in a polycentric urban area: the case of Chicago. *Environment and Planning A*, 22, 1561-74.
- Mejia-Dorantes, L., Paez, A., & Vassallo, J. M. (2012) Transportation infrastructure impacts on firm location: the effect of a new metro line in the suburbs of Madrid. *Journal of Transport Geography* 22, 236–250.
- Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics* 39(3), 332–342.
- Melo, P.C., Graham, D.J., & Noland, R.B. (2010) Impact of Transport Infrastructure on Firm Formation: Evidence from Portuguese Municipalities. *Transportation Research Record: Journal of the Transportation Research Board* 2163(1), 133–143.
- Mills, E. (1972). *Urban Economics*, Scott Foresman, Glenview, Illinois, USA.

- Mohammad, S.I., Graham, D.J., Melo, P.C., & Anderson, R.J. (2013) A Meta-analysis of the Impact of Rail Projects on Land and Property Values, *Transportation Research Part A*, 50, 158-170.
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18, 143-150.
- Muth, R. F. (1969) *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. University of Chicago Press, Chicago, US.
- NASA (2016). Earth at Night: Flat Map. *Earth observatory NASA*.  
<https://earthobservatory.nasa.gov/features/NightLights>
- Nelson, A. C. (1999) Transit stations and commercial property values: a case study with policy and land-use implications. *Journal of Public Transportation*, 2(3), 77-95.
- Nilsson, I. M., & Smirnov, O. A. (2016) Measuring the effect of transportation infrastructure on retail firm co-location patterns. *Journal of Transport Geography* 51, 110–118.
- O'Connor, K. (1989) Australian ports, metropolitan areas and trade-related services. *Australian Geographer* 20(2), 167–172.
- Pe'er, A. & Keil, T. (2013) Are all startups affected similarly by clusters? Agglomeration, competition, firm heterogeneity, and survival. *Journal of Business Venturing*, 28(3), 354-372.
- Pe'er, A., Vertinsky, I. & King, A. (2008) Who enters, where and why? The influence of capabilities and initial resource endowments on the location choices of de novo enterprises. *Strategic Organization*, 6(2), 119-149.
- Peeters, D., Thisse, J.-F., & Thomas, I. (2000) On High-Speed Connections and the Location of Activities. *Environment and Planning A: Economy and Space*, 32(12), 2097–2112.
- Percoco, M. (2016) Highways, local economic structure and urban development. *Journal of Economic Geography*, 16(5), 1035–1054.
- 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.
- Petráš, M., & Květoň, V. (2020). Spatial changes in distribution of firms and selected industries around the new motorway. *Applied Geography*, 121, 1–15.
- Phillips, B. D., & Kirchoff, B. A. (1989). Formation, growth and survival; Small firm dynamics in the U.S. economy. *Small Business Economics*, 1, 65–74.
- Porter, M.E. (1990). *The Competitive Advantage of Nations*. Free Press, New York.
- Porter, M.E. (1998) *On Competition*. Harvard Business School Press, Boston.
- Radio-Canada. (May 2018) La CAQ ne veut pas d'une ligne rose de métro à Montréal. *Radio-Canada*. <https://ici.radio-canada.ca/nouvelle/1103353/legault-metro-montreal>
- Ramaul, N. K., & Ramaul, P. (2016) Determinants of Industrial Location Choice in India: A Polychoric Principal Component Analysis Approach. *Journal of Quantitative Economics*, 14(1), 29–56.



- Ramón-solans Prat, J.C., & Marcén, R.F. (2006) Influential Factors in Location Choice of Spanish Businesses in Aragon. *The Journal of Entrepreneurship*, 15(1), 63–81.
- Ray, R. (2017). Open for business? Effects of Los Angeles metro rail construction on adjacent businesses. *Journal of Transport and Land Use*, 10, 725–742. doi:10.5198/jtlu.2017.932
- Renski, H. (2011) External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science*, 90(3), 473-502.
- Rich, J.T., Neely, J.G., Paniello, R.C., Voelker, C.C., Nussenbaum, B., & Wang, E.W. (2010) A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg.*, 143(3), 331-336.
- Romer, P. (1986). Increasing Returns and Long-Run Growth. *The Journal of Political Economy*, 94(5), 1002-1037.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*, 79(387), 516-524.
- Roth, J., Sant’Anna, P. H., Bilinski, A., & Poe, J. (2023). What’s trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218-2244.
- Rothengatter, W. (2017). Wider economic impacts of transport infrastructure investments: Relevant or negligible?. *Transport Policy*, 59, 124-133.
- Santarelli, E. (2000) The duration of new firms in banking: An application of Cox regression analysis. *Empirical Economics*, 25, 315-325.
- Schoenfeld, D. (1982) Partial Residuals for The Proportional Hazards Regression Model. *Biometrika*, 69(1), 239–241.
- Schoenmaker, D.A.J., & Van der Vlist, A.J. (2015). On real estate development activity: the relationship between commercial and residential real estate markets. *Lett. Spat. Resour. Sci.*, 8, 219–232.
- Schuetz, J. (2015). Do rail transit stations encourage neighbourhood retail activity? *Urban Studies*, 52(14), 2699-2723.
- Scoppa, M. D., & Peponis, J. (2015) Distributed Attraction: The Effects of Street Network Connectivity Upon the Distribution of Retail Frontage in the City of Buenos Aires. *Environment and Planning B: Planning and Design*, 42(2), 354–378.
- Secrétariat du Conseil du Trésor. (March 2023). Métro de Montréal, Ligne Bleue de la station Saint-Michel à Anjou- Prolongement. *Quebec Government* [https://www.tresor.gouv.qc.ca/infrastructures-publiques/tableau-de-bord?tx\\_tdbpip\\_tdbpip%5Baction%5D=show&tx\\_tdbpip\\_tdbpip%5Bcontroller%5D=Projet&tx\\_tdbpip\\_tdbpip%5BcurrentPage%5D=1&tx\\_tdbpip\\_tdbpip%5Bprojet%5D=39&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5Bactivite%5D=0&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5Bavancement%5D=0&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5BmotsCles%5D=Bleue&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5Bnouveau%5D=0&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5Bnumero%5D=0](https://www.tresor.gouv.qc.ca/infrastructures-publiques/tableau-de-bord?tx_tdbpip_tdbpip%5Baction%5D=show&tx_tdbpip_tdbpip%5Bcontroller%5D=Projet&tx_tdbpip_tdbpip%5BcurrentPage%5D=1&tx_tdbpip_tdbpip%5Bprojet%5D=39&tx_tdbpip_tdbpip%5BrechercheInfo%5D%5Bactivite%5D=0&tx_tdbpip_tdbpip%5BrechercheInfo%5D%5Bavancement%5D=0&tx_tdbpip_tdbpip%5BrechercheInfo%5D%5BmotsCles%5D=Bleue&tx_tdbpip_tdbpip%5BrechercheInfo%5D%5Bnouveau%5D=0&tx_tdbpip_tdbpip%5BrechercheInfo%5D%5Bnumero%5D=0)

[5D=&tx\\_tdbpip\\_tdbpip%5BrechercheInfo%5D%5Bregion%5D=6&cHash=dee606a97b3bd3dd65d16b6001ca304d](#)

Seo, K., Salon, D., Kuby, M., & Golub, A. (2019) Hedonic modeling of commercial property values: distance decay from the links and nodes of rail and highway infrastructure. *Transportation (Amst)*. 46, 859–882

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin, Boston, USA.

Shelby, L. B., & Vaske, J. J. (2008) Understanding Meta-Analysis: A Review of the Methodological Literature. *Leisure Sciences*, 30, 96–110.

Shiferaw, A., Söderbom, M., Siba, E., & Alemu, G. (2015) Road Infrastructure and Enterprise Dynamics in Ethiopia. *The Journal of Development Studies* 51(11), 1541–1558.

Simeonova, E., Currie, J., Nilsson, P., & Walker, R. (2021). Congestion Pricing, Air Pollution, and Children's Health. *Journal of Human Resources*, 56, 971-96.

Slusky, D. J. G. (2017). Significant placebo results in difference-in-differences analysis: The case of the ACA's parental mandate. *Eastern Economic Journal*, 43, 580-603.

Société des transports de Montréal (STM). (2020). Métro.  
<http://www.stm.info/en/info/networks/metro>

Song, Y., Lee, K., Anderson, W. P., & Lakshmanan, T. R. (2012) Industrial agglomeration and transport accessibility in metropolitan Seoul. *Journal of Geographical Systems*, 14, 299–318.

Spruance, S. L., Reid, J. E., Grace, M., & Samore, M. (2004) Hazard ratio in clinical trials. *Antimicrobial agents and chemotherapy*, 48(8), 2787-2792.

Stata. (2022). *Logistic regression, reporting coefficients*. <https://www.stata.com/manuals/rlogit.pdf>

Stata. (2021) *Tests of proportional-hazards assumption after stcox*.  
<https://www.stata.com/manuals/ststcoxph-assumptiontests.pdf>

Statistics Canada. (Feb. 2021) *Census 2016*. <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Comprehensive.cfm>

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.

Stepanova, M., & Thomas, L. (2002) Survival analysis methods for personal loan data. *Operations Research*, 50(2), 277-289.

Sterne, J. A. C., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *The Stata Journal*, 4(2), 127-141.

Stinchcombe A. (1965) Social structures and organizations. In: *Handbook of Organizations*, eds J. G. March, pp.142–193. Rand McNally, Chicago.

Stinchcombe A. (1965) Social structures and organizations. In: *Handbook of Organizations*, eds J. G. March, pp.142–193. Rand McNally, Chicago, USA.

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.

- Stringham, M. (1982). Travel behavior associated with land uses adjacent to rapid transit stations. *ITE Journal*, 52(4): 18-22
- Sukaryavichute, E., Delmelle, E., & Hammelman, C. (2021). Opportunities and challenges for small businesses in new transit neighborhoods: Understanding impacts through in-depth interviews. *Regional Science Policy and Practice*, 1025–1041.
- Takeda, Y., Kajikawa, Y., Sakata, I., & Matsushima, K. (2008) An analysis of geographical agglomeration and modularized industrial networks in a regional cluster: A case study at Yamagata Prefecture in Japan. *Technovation* 28(8), 531–539.
- Tiranti, R. (Oct. 2021) La ligne rose mise de côté pour le REM de l'Est. *Journal Métro*. <https://journalmetro.com/local/2708319/la-ligne-rose-mise-de-cote-pour-le-rem-de-lest/>
- Tornabene, S., & Nilsson, I. (2021). Rail transit investments and economic development: Challenges for small businesses. *Journal of Transport Geography*, 94, 103087.
- Tveterås, R. & Eide, G. E. (2000). Survival of New Plants in Different Industry Environments in Norwegian Manufacturing: A Semi-Proportional Cox Model Approach. *Small Business Economics*, 14, 65–82.
- Untermann, R. (1984). Adapting neighborhoods for walking and bicycling. *Accommodating the pedestrian*. Van Nostrand Reinhold, New York, USA.
- van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, 66, 206-22.
- Varian, H. R. (2015). *Introduction à la microéconomie* (8<sup>e</sup> Ed.). De Boeck Supérieur.
- Verhetsel, A., Kessels, R., Goos, P., Zijlstra, T., Blomme, N., & Cant, J. (2015) Location of logistics companies: a stated preference study to disentangle the impact of accessibility. *Journal of Transport Geography*, 42, 110–121.
- von Thünen, J.H., 1826. *Der Isolierte Staaat in Beziehung auf Landschaft und Nationalökonomie*. Hamburg. Pergamon Press, Oxford (English translation by C.M. Wartenburg, von Thunen's Isolated State)
- Weber, A. (1909). *Über den Standort der Industrien*, Tübingen. Tübingen: J. C. B. Mohr. (English translation by C. Friedrich *Theory of the Location of Industries*. Chicago: University of Chicago Press.)
- Weisbrod, G., Duncan, C., & Moses, S. J. (2014). Evolving Connection of Transit, Agglomeration, and Growth of High-Technology Business Clusters. *Transportation Research Record: Journal of the Transportation Research Board*, 2452(1), 11–18.
- Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, 40:1, 7-20.
- Wong, P. K., & He, Z. L. (2002). The impacts of knowledge interaction with manufacturing clients on KIBS firms innovation behaviour. *UNU-WIDER*, 2002(69), 1–20.
- Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. [https://harris.uchicago.edu/files/wooldridge\\_ppe\\_seminar\\_paper\\_6-1-22\\_0.pdf](https://harris.uchicago.edu/files/wooldridge_ppe_seminar_paper_6-1-22_0.pdf)

- Wu, K., Wang, Y., Ye, Y., Zhang, H., & Huang, G. (2019). Relationship Between the Built Environment and the Location Choice of High-Tech Firms: Evidence from the Pearl River Delta. *Sustainability*, 11(13), 1–21.
- Xu, T., Zhang, M., & Aditjandra, P. T. (2016) The impact of urban rail transit on commercial property value: New evidence from Wuhan, China. *Transportation Research Part A: Policy and Practice*, 91, 223-235.
- Yao, L., & Hu, Y. (2020) The impact of urban transit on nearby startup firms: Evidence from Hangzhou, China. *Habitat International*, 99, 1–10.
- Yi, C., & Nam, J. (2019) Interaction effects between the agglomerative externalities affecting the survival of small startups. *Growth and Change*, 50(4), 1311-1337.
- Zhang, M., & Xu, T. (2017) Uncovering the potential for value capture from rail transit services. *J Urban Plan Dev.*, 143, 04017006
- Zhong, N., Cao, J., & Wang, Y. (2017). Traffic congestion, ambient air pollution, and health: evidence from driving restrictions in Beijing. *Journal of the Association of Environmental and Resource Economists*, 4(3), 821–856.

## Annexe du Chapitre 1

**Table A1.1 - Selected papers used in the systematic review and meta-analysis**

1	Alañón-Pardo, Á., and Arauzo-Carod, J.-M. (2013) Agglomeration, accessibility and industrial location: Evidence from Spain. <i>Entrepreneurship &amp; Regional Development</i> 25(3–4), 135–173.
2	Beckerich, C., Benoit, S., and Delaplace, M. (2019) Are the reasons for companies to locate around central versus peripheral high-speed rail stations different? The cases of Reims central station and Champagne-Ardenne station. <i>European Planning Studies</i> 27(3), 574–594.
3	Bhat, C. R., Paleti, R., and Singh, P. (2014) A Spatial Multivariate Count Model for Firm Location Decisions: Firm Location Decisions. <i>Journal of Regional Science</i> 54(3), 462–502.
4	Bodenmann, B. R., and Axhausen, K. W. (2012) Effects and Side Effects of Measures to Attract Firms—A Micro-Simulation Study of Firm Location Choice. <i>disP—The Planning Review</i> 48(4), 14–28.
5	Castillo-Manzano, J. I., and López-Valpuesta, L. (2009) Urban retail fabric and the metro: A complex relationship. Lessons from middle-sized Spanish cities. <i>Cities</i> 26(3), 141–147.
6	Chang, T., Yang, D., Yang, Y., Huo, J., Wang, G., and Xiong, C. (2018) Impact of Urban Rail Transit on Business Districts Based on Time Distance: Urumqi Light Rail. <i>Journal of Urban Planning and Development</i> 144(3), 1–14.
7	Chatman, D. G., Noland, R. B., and Klein, N.J. (2016) Firm Births, Access to Transit, and Agglomeration in Portland, Oregon, and Dallas, Texas. <i>Transportation Research Record: Journal of the Transportation Research Board</i> 2598(1), 1–10.
8	Coughlin, C. C., and Segev, E. (2000) Location Determinants of New Foreign-Owned Manufacturing Plants. <i>Journal of Regional Science</i> 40(2), 323–351.
9	Credit, K. (2019) Transitive properties: a spatial econometric analysis of new business creation around transit. <i>Spatial Economic Analysis</i> 14(1), 26–52.
10	Curtis, C., and Mellor, R. (2011) Anticipating a New Railway: The Relevance of Transit-Oriented Development to Businesses in a Car-Oriented City. <i>Urban Policy and Research</i> 29(2), 141–165.
11	de Bok, M. (2009) Estimation and validation of a microscopic model for spatial economic effects of transport infrastructure. <i>Transportation Research Part A: Policy and Practice</i> 43(1), 44–59.
12	Forkenbrock, D. J., and Foster, N., S. J. (1996) Highways and Business Location Decisions. <i>Economic Development Quarterly</i> 10(3), 239–248.
13	Ghani, E., Goswami, A. G., and Kerr, W. R. (2016) Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. <i>The Economic Journal</i> 126(591), 317–357.
14	Gingerich, K., and Maoh, H. (2019) The role of airport proximity on warehouse location and associated truck trips: Evidence from Toronto, Ontario. <i>Journal of Transport Geography</i> 74, 97–109.
15	Gómez-Antonio, M., and Sweeney, S. (2018) Firm location, interaction, and local characteristics: A case study for Madrid’s electronics sector: Firms location choice interaction and local characteristics. <i>Papers in Regional Science</i> 97(3), 663–685.
16	Holl, A. (2004a) Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. <i>Regional Science and Urban Economics</i> 34(3), 341–363.
17	Holl, A. (2004b) startups and relocations: Manufacturing plant location in Portugal. <i>Papers in Regional Science</i> 83(4), 649–668.
18	Holl, A. (2004c) The role of transport in firms’ spatial organization: evidence from the Spanish food processing industry. <i>European Planning Studies</i> 12(4), 537–550.

19	Holl, A. (2004d) Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal. <i>Journal of Regional Science</i> 44(4), 693–712.
20	Holl, A., and Mariotti, I. (2018) The Geography of Logistics Firm Location: The Role of Accessibility. <i>Networks and Spatial Economics</i> 18(2), 337–361.
21	Hong, J. (2007a) Location Determinants and Patterns of Foreign Logistics Services in Shanghai, China. <i>The Service Industries Journal</i> 27(4), 339–354.
22	Hong, J. (2007b) Transport and the location of foreign logistics firms: The Chinese experience. <i>Transportation Research Part A: Policy and Practice</i> 41(6), 597–609.
23	Hong, J., and Chin, A.T.H. (2007) Modeling the location choices of foreign investments in Chinese logistics industry. <i>China Economic Review</i> 18(4), 425–437.
24	Hurst, N.B., and West, S. E. (2014) Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. <i>Regional Science and Urban Economics</i> 46, 57–72.
25	Iseki, H., and Eom, H. (2019) Impacts of Rail Transit Accessibility on Firm Spatial Distribution: Case Study in the Metropolitan Area of Washington, DC. <i>Transportation Research Record: Journal of the Transportation Research Board</i> 2673(11), 220–232.
26	Iseki, H., and Jones, R. P. (2018) Analysis of firm location and relocation in relation to Maryland and Washington, DC metro rail stations. <i>Research in Transportation Economics</i> 67, 29–43.
27	Jiang, Y., Ng, A.K.Y., Wang, Y., Wang, L., and Yu, B. (2018a) Locational Characteristics of Firms in the Business Service Industry in Airport Economic Zones: Case of Shanghai Hongqiao International Airport. <i>Journal of Urban Planning and Development</i> 144(1), 1–13.
28	Jiang, H., Xu, W., and Zhang, W. (2018b) Transportation Accessibility and Location Choice of Japanese-Funded Electronic Information Manufacturing Firms in Shanghai. <i>Sustainability</i> 10(2), 1–21.
29	Kang, S. (2020) Warehouse location choice: A case study in Los Angeles, CA. <i>Journal of Transport Geography</i> 88, 1–11.
30	Kang, S. J., and Lee, H. S. (2007) The determinants of location choice of South Korean FDI in China. <i>Japan and the World Economy</i> 19(4), 441–460.
31	Kim, H., Ahn, S., and Ulfarsson, G. F. (2018) Transportation infrastructure investment and the location of new manufacturing around South Korea’s West Coast Expressway. <i>Transport Policy</i> 66, 146–154.
32	Leitham, S., McQuaid, R.W., and Nelson, J. D. (2000) The influence of transport on industrial location choice: a stated preference experiment. <i>Transportation Research Part A</i> 34, 515–535.
33	López, F. A., and Páez, A. (2017) Spatial clustering of high-tech manufacturing and knowledge-intensive service firms in the Greater Toronto Area: Spatial clustering in the GTA. <i>The Canadian Geographer/Le Géographe canadien</i> 61(2), 240–252.
34	Maoh, H. and Kanaroglou, P. (2009) Intrametropolitan Location of Business Establishments: Microanalytical Model for Hamilton, Ontario, Canada. <i>Transportation Research Record: Journal of the Transportation Research Board</i> 2133(1), 33–45.
35	Matas, A., Raymond, J. L., and Roig, J. L. (2020) Evaluating the impacts of HSR stations on the creation of firms. <i>Transport Policy</i> 99, 396–404.
36	Mayer, T., and Trevien, C. (2017) The impact of urban public transportation evidence from the Paris region. <i>Journal of Urban Economics</i> 102, 1–21.
37	Mejia-Dorantes, L., Paez, A., and Vassallo, J. M. (2012) Transportation infrastructure impacts on firm location: the effect of a new metro line in the suburbs of Madrid. <i>Journal of Transport Geography</i> 22, 236–250.

38	Melo, P.C., Graham, D.J., and Noland, R.B. (2010) Impact of Transport Infrastructure on Firm Formation: Evidence from Portuguese Municipalities. <i>Transportation Research Record: Journal of the Transportation Research Board</i> 2163(1), 133–143.
39	Nilsson, I. M., and Smirnov, O. A. (2016) Measuring the effect of transportation infrastructure on retail firm co-location patterns. <i>Journal of Transport Geography</i> 51, 110–118.
40	O'Connor, K. (1989) Australian ports, metropolitan areas and trade-related services. <i>Australian Geographer</i> 20(2), 167–172.
41	Peeters, D., Thisse, J.-F., and Thomas, I. (2000) On High-Speed Connections and the Location of Activities <i>Environment and Planning A: Economy and Space</i> 32(12), 2097–2112.
42	Percoco, M. (2016) Highways, local economic structure and urban development. <i>Journal of Economic Geography</i> 16(5), 1035–1054.
43	Petráš, M., and Květoň, V. (2020) Spatial changes in distribution of firms and selected industries around new motorway. <i>Applied Geography</i> 121, 1–15.
44	Ramaul, N. K., and Ramaul, P. (2016) Determinants of Industrial Location Choice in India: A Polychoric Principal Component Analysis Approach. <i>Journal of Quantitative Economics</i> 14(1), 29–56.
45	Ramón-solans Prat, J.C., and Marcén, R.F. (2006) Influential Factors in Location Choice of Spanish Businesses in Aragon. <i>The Journal of Entrepreneurship</i> 15(1), 63–81.
46	Scoppa, M. D., and Peponis, J. (2015) Distributed Attraction: The Effects of Street Network Connectivity Upon the Distribution of Retail Frontage in the City of Buenos Aires. <i>Environment and Planning B: Planning and Design</i> 42(2), 354–378.
47	Shiferaw, A., Söderbom, M., Siba, E., and Alemu, G. (2015) Road Infrastructure and Enterprise Dynamics in Ethiopia. <i>The Journal of Development Studies</i> 51(11), 1541–1558.
48	Song, Y., Lee, K., Anderson, W. P., and Lakshmanan, T. R. (2012) Industrial agglomeration and transport accessibility in metropolitan Seoul. <i>Journal of Geographical Systems</i> 14, 299–318.
49	Takeda, Y., Kajikawa, Y., Sakata, I., and Matsushima, K. (2008) An analysis of geographical agglomeration and modularized industrial networks in a regional cluster: A case study at Yamagata Prefecture in Japan. <i>Technovation</i> 28(8), 531–539.
50	Verhetsel, A., Kessels, R., Goos, P., Zijlstra, T., Blomme, N., and Cant, J. (2015) Location of logistics companies: a stated preference study to disentangle the impact of accessibility. <i>Journal of Transport Geography</i> 42, 110–121.
51	Wu, K., Wang, Y., Ye, Y., Zhang, H., and Huang, G. (2019) Relationship Between the Built Environment and the Location Choice of High-Tech Firms: Evidence from the Pearl River Delta. <i>Sustainability</i> 11(13), 1–21.
52	Yao, L., and Hu, Y. (2020) The impact of urban transit on nearby startup firms: Evidence from Hangzhou, China. <i>Habitat International</i> 99, 1–10.

**Table A1.2 - Main conclusions of the 52 selected papers**

<b>Reference</b>	<b>Main Approach; Data</b>	<b>Transport Infrastructure</b>	<b>Main Conclusions</b>
Alañón-Pardo and Arauzo-Carod, 2013	Count model (NB); CSp	Highway	These results mean that the higher the accessibility is, the higher the location of new manufacturing plants will be. [...] The results of accessibility clearly show that a road connection between cities plays a key role in location decisions in all industries (p. 165).
Beckerich et al., 2019	Descriptive analysis; StP	Railway	The most commonly self-reported factor was office availability (20.23% for the central station area and 14.93% for the peripheral station area). [...] While HSR is not very important for businesses in the peripheral district (5.97%), it is the second most cited factor in central district (11.90%) (p. 588).
Bhat et al., 2014	Spatial Durbin model; CS	Highway; Airport; Port	[T]he industrial specialization and human capital considerations, according to our analysis, influence the count of new businesses solely through the demand intensity, while the transportation infrastructure and land development characteristics impact the count of new businesses solely through the supply function (p. 499).
Bodenmann and Axhausen, 2012	Nested Logit; CSp	Highway	[M]unicipalities along this new highway are impacted significantly. The positive effects are even further enhanced due to direct access to the motorway [...] Interventions regarding transport infrastructure have a larger impact in densely populated areas. Negative effects can be shown for areas along existing important traffic corridors [...] but also in the hinterland of municipalities directly concerned (p. 23).
Castillo-Manzano and López-Valpuesta, 2009	Censored Regression model; StP	PT	[W]hen the Metro comes into operation it becomes an important instrument for revitalizing retail commerce in historical city districts as it improves access to these from peripheral neighborhoods [...] [T]his explains why it is those retailers who are nearest the city center who most value the Metro (p. 146).
Chang et al., 2018	Metacartography ; CS	PT	Business districts are centered in the downtown area and distributed in circular shapes. The central business districts will be strengthened. Downtown business districts still have a regional advantage [...] The districts on the edge will produce a centripetal force because of a shortened time distance. Light rail construction will greatly promote the development of business districts on the edge of business districts (p. 11).
Chatman et al., 2016	Count model (NB); LD	PT	[T]his study does not find much evidence that rail influences firm births through agglomeration. The mechanism at work here may instead simply be accessibility to the larger metropolitan area (p. 8).



Coughlin and Segev, 2000	Count model (NB); CS	Highway	Manufacturing agglomeration, interstate highways, population and educational level were all positively related to the number of new plants. Examining the mean values for these variables, one sees that, except for manufacturing agglomeration, the urban value exceeds the rural value, placing urban counties at a relative advantage (p. 348).
Credit, 2019	Spatial Durbin model; CS	PT	[T]he results show that proximity to rail transit stations does, in fact, have a positive overall relationship with new business starts, even while controlling for several forms of spatial dependence, total existing and new business activity in the block, and other socio-demographic factors. [...] New retail, services, and food business and knowledge businesses are most consistently associated with rail transit variables (p. 47).
Curtis and Mellor, 2011	Descriptive analysis; StP	Railway	In Wellard [outer suburban metropolitan area – 39 km to CBD], 54 per cent of responding businesses indicated that they chose their business location because the new railway would grow business opportunities, compared to 37 per cent of businesses in Cockburn Central [suburban metropolitan area – 21 km to CBD] and only 14 per cent in Bull Creek [low-density area – 12 km to CBD] (p. 157).
De Bok, 2009	Simulation; LD	Highway; Railway	Highway proximity was significant for industries that are more likely to dependent on road networks for their business activities, such as business services, construction, logistics, financial sectors. Train station proximity is preferred by firms in the finance, education and restaurants [...] The agglomeration of firms from the same industry sectors is significant in some specific industry sectors: finance, business services, manufacturing, construction and logistics (p. 50).
Forkenbrock and Foster, 1996	Descriptive analysis; StP	Highway	Labor quality was the most important factor, followed by labor costs. Transportation services ranked third in Missouri (44 of 235 respondents ranked it first) and fifth in Iowa (7 of 219 respondents ranked it first). The Iowa results are particularly consistent with the research cited earlier: Proximity to markets and materials is more important than transportation per se (p. 244).
Ghani et al., 2016	OLS; StP	Highway	[W]eak effects accrued to districts 10–50 kilometres from the GQ network, which accounted for a little under 10% of initial output levels. [...] The biggest growth comes from the 0 to 10 kilometres districts, which accounted for 34% of initial levels [...] This would suggest a 17% output growth contribution overall for India’s manufacturing sector (p.354).
Gingerich and Maoh, 2019	Multinomial Logit; CS	Highway; Railway; Airport	A multinomial logit model confirmed that access to Pearson International Airport and the highway infrastructure are both significant factors in the location of Toronto based warehouses (p. 109).
Gómez-Antonio and Sweeney, 2018	Gibbs model; CS	Highway; PT; Airport; Railway	Firms prefer sites that are in proximity to train stations or metro transportation systems. Even large firms that locate in sites far from the city which might not be covered by the metro network are always in proximity to ‘local train stations.’ There is evidence of a significant preference for locations in proximity to Madrid International Airport and this is particularly the case for larger firms (p. 679).

Holl and Mariotti, 2018	Count model (Poisson, NB & Zero Inflated NB); CSp	Highways, Airport, Port	Access to transportation infrastructure and proximity to urban areas (especially Madrid and Barcelona) are important location determinants for logistics firms, thus confirming the attitude of this sector, which is market oriented, to be located in or close to the consumer market. Access to efficient transportation infrastructure is important as transportation speeds and costs are crucial determinants of efficient logistics (p. 355).
Holl, 2004a	Count model (Poisson); CSp	Highway	Manufacturing plant openings drop in municipalities outside the 10-km corridors relative to municipalities within the first 10 km from the motorways as a result of the motorway construction. Compared with the first 10 km, being outside the 10-km transport corridors reduces the expected number of new manufacturing plant openings by 12– 94%, depending on the distance to the new infrastructure and the sector (pp. 355–357).
Holl, 2004b	Count model (Pooled NB); CSp	Highway	[T]he negative signs of the motorway dummy variables indicate a clear preference for locating in the first 10 kilometers around the new motorways. Firm birth drops in 15 sectors in municipalities outside the 10-kilometer corridors relative to municipalities within the first 10 kilometers. [...] This suggests that the positive effects of transport investment are concentrated near the new corridors, and that by attracting firm birth into the new corridors, transport infrastructure investment can lead to negative spillover effects in adjacent areas (p. 704).
Holl, 2004c	Descriptive analysis; StP	Highway	First, roads matter for the interviewed firms, whereas other transport modes are only used in exceptional cases. As a location factor, transport costs and access to the primary road network are among the most frequently mentioned factors in having influenced the location decision of the interviewed companies in the NIV [highway] corridor (p. 548).
Holl, 2004d	Count model (Poisson & NB); CSp	Highway	The coefficient estimates predict that municipalities outside the 10-kilometre motorway corridors show between 11 and 31% less predicted manufacturing plant births [...] The negative distance effects are, in general, most strongly felt in municipalities between 20 and 30 kilometres from the nearest motorway [...] These results [...] suggest negative spillover effects in the sense that new motorways increase the attractiveness of municipalities in the new corridors at the cost of adjacent municipalities (p. 663).
Hong and Chin, 2007	Nested Logit; TS	Highway; Railway; Port	The transportation network capacity in terms of seaway, roadway and railway also played an important role in attracting FDI. Foreign logistics investors preferred cities with existing concentrations of logistics firms (p. 436).
Hong, 2007a	Nested Logit; CSp	Highway, Railway, Airport, Port	[Results] show that accessibility to transportation infrastructure is a significant attraction to foreign logistics investors. The coefficient estimate of [Port Dummy] ranges from 0.99 to 2.57, suggesting that proximity to a seaport is an important consideration when foreign logistics firms make their location choices. Accessibility to an airport also attracts logistics establishments, as indicated by the coefficient estimates for [Distance to Airport]. The above estimation results are consistent with our observation that many logistics establishments cluster near transport nodes (p. 347).

Hong, 2007b	Conditional Logit; CS	Highway, Railway, Port	[T]he empirical results in this paper have confirmed the importance of transport in attracting foreign logistics investment. In addition, while the impact of seaway and railway transport was quite consistent across sub-samples, different groups of firms might have quite different preferences on roadway transport infrastructure. [...] We also noted that the importance of roadway transport increased over time (p. 608).
Hurst and West, 2014	Logit; LD	PT	Results show that properties within ½ mile of operational LRT stations experience a small increase in the likelihood of land use change relative to when the LRT is under construction [...] Within the corridor, proximity to LRT increases the likelihood of land use change on single-family and industrial properties, but appears to have no effect on vacant land, commercial properties, and multi-family properties (p. 57).
Iseki and Eom, 2019	Multinomial Logit; TS	Highway; PT; Airport	The [Multinomial Logit] analysis results showed that metro station proximity increases the likelihood that a firm belongs to the industry [Information, Professional and Management; FIRE; Education and Health; Arts, Accommodation and Food; Public Administration] especially within 5 km of Metrorail stations even after controlling for land-use variables, distances to highways and airports, and other influential variables. This information, professional and management industry was found to have the second highest clustering tendency around Metrorail stations, following public administration (p. 230).
Iseki and Jones, 2018	Descriptive Analysis; LD	Highway; PT; Airport	Our analysis showed that the levels of ability to attract firms and of firm concentration substantially vary by jurisdiction, as well as by station area, especially in relation to the suburbanization of firm locations [...] Such factors include access to transportation facilities (highways, airports, intercity train stations, as well as metro stations), access to potential employees and customers, rents, land use regulations, planning process related to station area development, and policies of local governments (p. 42).
Jiang et al., 2018a	Count model (NB); CS	Airport	The results show that airport travel time, city center travel time, location in the industrial park or not, and intra-agglomeration are important factors in the locational decisions of both domestic and foreign firms. However, the travel time of both the airport and city center are the foremost locational factors of domestic firms, while the travel time of the city center is the crucial locational factor of foreign firms (pp. 11–12).
Jiang et al., 2018b	Count model (Poisson); CS <sub>p</sub>	Railway; PT; Airport; Port	[F]irms tend to select the areas with better accessibility to transportation facilities. [...] [W]e identified two types of spatial coupling models between transportation facility accessibility and firm location. The first, represented by the airports and railway stations, shows an inverse U-shaped curve relationship. The second type is associated with accessibility to the subway stations and the ports, and the spatial coupling relationship degree between transportation facility accessibility and firm distribution is of negative linearity (p. 18).

Kang, 2020	Multinomial Logit; CSp	Airport; Port	Most warehouses built in the early period (1951–1980) are located around the 20-mile band from the CBD (traditional industrial zones around the central urban areas and the vicinity of the port complex), whereas the recent development has been around the 40- and 60-mile bands from the CBD (p. 9).
Kang and Lee, 2007	Conditional Logit; CSp	Highway; Railway; Port	Various indicators have been used for transport infrastructures in many FDI studies. They are the lengths of paved roads, the lengths of inland waterways, and the lengths of highways or railways. We tried these three variables and found that the length of railways per km <sup>2</sup> is quite robust [...] [T]he coefficient for the railroad variable is robust to various combinations of transport infrastructure variables. Even with or without other transport infrastructure variables, the inland waterways per km <sup>2</sup> is negative and significant (p. 457).
Kim et al., 2018	Count model (NB); LD	Highway	When the distance from a West Coast Expressway IC [Highway] increased by 1 km, the expected number of new manufacturing establishments decreased by 0.26%. This indicates that the investment in new transport infrastructure changes the new manufacturing establishment's preference for location and consequently has a significant impact on the spatial distribution of economic activity. Such an explanation is possible because this analysis effectively controls for the potential endogeneity and simultaneity between the location of new manufacturing establishments and the new investment in transport infrastructure (p. 151).
Leitham et al., 2000	Logit; StP	Highway; PT	UK-sourced branch plants were found to rate motorway links the highest of any of the groups of firms, underlining the dominance of road links for the distribution of goods and services in the UK. In contrast, overseas sourced branch firms found road links to be an unimportant factor, being outweighed primarily by consideration of workforce and premises [...] Good public transport did emerge as a statistically significant factor for larger intra-regional locating firms (p. 532).
López and Páez, 2017	Logit; CS	Highway; Railway; Airport	[P]roximity to transportation infrastructure has the expected effect. As distance from the nearest transportation facility increases, the probability that a firm will be a member of a cluster decreases [...] At distances beyond 1 km, the effect has already become vanishingly small [...] [T]he probability of being a member of a cluster decays more rapidly with distance in the case of KIS [knowledge-Intensive Services] than manufacturing, which is consistent with the notion that transportation accessibility is more highly priced by the service sector (p. 250).
Maoh and Kanaroglou, 2009	Multinomial Logit; CS	Highway	The [results] suggest that manufacturing, communication, transportation, and wholesale trade firms are strongly affiliated with land in close proximity to main roads and highways. The same could be said about new firms that belong to health and social services, and accommodation, food, and beverage industries [...] In contrast, construction firms do not show preference to locate close to main roads and highways, except for construction firms specializing in electrical work (p. 37).

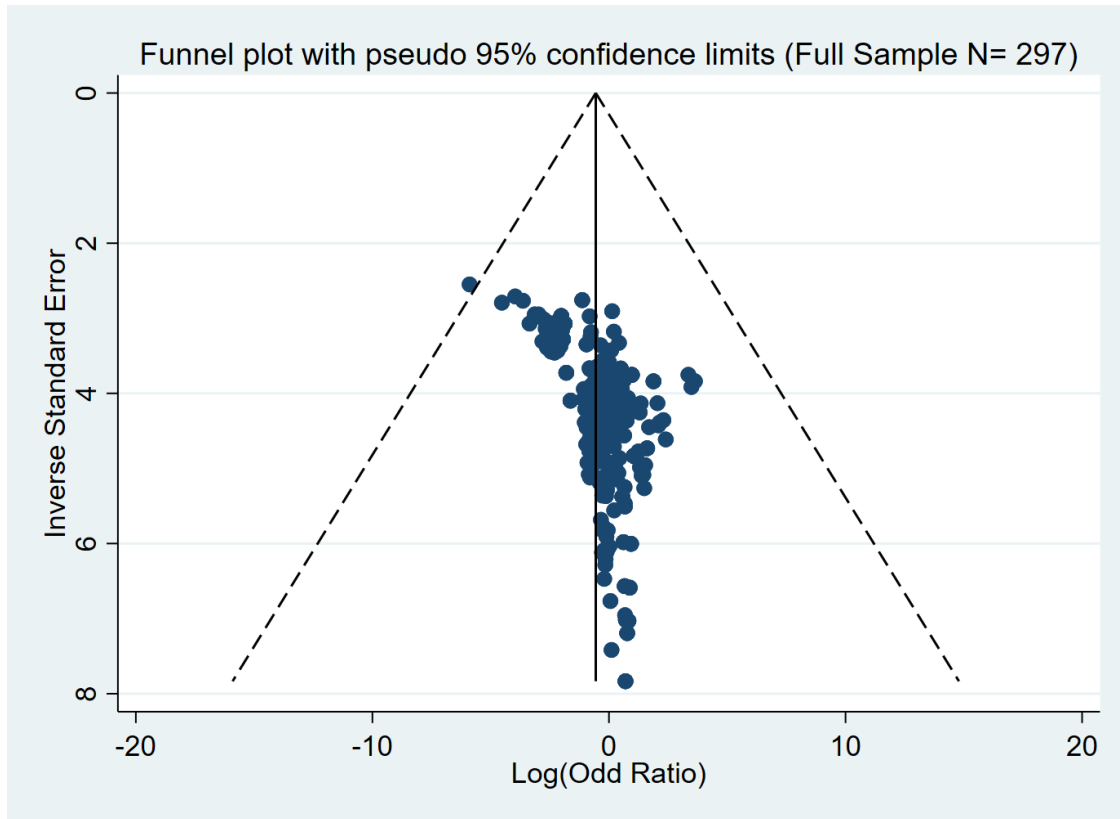
Matas et al., 2020	Non-Linear model; LD	Railway	The results point out that being connected to an HSR system may boost firm creation, but this is not always the case [...] Connection to the HSR network leads to higher growth in firm creation in sectors where face-to-face relations are more important. This is the case for service, tourism and knowledge-intensive activities. Furthermore, the magnitude of the effect depends on the province characteristics. Hence, once again, the evidence confirms that the results need to be contextualized (p. 403).
Mayer and Trevien, 2017	OLS; CSp	PT	[T]he positive effect of the RER [Railway] is also found for the number of firms [...] There are striking differences between domestic and foreign-owned firms. First, the RER treatment effect is much larger for the latter set of firms. The number of foreign-owned firms is 20.2% higher in treated municipalities with a one-minute drop in travel time: public transport affects foreign firms more than domestic firms. Second, access to the city center also appears to play a major role in the location decisions of foreign investors (the dummies for distance to Paris exhibit a clear core-oriented pattern) (p. 9).
Mejia-Dorantes et al., 2012	Multinomial Logit; LD	Highway; PT	The evidence is supported by means of a multinomial logistic regression, where we investigate how businesses by group sector [...] are related to urban characteristics, such as distance to different public transport infrastructure; population and street network densities; local CBD and a ratio of business occupation around each location point in 1998. Our estimates show that firms benefit from neighborhood characteristics. We conclude that the changes in accessibility related to Metrosur affected the locational pattern and generated a tendency to locate near Metro stations (pp. 247-248).
Melo et al., 2010	Count model (Pooled—Poisson & NB); LD	Highway; Railway	Expanding the local municipality railway and motorway networks by 10% can produce an increase in new plant openings from 0.9% to 2.7% and 0.7% to 2.6%, respectively. Positive spatial spillover effects from the transport infrastructure were also found: improving the motorway and railway networks of neighboring municipalities by 10% increases plant openings by [...] 2.7% to 5.1% and 2.4% to 6.7%, respectively (p. 133).
Nilsson and Smirnov, 2016	Multivariate Logit; CS	Highway	[T]he results indicate that access to common transportation infrastructure induces firms to locate in close proximity to one another; such interactions may not have been present in the absence of the transportation network. Many of the retail chains that displayed significant co-location tendencies in the overall population, and among outlets located in close proximity to important road transportation infrastructure, show no such tendencies among the outlets with no such access (p. 116).
O'Connor, 1989	Descriptive analysis; StP	Port	Data assembled on the location of these services in Australia [...] indicates their location is influenced by networks and linkages between firms; services are not necessarily located in the busiest port, but rather in the city with the most diverse service sector. This result illustrates the selectivity in location that is typical of producer services, and has important consequences for the role and influence of cities (p. 167).

Peeters et al., 2000	Simulation; CS	Railway	These results lead us to speculate that the construction of a high-speed [railway] connection should favor the regional center that has some comparative advantage [...] In addition, the lure of the primary regional centers should also be strengthened at the expense of more peripheral, secondary centers because the market area of firms located therein expands (p. 2111).
Percoco, 2016	OLS; LD	Highway	This result points at a confirmation of the positive impact of highways on local economic growth in an order of magnitude of 4–5% for employment growth and 2–3% for plants (p. 1050).
Petráš and Květoň, 2020	Descriptive analysis; CSp	Highway	The result show[s] that municipalities around new motorway have more positive development in the firm birth intensity also compared to municipalities on the original route. The situation on these old routes varies; however, both of these old routes have experienced decrease in manufacturing, transport and logistics firm birth intensity that is bigger than in other parts of studied region (p. 10).
Ramaul and Ramaul, 2016	Revealed Preference; StP	Highway; Railway; Airport	[T]ransport and utilities were rated much lower in terms of importance in the critical factors. The low importance to transportation may be explained by the low importance rating of airway and railway facility by the firms (p. 37).
Ramón-solans Prat and Marcén, 2006	Descriptive analysis; StP	Highway; Railway; Airport	The results of the frequency analysis led to the conclusion that location choices were influenced by factors such as good accessibility through an efficient road network, demand in the area, qualified soil and quality and availability of productive factors. Similar results were obtained by the comparative analysis of hierarchization (p. 63).
Scoppa and Peponis, 2015	OLS; CS	Highway; Railway; PT; Airport; Port	The land uses immediately associated with a street segment have a less significant effect on the distribution of movement. Thus, while shops located on a given street segment are likely to attract some movement that would otherwise not be directed there, they are more likely to take advantage of passing movement that is independent of their presence (p. 376).
Shiferaw et al., 2015	GMM; LD	Highway	Our results indicate that better road accessibility is associated with a town's desirability for manufacturing firms. System GMM results and estimates from regressions specified in long differences suggest that improved local road infrastructure—measured in our data set as the total distance that can be travelled during a 60-minute drive from the centre of town and the total area accessible during those 60 minutes of driving—is associated with a larger number of firms in that locality (p. 1556).
Song et al., 2012	Logit; CSp	Highway; PT	The effect of transport networks [...] captures the change of probability of a village being in an agglomeration when a regressor increases by one percent and other regressors stay at the mean. [...] None of the transport variables is found to have significant marginal effects in primary, manufacturing, sales, finance, and business activities sectors. Real estate is the only industry with significantly positive marginal effect from road density, which is only 0.27%. Distance to highway shows quite similar magnitude of marginal effects ranging [from] 0.3 to -1.4% where it is significant (p. 312).

Takeda et al., 2008	Network Analysis; CS	Highway; Railway	The first point indicates that infrastructure for transportation still plays a decisive role in terms of site location. The second point indicates that there is a tendency for firms to transact with those in close proximity, and it is driven by the importance of transaction costs or the importance of face-to-face communication with suppliers and customers (p. 538).
Verhetsel et al., 2015	Mixed Multinomial Logit; StP	Highway; Railway; Port	An analysis of the resulting data by means of discrete choice modeling revealed that land rent is the most important factor in the location choice of logistics companies in Flanders. Access to a port is the second most important factor, followed by access to a motorway, the location in a business park and an inland navigation terminal, which are all about equally important. Access to a rail terminal plays no significant role in the location choice of logistics companies in Flanders (p. 110).
Wu et al., 2019	Spatial Econometrics (SEM; SLM); CS	Highway	[T]he results of the spatial regression model demonstrate a significant relationship between the [Transport Infrastructure Level] and the distribution of high-tech firms [...] Similar findings were found in earlier studies, which revealed positive impacts of the [Transport Infrastructure Level] or accessibility condition on the agglomeration of high-technology firms and manufacturing companies [...] The [Transport Infrastructure Level] is always an important factor for determining the location of firms. Even with well-developed information and communication technology (ICT), in the contemporary world, high-tech firms still prefer locations with good accessibility conditions (p. 14).
Yao and Hu, 2020	Count model (Poisson); LD	PT	[W]e show that in total the transit areas experienced about 28% more startups than the controlled ones. Specifically, the numbers are most significant for retail, construction, and technology firms with about 42%, 30.9%, and 29% gain respectively. [...] We find that the effect is mostly on firms with larger registered capital — 64.5% for retail and 40.9% for technology firms. In contrast, we find little evidence of influence on manufacturing firms and firms with small registered capital (p. 9).

Legend (Data): CS, Cross-Sectional Data. CSp, Cross-Sectional Pooled. LD, Longitudinal Data. StP, Stated Preferences (Surveys). TS, Time Series. The data were defined based on Dubé and Legros (2014). Legend (Transport Infrastructure): PT, Public Transit.

**Figure A1.3 - Evaluating the potential publication bias**





**Table A1.4.1 - Ordered-logit regression, pairing negative and significant effects with non-significant effects.**

Variable		Odd ratio	Std. Error	sign.
Sector	Manufacturing	Reference	--	
	Agriculture and Mining	1.80e-8	0.0000	
	Construction	1.3203	0.8587	
	Logistics	2.1590	1.2818	
	Wholesale	1.2325	0.5059	
	Retail	1.1636	0.6950	
	Business Services	0.7989	0.4924	
	FIRE	3.7525	2.4892	**
	KIBS	3.9814	3.7984	
	Education/Health/Food and Accommodation/Arts and Entertainment	1.7787	0.9082	
	Public Administration	1.1209	1.1699	
Transport	Highway	Reference	--	
	Railway	3.1543	1.5683	**
	Airport	1.0839	0.6391	
	Port	1.90e7	4.27e10	
	Public Transit	1.5890	0.7180	
Location	Other	Reference	--	
	United States	0.0478	0.0533	***
	Europe	0.1657	0.1597	*
	Asia	0.1440	0.1277	**
Methodology	Other Models	Reference	--	
	Count Model	5.1623	3.014	***
	Logit Model	1.7443	1.1052	
	Descriptive Analysis	0.4885	0.4361	
	Linear Model	3.6624	5.1153	
	Spatial Econometric Model	0.3634	0.4629	
Data	Cross-Sectional	Reference	--	
	Pooled Cross-Sectional	0.0531	0.0630	**
	Longitudinal/Panel	0.0459	0.0540	***
	Stated Preferences	0.0762	0.1064	*
	Time Series	7482872	7.85e9	

The option [or] in Stata was used in order to obtain the odd ratio as coefficient.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

**Table A1.4.2 - Ordered-logit regression, using positive and non-significant effects  
(discarding the negative effect)**

Variable		Odd ratio	Std. Error	sign.
Sector	Manufacturing	Reference	--	
	Agriculture and Mining	1.33e-8	0.00003	
	Construction	1.2257	0.7967	
	Logistics	1.7982	1.0748	
	Wholesale	1.1014	0.7792	
	Retail	1.2639	0.7763	
	Business Services	0.8381	0.5288	
	FIRE	3.4631	2.3593	*
	KIBS	3.8357	3.6396	
	Education/Health/Food and Accommodation/Arts and Entertainment	1.7562	0.9165	
	Public Administration	1.0706	1.1172	
Transport	Highway	Reference	--	
	Railway	3.0121	1.4776	**
	Airport	1.1942	0.6880	
	Port	2.27e7	5.42e10	
	Public Transit	1.7175	0.8037	
Location	Other	Reference	--	
	United States	0.0314	0.0415	***
	Europe	0.1142	0.1388	*
	Asia	0.0670	0.0776	**
Methodology	Other Models	Reference	--	
	Count Model	4.9686	2.9357	***
	Logit Model	2.4327	1.6309	
	Descriptive Analysis	0.5068	0.5435	
	Linear Model	3.2439	4.8101	
	Spatial Econometric Model	0.7314	0.7874	
Data	Cross-Sectional	Reference	--	
	Pooled Cross-Sectional	0.1188	0.1153	**
	Longitudinal/Panel	0.0980	0.0928	**
	Stated Preferences	0.1669	0.2201	
	Time Series	1.23e7	1.39e10	

The option [or] in Stata was used in order to obtain the odd ratio as coefficient.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

**Table A1.5.1 - Meta-analysis results, using cross-product multiplication accounting for freight transportation**

Variable		Odd ratio	Std. Error	sign.
Sector	Manufacturing	Reference	--	
	Agriculture and Mining	0.209	0.151	**
	Construction	1.282	0.798	
	Logistics	2.146	1.226	
	Wholesale	1.262	0.854	
	Retail	1.192	0.983	
	Business Services	1.225	1.009	
	FIRE	4.297	3.793	*
	KIBS	5.116	5.529	
	Education/Health/Food and Accommodation/Arts and Entertainment	1.987	1.547	
Public Administration	1.144	1.154		
Transport	Highway	Reference	--	
	Railway	6.188	4.437	**
	Airport	0.889	0.722	
	Port	7.91e6	1.16e10	
	Public Transit	1.227	0.652	
Transport Cross Variable	Highway x Freight Transp.	1.475	1.145	
	Railway x Freight Transp.	0.434	0.436	
	Airport x Freight Transp.	1.816	2.153	
	Port x Freight Transp.		Omitted	
	Public Transit x Freight Transp.		Omitted	
Location	Other	Reference	--	
	United States	0.060	0.064	***
	Europe	0.191	0.184	*
	Asia	0.170	0.153	**
Methodology	Other Models	Reference	--	
	Count Model	5.425	3.031	***
	Logit Model	1.411	0.897	
	Descriptive Analysis	0.173	0.149	**
	Linear Model	2.922	3.978	
	Spatial Econometric Model	0.642	0.668	
Data	Cross-Sectional	Reference	--	
	Pooled Cross-Sectional	0.132	0.124	**
	Longitudinal/Panel	0.071	0.065	***
	Stated Preferences	0.382	0.447	
	Time Series	7.61e6	5.31e9	

The option [or] in Stata was used in order to obtain the odd ratio as coefficient.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

**Table A1.5.2 - Meta-analysis results, using cross-product multiplication accounting for passengers' transportation**

Variable		Odd ratio	Std. Error	sign.
Sector	Manufacturing	Reference	--	
	Agriculture and Mining	0.209	0.151	**
	Construction	1.282	0.798	
	Logistics	2.146	1.226	
	Wholesale	1.262	0.854	
	Retail	1.192	0.983	
	Business Services	1.225	1.009	
	FIRE	4.297	3.793	*
	KIBS	5.116	5.529	
	Education/Health/Food and Accommodation/Arts and Entertainment	1.987	1.547	
	Public Administration	1.144	1.154	
Transport	Highway	Reference	--	
	Railway	1.821	1.018	
	Airport	1.096	0.785	
	Port	5.37e6	7.85e9	
	Public Transit	0.832	0.529	
Transport Cross Variable	Highway x Passengers' Transp.	0.678	0.527	
	Railway x Passengers' Transp.	2.304	2.316	
	Airport x Passengers' Transp.	0.551	0.653	
	Port x Passengers' Transp.	Empty (no observation)		
	Public Transit x Passengers' Transp.	Omitted		
Location	Other	Reference	--	
	United States	0.060	0.064	***
	Europe	0.191	0.184	*
	Asia	0.170	0.153	**
Methodology	Other Models	Reference	--	
	Count Model	5.425	3.031	***
	Logit Model	1.411	0.897	
	Descriptive Analysis	0.173	0.149	**
	Linear Model	2.922	3.978	
	Spatial Econometric Model	0.642	0.668	
Data	Cross-Sectional	Reference	--	
	Pooled Cross-Sectional	0.132	0.124	**
	Longitudinal/Panel	0.071	0.065	***
	Stated Preferences	0.382	0.447	
	Time Series	7.61e6	5.31e9	

The option [or] in Stata was used in order to obtain the odd ratio as coefficient.  
 \*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

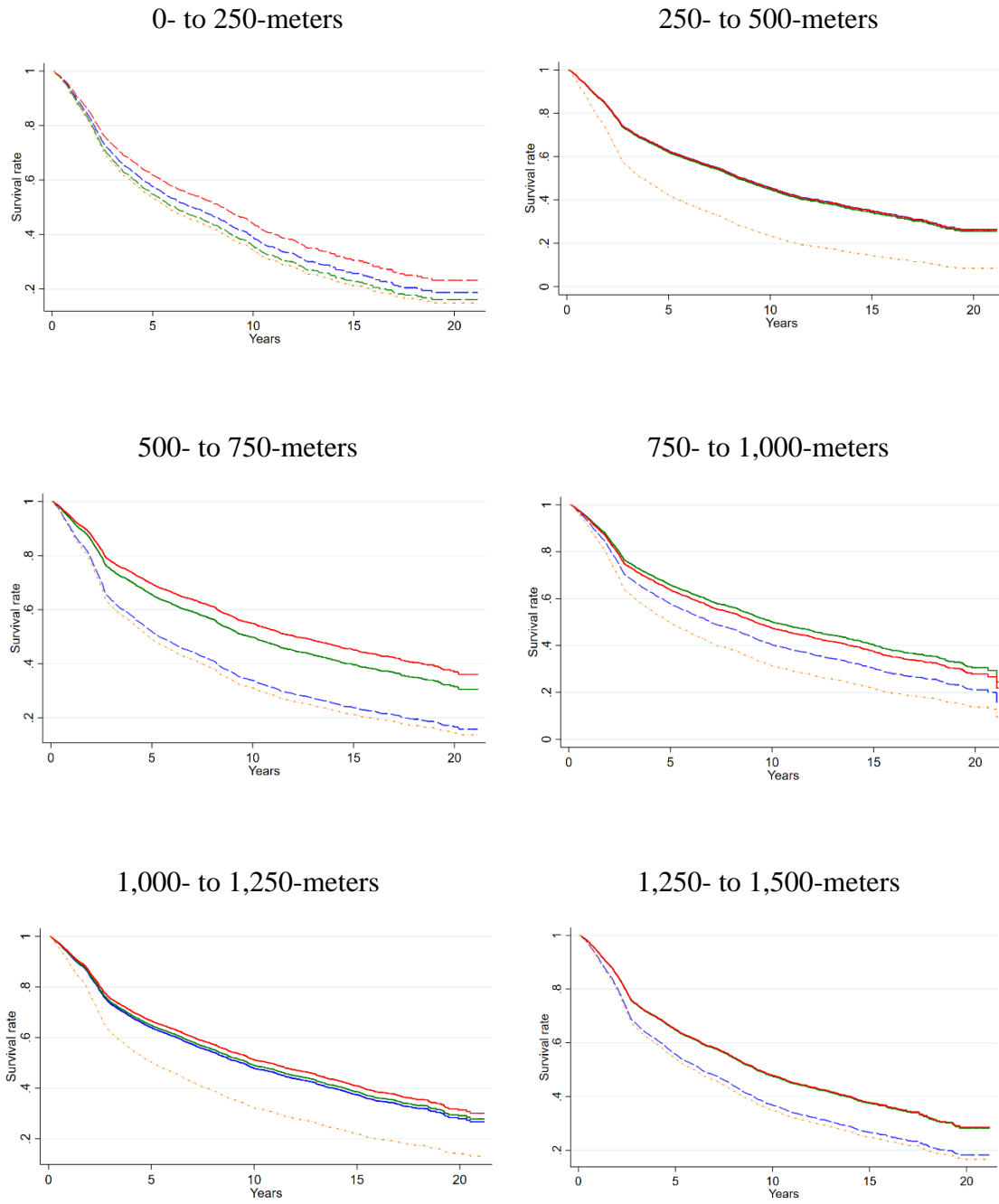
## Annexe du Chapitre 4

**Table A4.1 - Repartition of economic sectors in 15 categories**

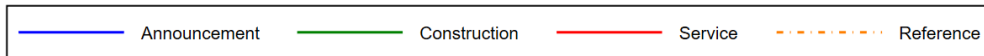
Category	Description
1	Primary (Agriculture, Forests, Fisheries, and Mines)
2	Manufacturing
3	Construction
4	Logistics, Transportation and Warehousing
<b>5</b>	<b>Wholesale</b>
<b>6</b>	<b>Retail</b>
<b>7</b>	<b>Finance, Insurance, and Real Estate (FIRE)</b>
<b>8</b>	<b>Business Services, and Knowledge Intensive Business Services (KIBS)</b>
9	Education
10	Health and Social Services
<b>11</b>	<b>Food and Accommodation</b>
<b>12</b>	<b>Arts and Entertainment</b>
<b>13</b>	<b>Personal Services</b>
14	Public Administration
15	Other

Note: All sectors used in the regression are indicated in bold.

**Figure A4.2.1 - Kaplan-Meier survival plots, for all economic sectors**

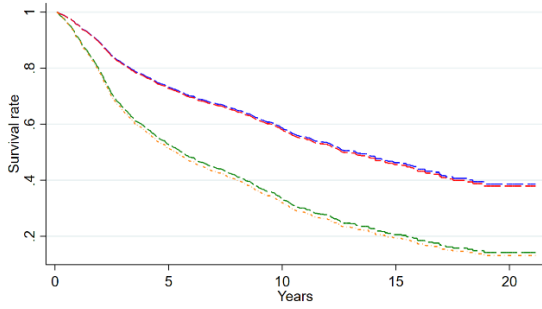


Legend: Dash line represents non-significant results.

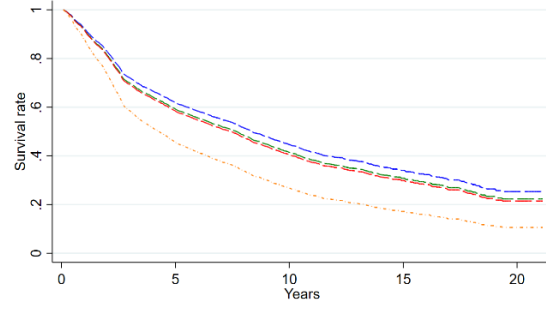


**Figure A4.2.2 - Kaplan-Meier survival plots, for retail activities**

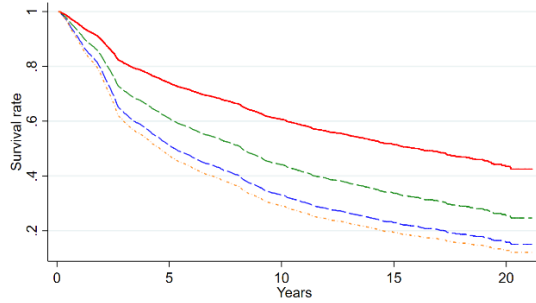
0- to 250-meters



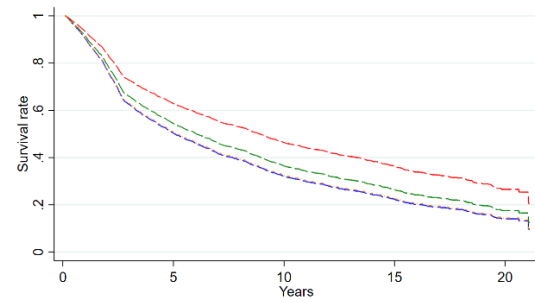
250- to 500-meters



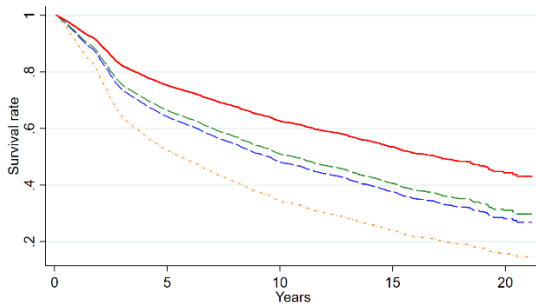
500- to 750-meters



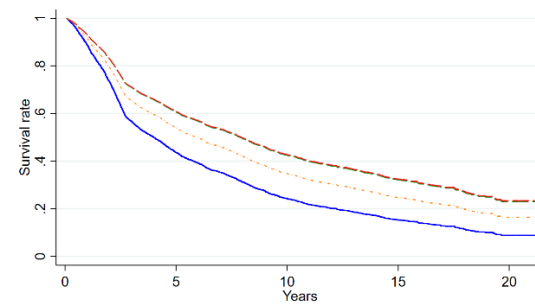
750- to 1,000-meters



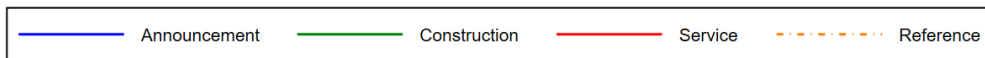
1,000- to 1,250-meters



1,250- to 1,500-meters

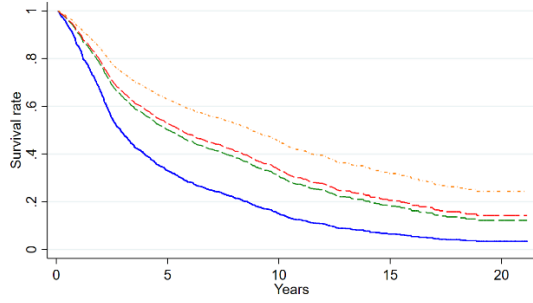


Legend: Dash line represents non-significant results.

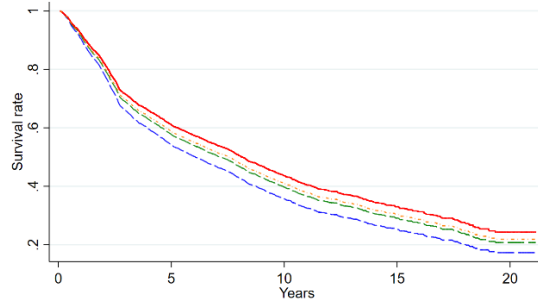


**Figure A4.2.3 - Kaplan-Meier survival plots, for FIRE and KIBS**

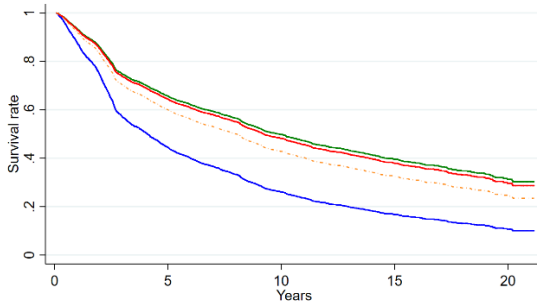
0- to 250-meters



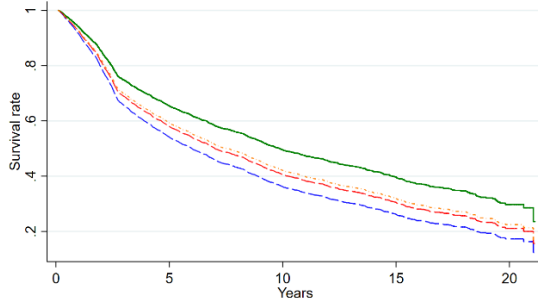
250- to 500-meters



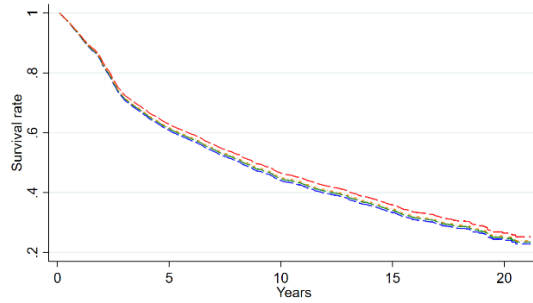
500- to 750-meters



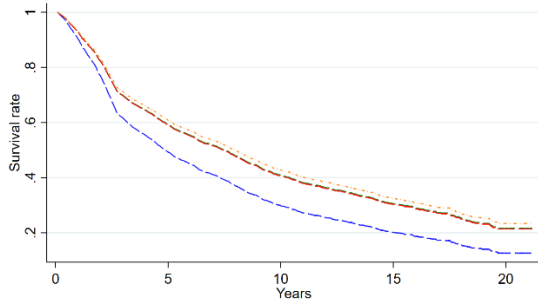
750- to 1,000-meters



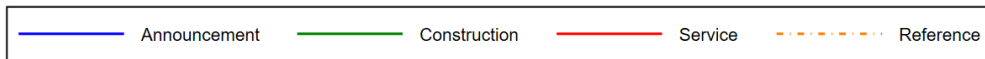
1,000- to 1,250-meters



1,250- to 1,500-meters



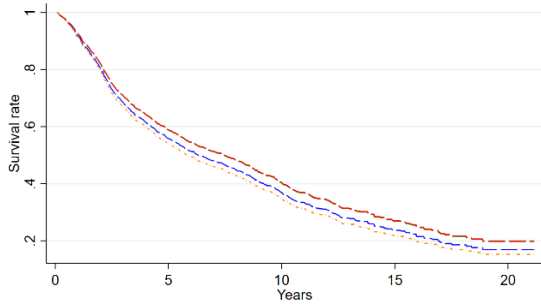
Legend: Dash line represents non-significant results.



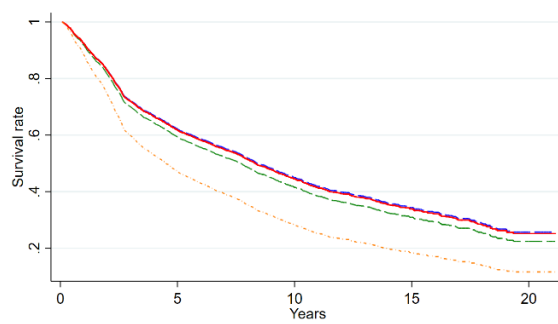


**Figure A4.2.4 - Kaplan-Meier survival plots, for local activities**

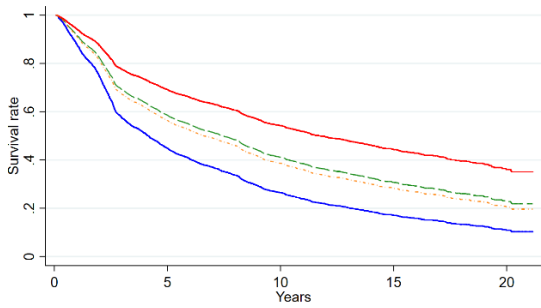
0- to 250-meters



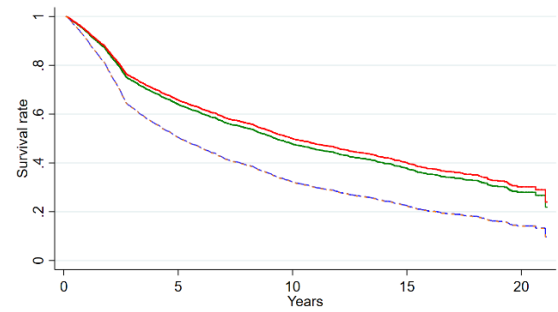
250- to 500-meters



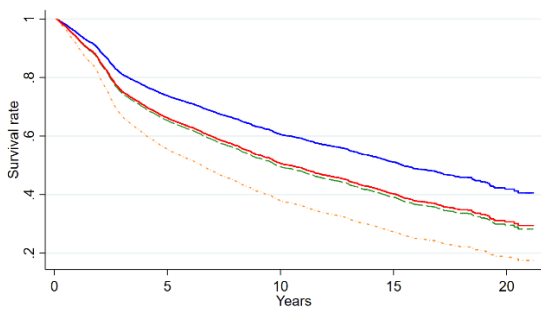
500- to 750-meters



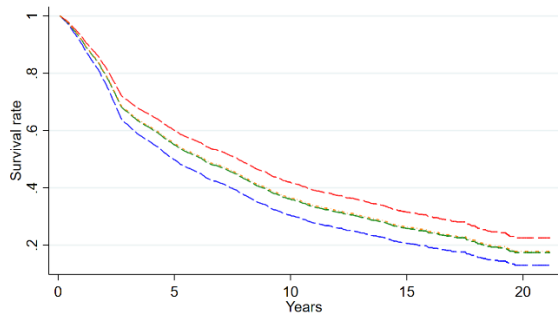
750- to 1,000-meters



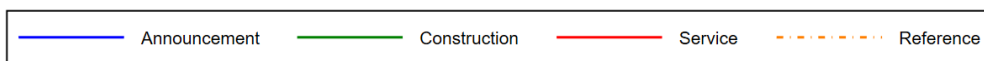
1,000- to 1,250-meters



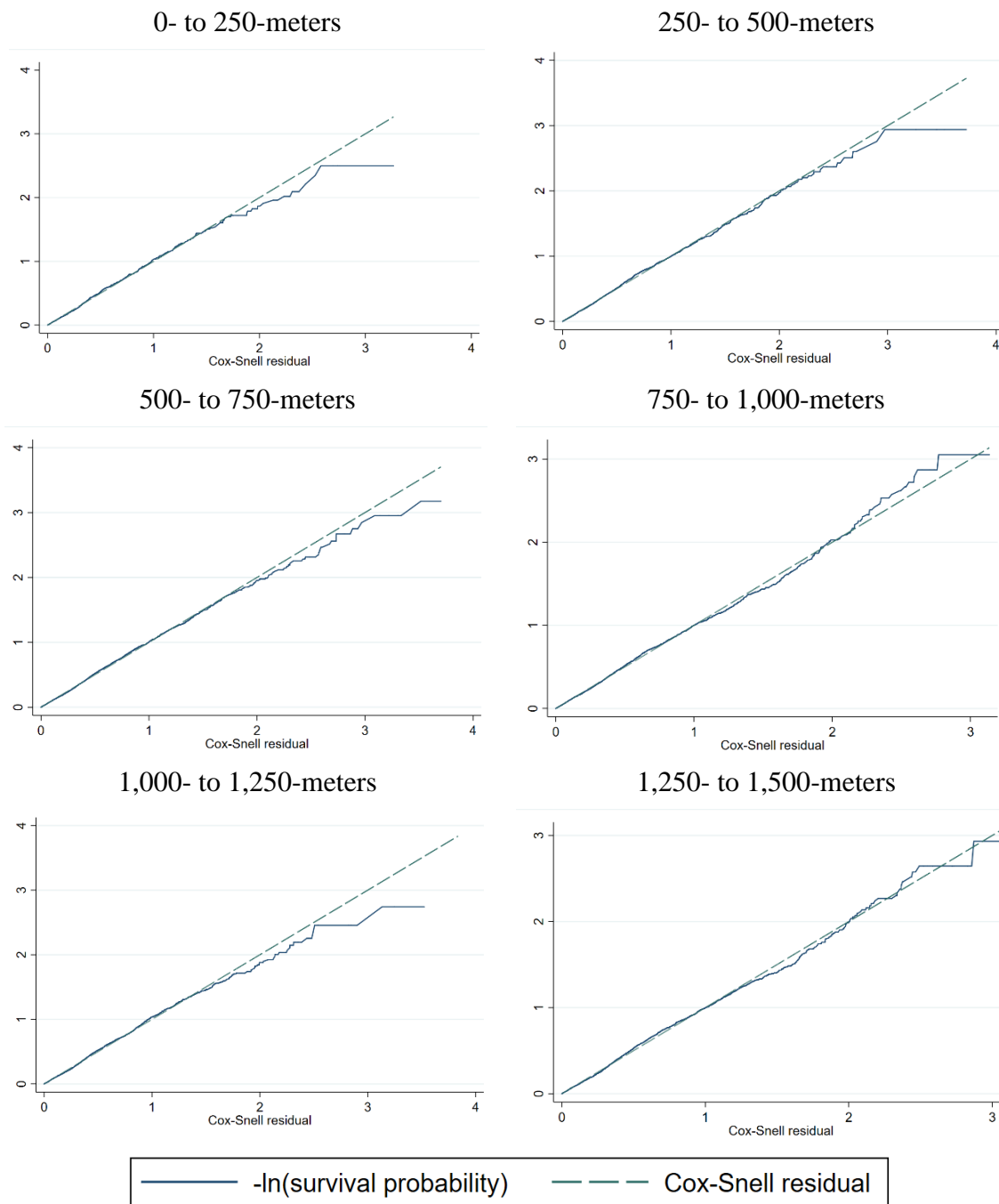
1,250- to 1,500-meters



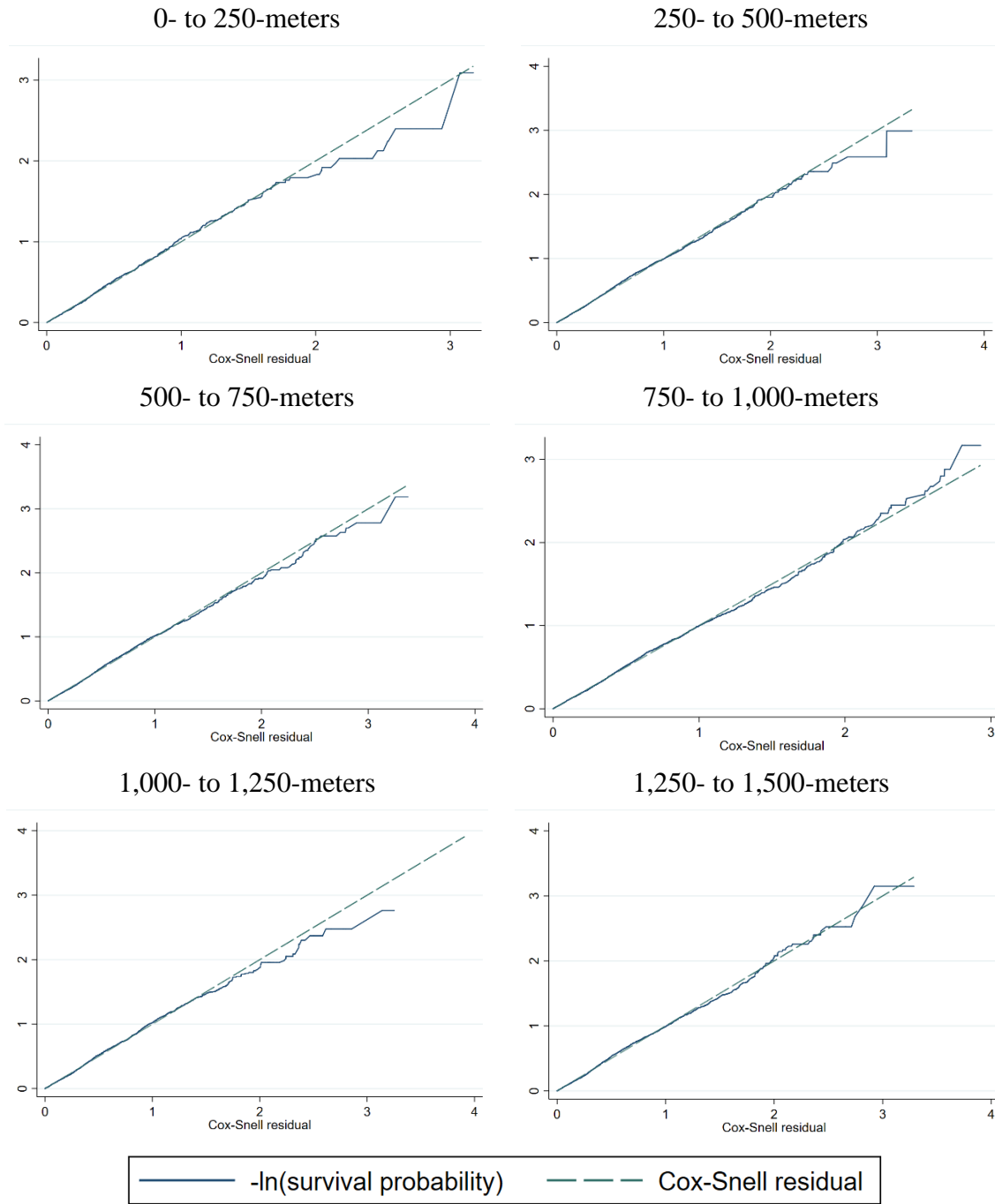
Legend: Dash line represents non-significant results.



**Figure A4.3 - Cox-Snell Residuals (Model 1)**



**Figure A4.4 - Cox-Snell Residuals (Model 2)**



**Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 1**

		0m - 250m		250m - 500m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
Reference (0 to 1 employee)									
Size (employees)	2 - 10	0.776 (0.065)	***	0.772 (0.065)	***	0.910 (0.048)	*	0.914 (0.048)	*
	11 - 25	0.791 (0.200)		0.784 (0.199)		0.568 (0.144)	**	0.572 (0.146)	**
	26 - 49	0.469 (0.214)	*	0.466 (0.212)	*	0.479 (0.216)		0.466 (0.210)	*
	50 - 99	0.317 (0.161)	**	0.268 (0.138)	**	0.445 (0.317)		0.466 (0.332)	
	More than 100	1.855 (0.634)	*	1.822 (0.624)	*	2.843 (0.537)	***	2.927 (0.551)	***
Reference (Wholesale)									
Sector	Retail	1.236 (0.181)		1.290 (0.195)	*	1.088 (0.091)		1.149 (0.105)	
	FIREKIBS	0.971 (0.129)		0.935 (0.129)		0.727 (0.055)	***	0.765 (0.065)	***
	Loc. Act.	1.224 (0.174)		1.221 (0.181)		1.003 (0.080)		1.075 (0.095)	
Local Indexes	LQ (Retail)	0.381 (0.499)		0.423 (0.554)		0.544 (0.527)		0.524 (0.509)	
	LQ (FIRE)	0.790 (1.478)		0.959 (1.799)		0.586 (0.783)		0.592 (0.794)	
	LQ (KIBS)	0.564 (0.453)		0.548 (0.442)		0.712 (0.405)		0.687 (0.392)	
	LQ (Food & Acc.)	3.889 (4.951)		3.453 (4.394)		2.871 (2.603)		2.859 (2.600)	
	LQ (Arts & Ent.)	0.395 (0.460)		0.393 (0.461)		0.678 (0.559)		0.646 (0.534)	
	LQ (Pers. Serv.)	0.501 (0.461)		0.480 (0.443)		0.471 (0.316)		0.508 (0.342)	
	HI	8398.9 (210921)		704.83 (17727)		9042.5 (163717)		16335.5 (296463)	
	Orange (Metro)	1.159 (0.278)		1.124 (0.232)		1.362 (0.158)	***	1.192 (0.114)	*
Metro	Orange x Ann.	1.037 (0.317)		.		0.728 (0.103)	**	.	
	Orange x Const.	1.134 (0.317)		.		0.741 (0.104)	**	.	
	Orange x Serv.	0.851 (0.255)		.		0.682 (0.095)	***	.	
Retail	Orange x Ann.	.		0.589 (0.246)		.		0.762 (0.155)	
	Orange x Const.	.		1.238 (0.622)		.		0.826 (0.197)	
	Orange x Serv.	.		0.675 (0.314)		.		0.942 (0.193)	
FIRE & KIBS	Orange x Ann.	.		2.103 (0.844)	*	.		0.973 (0.157)	
	Orange x Const.	.		1.294 (0.374)		.		0.865 (0.135)	

	Orange x Serv.	.	1.071 (0.369)	.	0.713 (0.114)	**
Loc. Act.	Orange x Ann.	.	1.097 (0.426)	.	0.762 (0.151)	
	Orange x Const.	.	1.001 (0.384)	.	0.838 (0.164)	
	Orange x Serv.	.	0.927 (0.347)	.	0.712 (0.129)	*

**Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 2**

		500m - 750m		750m - 1,000m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
Reference (0 to 1 employee)									
Size (employees)	2 - 10	0.903 (0.043)	**	0.900 (0.043)	**	0.876 (0.043)	***	0.872 (0.042)	***
	11 - 25	0.759 (0.187)		0.758 (0.187)		0.429 (0.116)	***	0.434 (0.117)	***
	26 - 49	1.406 (0.412)		1.395 (0.409)		0.758 (0.256)		0.781 (0.263)	
	50 - 99	0.656 (0.270)		0.658 (0.271)		0.541 (0.383)		0.545 (0.387)	
	More than 100	1.217 (0.205)		1.230 (0.207)		1.923 (0.293)	***	1.972 (0.300)	***
	Reference (Wholesale)								
Sector	Retail	0.999 (0.079)		1.057 (0.090)		1.038 (0.084)		1.062 (0.094)	
	FIRE & KIBS	0.698 (0.051)	***	0.713 (0.056)	***	0.767 (0.057)	***	0.801 (0.064)	***
	Act Loc	0.790 (0.061)	***	0.818 (0.068)	**	1.014 (0.079)		1.066 (0.091)	
Local Indexes	LQ(Retail)	0.153 (0.130)	**	0.147 (0.125)	**	0.170 (0.151)	**	0.158 (0.141)	**
	LQ(FIRE)	0.326 (0.399)		0.347 (0.425)		0.163 (0.202)		0.190 (0.237)	
	LQ(KIBS)	0.707 (0.369)		0.681 (0.356)		0.770 (0.423)		0.724 (0.398)	
	LQ(Food & Acc.)	5.093 (4.361)		4.613 (3.952)	*	7.369 (6.146)	**	6.788 (5.680)	**
	LQ(Arts & Ent.)	0.747 (0.555)		0.762 (0.567)		0.634 (0.477)		0.618 (0.466)	
	LQ(Pers. Serv.)	0.936 (0.569)		1.023 (0.622)		0.981 (0.620)		1.087 (0.683)	
	HI	6.46e7 (1.06e9)		7.9e7 (1.29e9)		6.35e11 (1.06e13)		2.57e11 (4.29e12)	
	Orange (Metro)	1.181 (0.109)	*	1.054 (0.087)		1.130 (0.108)		0.986 (0.082)	
Metro	Orange x Ann.	1.163 (0.135)		.		0.933 (0.111)		.	
	Orange x Const.	0.726 (0.086)	***	.		0.686 (0.081)	***	.	
	Orange x Serv.	0.587 (0.070)	***	.		0.707 (0.083)	***	.	
Retail	Orange x Ann.	.		7.9e7 (1.29e9)		.		1.180 (0.199)	
	Orange x Const.	.		1.176 (0.198)		.		1.039 (0.184)	
	Orange x Serv.	.		0.867 (0.157)		.		0.733 (0.133)	*
FIRE KIBS	Orange x Ann.	.		0.553 (0.106)	***	.		1.058 (0.148)	

	Orange x Const.	.	1.457 (0.219)	**	.	0.714 (0.102)	**
	Orange x Serv.	.	0.740 (0.110)	**	.	0.882 (0.113)	
Loc. Act.	Orange x Ann.	.	0.794 (0.114)		.	1.175 (0.199)	
	Orange x Const.	.	1.408 (0.226)	**	.	0.754 (0.121)	*
	Orange x Serv.	.	0.930 (0.175)		.	0.718 (0.120)	**

**Table A4.5 - Results by model using a 7-year operation period (2007-2013) part 3**

		1,000m - 1,250m		1,250m - 1,500m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
Reference (0 to 1 employee)									
Size (employees)	2 - 10	0.867 (0.042)	***	0.860 (0.041)	***	0.864 (0.044)	***	0.868 (0.045)	***
	11 - 25	0.388 (0.102)	***	0.401 (0.105)	***	0.352 (0.099)	***	0.355 (0.100)	***
	26 - 49	0.526 (0.187)	*	0.519 (0.185)	*	0.235 (0.166)	**	0.229 (0.162)	**
	50 - 99	0.422 (0.244)		0.402 (0.233)		0.377 (0.219)	*	0.374 (0.217)	*
	More than 100	1.847 (0.298)	***	1.792 (0.290)	***	1.890 (0.351)	***	1.934 (0.360)	***
Reference (Wholesale)									
Sector	Retail	0.909 (0.075)		0.988 (0.087)		0.942 (0.077)		0.920 (0.083)	
	FIREKIBS	0.739 (0.054)	***	0.738 (0.059)	***	0.744 (0.055)	***	0.731 (0.059)	***
	Loc. Act.	0.853 (0.068)	**	0.919 (0.077)		0.887 (0.070)		0.863 (0.074)	*
Local Indexes	LQ (Retail)	0.433 (0.355)		0.375 (0.306)		0.659 (0.608)		0.551 (0.509)	
	LQ (FIRE)	1.869 (2.200)		1.643 (1.937)		10.738 (15.08)	*	11.339 (15.948)	*
	LQ (KIBS)	2.476 (1.311)	*	2.296 (1.215)		1.502 (0.896)		1.374 (0.817)	
	LQ (Food & Acc.)	0.462 (0.384)		0.505 (0.420)		0.117 (0.117)	*	0.114 (0.114)	**
	LQ (Arts & Ent.)	0.853 (0.593)		0.901 (0.630)		1.208 (0.938)		1.170 (0.908)	
	LQ (Pers. Serv.)	1.033 (0.600)		1.132 (0.651)		0.653 (0.429)		0.758 (0.494)	
	HI	0.003 (0.049)		0.033 (0.525)		2.1e-10 (3.89e-9)		8.69e-10 (1.61e-8)	
Metro	Orange (Metro)	1.264 (0.124)	**	1.217 (0.103)	**	1.048 (0.126)		0.859 (0.085)	
	Orange x Ann.	0.821 (0.100)		.		1.055 (0.151)		.	
	Orange x Const.	0.792 (0.096)	*	.		0.775 (0.109)	*	.	
	Orange x Serv.	0.721 (0.089)	***	.		0.788 (0.117)		.	
Retail	Orange x Ann.	.		0.822 (0.164)		.		1.487 (0.274)	**
	Orange x Const.	.		0.759 (0.159)		.		0.907 (0.188)	
	Orange x Serv.	.		0.511 (0.105)	***	.		0.904 (0.231)	
FIRE KIBS	Orange x Ann.	.	0.926 (0.121)		.		1.285 (0.204)		



	Orange x Const.	.	0.908 (0.119)	.	0.956 (0.149)
	Orange x Serv.	.	0.860 (0.120)	.	1.038 (0.165)
Loc. Act.	Orange x Ann.	.	0.550 (0.126)	***	1.279 (0.273)
	Orange x Const.	.	0.774 (0.137)	.	1.104 (0.194)
	Orange x Serv.	.	0.717 (0.121)	**	0.961 (0.190)

Note: The Cox-Snell's residuals were estimated for both models to ensure that the data used were well suited for the Cox model. The Schoenfeld's residuals were also computed for all explanatory variables, using the "estat phtest" command in Stata. Based on those tests, the data are well suited and respect the PH assumption.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

**Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 1**

		0m - 250m		250m - 500m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
Reference (0 to 1 employee)									
Size (employees)	2 - 10	0.761 (0.666)	***	0.755 (0.666)	***	0.903 (0.050)	*	0.907 (0.051)	*
	11 - 25	0.732 (0.197)		0.730 (0.197)		0.522 (0.142)	**	0.527 (0.143)	**
	26 - 49	0.461 (0.210)	*	0.459 (0.209)	*	0.404 (0.204)	*	0.385 (0.194)	*
	50 - 99	0.316 (0.160)	**	0.268 (0.139)	**	0.463 (0.330)		0.491 (0.350)	
	More than 100	1.861 (0.637)	*	1.830 (0.627)	*	2.977 (0.565)	***	3.059 (0.578)	***
Reference (Wholesale)									
Sector	Retail	1.135 (0.170)		1.187 (0.183)		1.073 (0.093)		1.131 (0.106)	
	FIREKIBS	0.929 (0.125)		0.893 (0.125)		0.719 (0.057)	***	0.746 (0.065)	***
	Loc. Act.	1.188 (0.172)		1.180 (0.178)		1.002 (0.083)		1.087 (0.100)	
Local Indexes	LQ (Retail)	0.432 (0.611)		0.488 (0.689)		1.050 (1.109)		1.014 (1.076)	
	LQ (FIRE)	0.993 (2.121)		1.293 (2.764)		2.170 (3.421)		2.375 (3.754)	
	LQ (KIBS)	0.599 (0.500)		0.582 (0.488)		0.639 (0.381)		0.606 (0.363)	
	LQ (Food & Acc.)	3.266 (5.451)		2.686 (4.478)		0.750 (0.927)		0.678 (0.839)	
	LQ (Arts & Ent.)	0.371 (0.451)		0.375 (0.457)		0.885 (0.759)		0.884 (0.759)	
	LQ (Pers. Serv.)	0.470 (0.458)		0.450 (0.440)		0.522 (0.371)		0.581 (0.413)	
	HI	483.11 (13581)		20.59 (578.79)		0.001 (0.018)		0.001 (0.016)	
	Orange (Metro)	1.151 (0.276)		1.097 (0.227)		1.337 (0.155)	**	1.168 (0.113)	
Metro	Orange x Ann.	1.053 (0.321)		.		0.733 (0.104)	**	.	
	Orange x Const.	1.131 (0.316)		.		0.747 (0.105)	**	.	
	Orange x Serv.	0.813 (0.263)		.		0.677 (0.100)	***	.	
Retail	Orange x Ann.	.		0.624 (0.261)		.		0.771 (0.158)	
	Orange x Const.	.		1.327 (0.669)		.		0.838 (0.200)	
	Orange x Serv.	.		0.502 (0.311)		.		0.994 (0.239)	
FIRE KIBS	Orange x Ann.	.	2.129 (0.855)	*	.		0.996 (0.162)		

	Orange x Const.	.	1.317 (0.381)	.	0.887 (0.140)	
	Orange x Serv.	.	1.072 (0.413)	.	0.781 (0.132)	
Loc. Act.	Orange x Ann.	.	1.105 (0.430)	.	0.750 (0.149)	
	Orange x Const.	.	1.014 (0.390)	.	0.826 (0.163)	
	Orange x Serv.	.	1.032 (0.434)	.	0.576 (0.131)	**

**Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 2**

		500m - 750m		750m - 1,000m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
Reference (0 to 1 employee)									
Size (employees)	2 - 10	0.907 (0.045)	**	0.903 (0.044)	**	0.863 (0.044)	***	0.860 (0.044)	***
	11 - 25	0.708 (0.186)		0.706 (0.185)		0.427 (0.120)	***	0.432 (0.121)	***
	26 - 49	1.164 (0.372)		1.155 (0.370)		0.755 (0.255)		0.778 (0.263)	
	50 - 99	0.701 (0.288)		0.703 (0.289)		0.538 (0.381)		0.550 (0.390)	
	More than 100	1.205 (0.203)		1.218 (0.206)		1.887 (0.288)	***	1.932 (0.295)	***
	Reference (Wholesale)								
Sector	Retail	0.989 (0.080)		1.042 (0.090)		1.039 (0.086)		1.056 (0.095)	
	FIRE & KIBS	0.683 (0.051)	***	0.696 (0.055)	***	0.765 (0.058)	***	0.790 (0.065)	***
	Loc. Act.	0.786 (0.620)	***	0.809 (0.069)	**	1.002 (0.080)		1.052 (0.092)	
Local Indexes	LQ (Retail)	0.201 (0.186)	*	0.197 (0.182)	*	0.179 (0.179)	*	0.179 (0.179)	*
	LQ (FIRE)	0.606 (0.857)		0.662 (0.938)		0.246 (0.368)		0.331 (0.498)	
	LQ (KIBS)	0.697 (0.375)		0.674 (0.364)		0.609 (0.346)		0.573 (0.325)	
	LQ (Food & Acc.)	2.666 (2.965)		2.353 (2.620)		4.471 (5.183)		3.563 (4.149)	
	LQ (Arts & Ent.)	0.842 (0.644)		0.859 (0.658)		0.834 (0.643)		0.835 (0.645)	
	LQ (Pers. Serv.)	0.975 (0.620)		1.069 (0.680)		1.321 (0.885)		1.471 (0.979)	
	HI	37839.5 (702973)		30230 (562443)		9.15e9 (1.8e11)		6.06e8 (1.19e10)	
Orange (Metro)		1.159 (0.108)		1.041 (0.087)		1.098 (0.106)		0.969 (0.081)	
Metro	Orange x Ann.	1.164 (0.136)		.		0.954 (0.114)		.	
	Orange x Const.	0.730 (0.087)	***	.		0.703 (0.084)	***	.	
	Orange x Serv.	0.591 (0.075)	***	.		0.736 (0.090)	**	.	
Retail	Orange x Ann.	.		1.166 (0.197)		.		1.193 (0.203)	
	Orange x Const.	.		0.865 (0.157)		.		1.050 (0.187)	
	Orange x Serv.	.		0.549 (0.114)	***	.		0.707 (0.147)	*
FIRE KIBS	Orange x Ann.	.		1.471 (0.222)	**	.		1.074 (0.151)	

	Orange x Const.	.	0.745 (0.111)	**	.	0.729 (0.105)	**
	Orange x Serv.	.	0.785 (0.122)		.	0.961 (0.130)	
Loc. Act.	Orange x Ann.	.	1.395 (0.224)		.	1.190 (0.203)	
	Orange x Const.	.	0.922 (0.173)		.	0.760 (0.123)	*
	Orange x Serv.	.	0.524 (0.116)		.	0.653 (0.129)	**

**Table A4.6 - Results by model using a 5-year operation period (2007-2011) part 3**

		1,000m - 1,250m		1,250m - 1,500m					
		Model 1	Model 2	Model 1	Model 2				
		Hazard Ratio (Std. Err.)		Hazard Ratio (Std. Err.)					
		Reference (0 to 1 employee)							
Size (employees)	2 - 10	0.863 (0.043)	** *	0.858 (0.043)	***	0.852 (0.045)	***	0.857 (0.045)	***
	11 - 25	0.369 (0.104)	** *	0.382 (0.107)	***	0.356 (0.100)	***	0.360 (0.102)	***
	26 - 49	0.533 (0.190)	* 	0.527 (0.188)	*	0.238 (0.169)	**	0.233 (0.165)	**
	50 - 99	0.439 (0.254)		0.422 (0.244)		0.388 (0.225)		0.383 (0.223)	*
	More than 100	1.891 (0.309)	** *	1.841 (0.301)	***	1.865 (0.347)	***	1.902 (0.354)	***
			Reference (Wholesale)						
Sector	Retail	0.908 (0.077)		0.982 (0.089)		0.963 (0.080)		0.939 (0.086)	
	FIREKIBS	0.723 (0.055)	** *	0.731 (0.060)	***	0.741 (0.056)	***	0.719 (0.060)	***
	Loc. Act.	0.859 (0.070)	* 	0.920 (0.079)		0.910 (0.728)		0.890 (0.078)	
Local Indexes	LQ (Retail)	1.478 (1.338)		1.285 (1.161)		0.410 (0.407)		0.348 (0.346)	
	LQ (FIRE)	22.486 (32.483)	**	19.941 (28.846)	**	4.51 (7.177)		5.418 (8.643)	
	LQ (KIBS)	2.858 (1.586)	*	2.601 (1.442)	*	1.315 (0.792)		1.200 (0.720)	
	LQ (Food & Acc.)	0.044 (0.050)	** *	0.047 (0.054)	***	0.249 (0.308)		0.212 (0.263)	
	LQ (Arts & Ent.)	0.996 (0.714)		1.074 (0.769)		1.268 (0.999)		1.262 (0.994)	
	LQ (Pers. Serv.)	0.892 (0.545)		0.996 (0.601)		0.767 (0.522)		0.906 (0.610)	
	HI	1.31e-16 (2.46e-15)	* 	1.3e-15 (2.45e-14)	*	0.000 (0.000)		0.0000 (0.0003)	
	Orange (Metro)	1.242 (0.123)	**	1.197 (0.103)	**	1.034 (0.124)		0.840 (0.084)	*
Metro	Orange x Ann.	0.821 (0.100)		.		1.060 (0.152)		.	
	Orange x Const.	0.810 (0.099)	*	.		0.771 (0.109)	*	.	
	Orange x Serv.	0.732 (0.096)	**	.		0.777 (0.121)		.	
	Orange x Ann.	.		0.821 (0.164)		.		1.475 (0.273)	**
Retail	Orange x Const.	.		0.778 (0.164)		.		0.897 (0.186)	
	Orange x Serv.	.		0.495 (0.113)	***	.		0.920 (0.268)	
	FIRE KIBS Orange x Ann.	.		0.921 (0.121)		.		1.333 (0.213)	*

	Orange x Const.	.	0.932 (0.123)	.	0.984 (0.154)
	Orange x Serv.	.	0.872 (0.134)	.	1.135 (0.195)
Loc. Act.	Orange x Ann.	.	0.552 (0.126)	***	1.262 (0.270)
	Orange x Const.	.	0.785 (0.140)	.	1.077 (0.191)
	Orange x Serv.	.	0.763 (0.139)	.	0.860 (0.186)

Note: The Cox-Snell's residuals were estimated for both models to ensure that the data used were well suited for the Cox model. The Schoenfeld's residuals were also computed for all explanatory variables, using the "estat phtest" command in Stata. Based on those tests, the data are well suited and respect the PH assumption.

\*: p-value = 10%, \*\*: p-value = 5%, \*\*\*: p-value = 1%.

## Annexe du Chapitre 5

**Table A5.1 - Placebo test: Testing the parallel trend assumption (part 1)**

Critical distance	100 meters		200 meters		300 meters	
	Haz. Ratio (Std. Err)	PHA	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA
Tramway	1.398 (0.339)	0.671	1.032 (0.149)	0.431	1.043 (0.115)	0.964
Placebo Period	0.328 (0.096) ***	0.136	0.262 (0.048) ***	0.022	0.239 (0.041) ***	0.004
Tramway x Placebo Period	0.921 (0.467)	0.578	0.677 (0.217)	0.572	0.985 (0.244)	0.558
Distance to the city center	0.892(0.096)	0.054	0.975 (0.055)	0.814	1.027 (0.045)	0.777
Distance to the closest terminus	1.267 (0.175) *	0.461	1.050 (0.079)	0.046	0.994 (0.060)	0.673
Herfindahl index	0.038 (0.099)	0.324	0.213 (0.267)	0.668	0.203 (0.229)	0.141
LQ FIRE	1.455 (0.619)	0.312	1.249 (0.264)	0.290	1.130 (0.199)	0.990
LQ KIBS	1.448 (0.275) *	0.208	1.086 (0.118)	0.356	1.136 (0.106)	0.309
LQ Business services	1.018 (0.162)	0.255	1.045 (0.108)	0.311	1.064 (0.083)	0.817
LQ Personal services	0.635 (0.115) **	0.968	0.861 (0.079)	0.507	0.935 (0.066)	0.996
LQ Food and Accommodation	1.089 (0.238)	0.179	0.993 (0.111)	0.000	1.089 (0.089)	0.070
LQ Wholesale	0.887 (0.180)	0.275	0.952 (0.093)	0.280	1.021 (0.062)	0.983
Global PHA		0.172		0.000		0.000

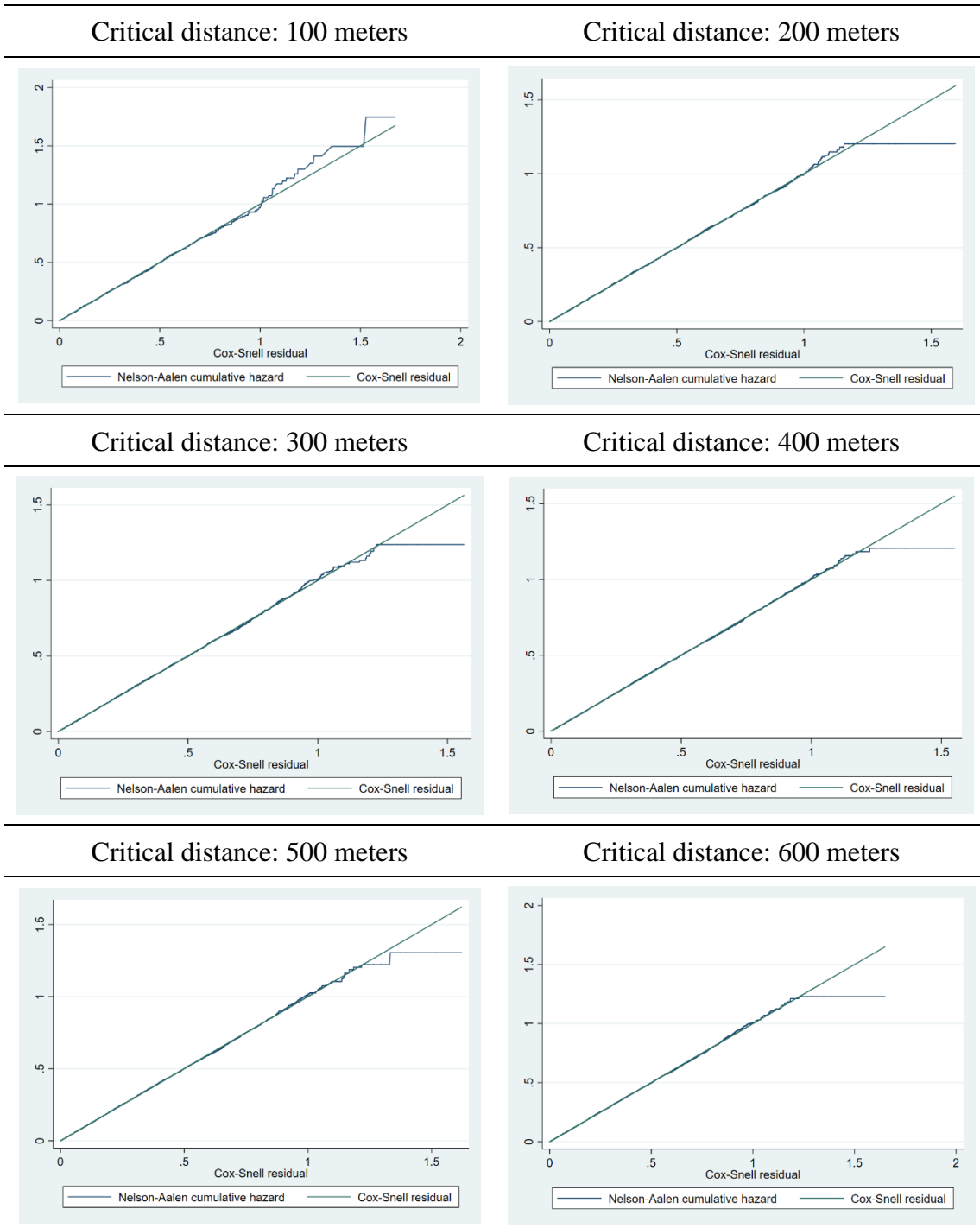


**Table A5.1 - Placebo test: Testing the parallel trend assumption (part 2)**

Critical distance	400 meters		500 meters		600 meters	
	Haz. Ratio (Std. Err)	PHA	Haz. Ratio (Std. Err.)	PHA	Haz. Ratio (Std. Err.)	PHA
Tramway	1.127 (0.111)	0.924	1.168 (0.111)	0.465	1.134 (0.107)	0.236
Placebo Period	0.282 (0.044) ***	0.000	0.321 (0.048) ***	0.000	0.316 (0.049) ***	0.001
Tramway x Placebo Period	0.843 (0.183)	0.616	0.790 (0.158)	0.584	0.827 (0.163)	0.461
Distance to the city center	0.993 (0.040)	0.634	0.970 (0.035)	0.458	0.967 (0.033)	0.276
Distance to the closest terminus	1.024 (0.056)	0.564	1.052 (0.055)	0.688	1.036 (0.053)	0.862
Herfindahl index	0.320 (0.300)	0.523	0.754 (0.567)	0.579	1.055 (0.721)	0.658
LQ FIRE	1.210 (0.190)	0.422	1.079 (0.141)	0.074	1.031 (0.126)	0.097
LQ KIBS	1.124 (0.092)	0.736	1.169 (0.085) **	0.858	1.149 (0.078) **	0.777
LQ Business services	1.090 (0.073)	0.801	1.072 (0.067)	0.852	1.100 (0.055) *	0.136
LQ Personal services	1.009 (0.059)	0.523	1.024 (0.052)	0.521	0.992 (0.046)	0.862
LQ Food and Accommodation	1.083 (0.076)	0.198	1.095 (0.071)	0.037	1.091 (0.069)	0.036
LQ Wholesale	1.020 (0.052)	0.132	1.009 (0.046)	0.051	0.998 (0.043)	0.120
Global PHA		0.000		0.000		0.000

Legend: \*, p-value < 10%, \*\*, p-value < 5%, \*\*\*, p-value < 1%. PHA: p-value of the Schoenfeld residuals test, testing for the Proportional Hazard Assumption.

**Figure A5.2 - Cox-Snell residuals: Testing the goodness-of-fit of the model**



Note: Variability in the right end tail is expected, when using the Nelson-Aalen cumulative hazard estimator, because of previous censoring and failures (Cleves et al., 2010).