ABSTRACT

Title of dissertation: A MICRO-LEVEL EXAMINATION OF THE IMPACT OF RAIL TRANSIT INVESTMENTS ON THE PATTERNS OF FIRM DYNAMICS

Basheer A M Saeed

Dissertation directed by: Professor Hiroyuki (Hiro) Iseki
School of Architecture, Planning and Preservation
University of Maryland, College Park

Transit-oriented development has been increasingly implemented at stations of both existing and new fixed transit systems across the U.S. to stimulate local economy and create livable communities. A common belief among planners in favor of transit-oriented development is that the provision of passenger rail systems promotes urban development around rail stations. There is a lack of empirical evidence, however, that supports this presumption. To address the gap in relevant literature, this dissertation examines the impact of passenger rail stations on the four different patterns of firm dynamics in the State of Maryland—firm birth and *inward* relocation as positive impacts, and firm closure and *outward* relocation as negative impacts. This dissertation uses both standard and propensity-score-weighted negative binomial regression methods to analyze the dependent variables of firm dynamics constructed from the National Establishment Time Series (NETS) panel data of the State of Maryland from 1990 to 2010. By examining both positive and negative impacts of firm dynamics, this dissertation estimates the likelihood of firm retainment and net relocation for areas in proximity of the passenger rail stations, while controlling for a number of potentially confounding factors.

Positive and statistically significant relationships are found between proximity to the passenger rail stations and the rates of firm births and *inward* relocating firms in Maryland, regardless of differences in the level of maturity of stations. From 1990 to 2010, the areas of passenger rail stations in Maryland experienced a wide range of rates of growth in firm density, depending on the year of station opening. The results of the four different patterns of firm dynamics suggest that areas near passenger rail stations gain belated economic benefits, well after the introduction of rail stations, shown by higher likelihood of firm retainment and net relocation around the mature rail stations opened before 1990. In comparison, areas near the less mature stations that opened after 1990 had predominantly lower likelihood of firm retainment and net firm relocation. Planners and policymakers should be proactive in directing development near rail stations by adopting a variety of measures and policies that support or at least consistent with transit-oriented development.

Key words: firm birth, firm closure, firm relocation, firm dynamics, National Establishment Time Series data (NETS), transit-oriented development (TOD), negative binomial, propensity score weights

A MICRO-LEVEL EXAMINATION OF THE IMPACT OF RAIL TRANSIT INVESTMENTS ON THE PATTERNS OF FIRM DYNAMICS

by

Basheer A M Saeed

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2018

Advisory Committee:

Professor Hiroyuki Iseki, Chair
Dr. Chao Liu
Professor Gerrit Knaap
Professor Lori Lynch
Professor Marie Howland

Dedication

In memory of my best friend, Hasan.

Acknowledgements

I would like to express my great appreciation to my advisor, Professor Hiroyuki Iseki, for his valuable and constructive feedback during the planning and development of this research work. I would also like to express my sincere gratitude to my committee members. Comments from Professor Gerrit Knapp and Professor Lori Lynch provided excellent suggestions to better focus this dissertation. I would like to give my sincere thanks to Professor Marie Howland and Dr. Chao Liu for showing great enthusiasm for my dissertation work and for their much-appreciated personal and intellectual guidance throughout my Ph.D. studies at the University of Maryland, College Park.

The completion of this dissertation would not have been possible without the help and support of my family in Kurdistan, Iraq, and in the United States. Most importantly, I wish to thank my loving and supportive wife, Laura Garnett, who has given me all her emotional and intellectual support since the beginning of my Ph.D. journey. My beloved mother and siblings have also been a tremendous source of emotional support that brought me up from my low points. I am also grateful to my father- and mother-in-law for their unconditional support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Maps

CHAPTER 1: INTRODUCTION

For more than 150 years the rail transit network has been playing a critical role in the U.S. transportation system and economy. Since the 1970's, the number of rail transit systems in the United States, i.e. heavy and light rail systems, has more than tripled from 22 rail transit systems in 1970 to 88 in 2015. Within the same period, rail transit ridership more than doubled, growing to more than 50 billion passenger-miles in 2014. Billions of U.S. dollars were spent in the development of these rail transit systems, made availabe through revenue from transit agencies and financial assistance from the state, local, and federal governments. In 2014 alone, around $36 billion was provided for capital investments and operation of rail transit systems (American Public Transportation Association, 2017).

Proponents of rail infrastructure often justify such substantial investments in rail transit systems because of their contributions to: (1) improved overall efficiency of transportation systems, (2) environmental sustainability, (3) reducing automobile dependence and congestion, and (4) promoting economic development. While the first three influences of rail investments are well-documented, the contribution of these investments to economic development is less understood and has recently attracted close attentions by transportation scholars and economists, as well as local officials and planners. A full understanding of the economic impacts of investments in rail transit system is critical to the decision-making process of policymakers.

The literature on the economic impacts of transit investments is largely focused on aspects related to property values and total employment. An important (and perhaps the least discussed) economic aspect of transit investments is their impact on changes in the patterns of firm dynamics. In the context of this research, firm dynamics refer to firm birth, firm closure, and firm relocation patterns. Some available evidence about firm dynamics suggests investments in rail transit contribute to denser employment clusters and even denser and more diverse cities in terms of economic activities, leading consequently to higher economic productivity (Chatman and Noland 2014). However, research examining the relationship between firm dynamics and transport infrastructure remains relatively limited, where most of the existing research examines the aggregate economic growth (Holl, 2006; Chatman et. al., 2016).

Due to a lack of empirical evidence, policymakers and academics disagree about the magnitude of impact that rail transit infrastructure has on the patterns of firm dynamics. The patterns of firm dynamics within regions and urban areas are important indicators of change in employment and economic growth. New firm birth to an urban economy signals innovation and is an indicator of economic growth (Reynolds, 1994; Chatman et. al., 2016). Firm birth alone does not capture, however, the overall spatial patters of firm dynamics. After all, new firms in a given location may either fail or decide to relocate to a more economically suitable location. Therefore, examining firm birth alone will inflate the estimated impact of rail investments on economic development. A careful examination of the influence that rail investments have on these patterns of firm dynamics will substantially enrich our understanding of their overall contribution to economic development.

Variation in the geographical scope considered in empirical studies is a contributing factor to the lack of consensus among scholars focusing on the determinants of firm dynamics. Rail transit infrastructure investments have numerous impacts on the location and spatial organization of firms at the micro-level, which cannot be captured at the macro-level (Holl, 2006). Further, there is a considerable variation in the patterns of firm dynamics happening at the micro scale (i.e. rates of firm births, closures, and relocations) because of differences in transport accessibility and agglomeration economies across localities. Nonetheless, a disproportionate share of literature on firm dynamics takes an aggregate scope toward examining the impact of transportation infrastructure, i.e. at regional levels (Smith and Florida, 1994; Manzato, et al., 2010; Nguyen et al., 2013). Larger geographic units of analysis hide micro-spatial patterns of firm dynamics that are essential to proper planning and justification of future transportation investments as a catalyst for local economic development. A micro-level examination can clarify the relationship between transportation infrastructure and firm dynamics, and a scarcity of empirical research on this relationship warrants further analysis of transportation infrastructure impacts at the micro-level.

Industrial aggregation is another equally important contributing factor to the discrepancy in the findings of empirical studies. Empirical research on firm dynamics often focuses on the manufacturing industrial sector alone or on all industrial sectors combined. Firms' sensitivity to transportation costs may vary across industry sectors, and these costs are linked to the availability and form of transportation infrastructure. Transport-dependent firms seek to minimize total transport costs, so they are more likely to benefit of a location that minimizes transport costs (such as areas proximate to rail stations) compared to less transport-dependent firms.

Retail firms, for example, are sensitive to changes in transportation costs. Retail customers often economize on travel costs through multipurpose shopping (Pellenbarg et al., 2002), which attracts retail firms to locations that are more accessible for their customers. Therefore, changes in transportation costs may considerably impact retail firm dynamics. Manufacturing firms, on the other hand, that are resource-oriented are less likely to be influenced (location wise) by changes in transportation costs (O'Sullivan, 2005). There are, however, manufacturing firms that can be sensitive to transportation costs, depending on the proportion of the transportation to total costs. Despite the extensive theoretical literature on the subject, empirical evidence remains inadequate on the sensitivity of firms across sectors to transportation accessibility in general and to passenger rail accessibility more specifically.

The objective of this research is to examine the impact of rail transit investments on the patters of firm dynamics, looking at firm birth, closure, and relocation patterns in areas within short distances from three passenger rail transit systems located within five jurisdictions of the State of Maryland. The following chapter reviews literature on the determinants of firm birth, closure, and relocation to provide a conceptual framework of the patterns of firm dynamics. The chapter provides a review of studies that empirically examine the association between passenger rail investments and the patterns of firm dynamics. Chapter 3 describe in detail both regression methods used for the analysis: (1) a standard negative binomial regression method, and (2) a PS-weighted negative binomial regression method to control for the endogeneity of rail transit investments. The research examines the patterns of firm dynamics of multiple industry sectors individually to determine their sensitivity to proximity to passenger rail stations using the two regression methods. The regression analyses examine firm birth, closure, and relocation impacts across multiple firm size categories and industry sectors. Chapters 4 discuss the analysis results of firm birth and closure, while Chapter 5 discuss the analysis results of firm relocation (inward and outward relocation). Chapter 6 covers conclusions, policy/planning implications, and future research agendas.

This research hypothesizes that areas within short walking distances to passenger rail stations experience, on average, positive net gain in firm birth and firm relocation (through improved transport accessibility) compared to areas farther away from the stations. The research also hypothesizes that the magnitude of effect experienced in areas near the transit stations varies across industry sectors. By determining these magnitudes of effect, policymakers who advocate for transit-oriented development can have a better understating on what industry sectors are more likely to thrive near rail transit stations, and consequently guiding future urban development policies.

CHAPTER 2: LITERATURE REVIEW

This chapter provides the foundation of existing literature on the determinants of firm dynamics with a specific focus on transportation-related determinants. The first section defines the nature of firm dynamics relevant to this research. The second section provides a theoretical and empirical review of the literature on the determinants of firm dynamics, i.e. firm birth, closure and survival, and relocation patterns. The third section reviews empirical studies that either directly or indirectly examine the influence of transportation-related factors on the patterns of firm dynamics. The last section provides more detail on the methodologies, geographical scopes, and industrial aggregations considered in the analyses of these empirical studies.

2.1. Firm dynamics: birth, closure, survival, and relocation

There is a common spatial element within the literature on the determinants of firm location decisions and on the determinants of the patterns of firm dynamics. The former is concerned with the location patterns of all existing firms (i.e. agglomeration forces), while the later examines variations in firm birth, firm closure or survival, or firm relocation patterns (i.e. firm dynamics) across certain geographic units of analysis, e.g. countries, regions, or counties. Nonetheless, these studies are both concerned with the question of how firms choose where to locate. This question has been under examination since the seminal work of Alfred Marshal (1890) titled "Principles of Economics." Studies on the determinants of firm dynamics do not examine only the location decisions of firms, but also take a step further to examine firm birth, closure, or relocation patterns (firm birth is the most commonly examined amongst the three patterns). There has been, however, no empirical study to date that examined the determinants of all these patterns of firm dynamics combined for a particular region. One objective of this research is to fill that gap.

What is the relevance of jointly examining all these firm dynamics in an economy instead of only examining, for instance, firm birth? Schumpeter's "creative destruction" argument (Schumpeter, 1934) provides a glimpse to the answer. Schumpeter argues that an industrial restructuring of a region occurs through replacement of less efficient and less innovative firms by those that are newer, smaller, and more innovative. Firm location, however, is the result of either a new firm birth into the economy or the relocation of an already established firm, yet most attention in literature has been given to the location of newly formed firms and their determinants. While a limited but sizable number of studies in the past empirically examined the determinants of the spatial patterns of firm closure and firm relocation, studies accounting for access to rail transit variables are lacking.

Earlier studies on the determinants of firm birth have used the total number of firm births in a region as the dependent variable. Since regions vary in size, it can be misleading to only use the number of firm births when examining their variation across different regions (Armington and Acs, 2002). To standardize the number of firm births, two empirical methods for operationalizing firm birth as a dependent variable are notable in literature (Audretch and Fritsch, 1994; Armington and Acs, 2002; Sutaria, and Hicks, 2004; Lasch, Robert and Le Roy, 2013). The first approach is known as the 'ecological approach' in which the total number of firm births in any given geographical unit of analysis is divided by the total number of existing firms within that geographical unit. The second approach, named the 'labor market approach', standardizes the total number of firm births relative to the size of the labor force (Audretch and Fritsch, 1994a, 1994b). These studies, however, do not take into account differences between the number of firm births and the number of firm closures across the units of analysis.

The relationship between firm closure and firm birth is not straightforward. In an economic analysis of firm dynamics, understanding the spatial patterns of firm closure is equally important to the understanding of the spatial patterns of firm birth. Nonetheless, most empirical studies fail to control for, let alone analyzing, the rates of firm closure when examining the determinants of firm birth. The rate of firm closure (also known as firm exit, destruction, or failure) is included as an independent variable in a few empirical studies examining the determinants of firm birth (Sutaria and Hicks, 2004; Chatman, Noland, and Klein, 2016) (the following section provides detailed discussion on the determinants of firm birth and firm closure). Relatively high number of firm births in an area may not necessarily mean a positive economic trend if the incidents of firm closure are higher in that area.

The relationship between firm birth and firm closure can be either positive or negative. Over time, more firm births may lead to more firm closures when a process called "competition effect" is at work. This means existing firms fail to compete with newly formed firms to meet market demand and then subsequently exit the economy. On the other hand, more firm births may lead to less firm closures when the market demand increases for business products and services in a process called "multiplier effect" (Johnson and Parker, 1996; Sutaria and Hicks, 2004; Cainelli, 2014). The multiplier effect hypothesizes that firm births cause more future firm births and impede future firm closures, or that firm closures cause more future firm closures and impede future firm births.

Moreover, firm dynamics are not limited to birth and closure. Firms operate in a dynamic environment where their internal and external contexts are continuously prone to change. Such forces can either attract or compel firms to relocate. Factors external to firms are often referred to as push factors (i.e. relocate out of the exiting location) and pull factors (i.e. relocate to the attracting location) (Risselada et al. 2012). The push and pull (or keep) factors are firm-related, location-specific factors and typically similar to those considered within studies examining the determinants of firm birth. Push factors are negative since they drive firms to out-migrate or steer firms away from relocating inward. Pull factors, on the other hand, are positive since they attract firms to relocate inward while retaining the existing ones. The examination of firm relocation and closure dynamics, however, requires an account of factors specific/internal to firms. Changes in internal factors such as age, size, and structure may potentially influence firms to relocate or close. The following section on the determinants of firm dynamics (section 2.1) provides more details on these factors.

Therefore, in an economy, net growth or decline in economic activities cannot be fully understood without a broader analysis of firm dynamics. When a firm enters an economy, other things being equal, it signals innovation and positive contribution to economic development (Reynolds, 2014; Armington and Acs, 2002; Chatman et. al., 2016). Other things are not equal, however. Firms also close and relocate within confined economies. Low closure rates or high survival rates of newly formed firms signal a prevalence of economic opportunities that maintain the economy and promote economic growth (Jostarndt and Rudolph, 2007). High rates of firm relocation toward an area signal a flourishing and attractive economy of that area and contribute consequently to economic growth (Neumark, Zhang, and Wall 2005).

A combined record of firm birth, closure, and relocation patterns provides a stronger measurement of net economic growth or decline in an area. Across different areas of a region, the net difference in the incidents between firm birth and closure are unlikely similar. The same dissimilarity applies to the incidents of inward and outward firm relocation in an area. Firm closure and relocation dynamics, in general, are not well studied. The role of rail transit investments in these dynamics are particularly neglected in most empirical studies on the determinant of firm birth; it is not clear whether relocated firm records are included or excluded from the records of firm birth (see for example, Smith and Florida, 1994; Coughlin and Segev, 2000; Armington and Acs, 2002), while only a few empirical studies distinguish between surviving and early failing new firm entrants when examining their determinants (Elert, 2014). This research addresses this gap in the literature by examining the determinants of firm birth, closure, and relocation patterns over time and across micro-level geographic units of analysis.

2.2. Determinants of firm dynamics: theory and practice

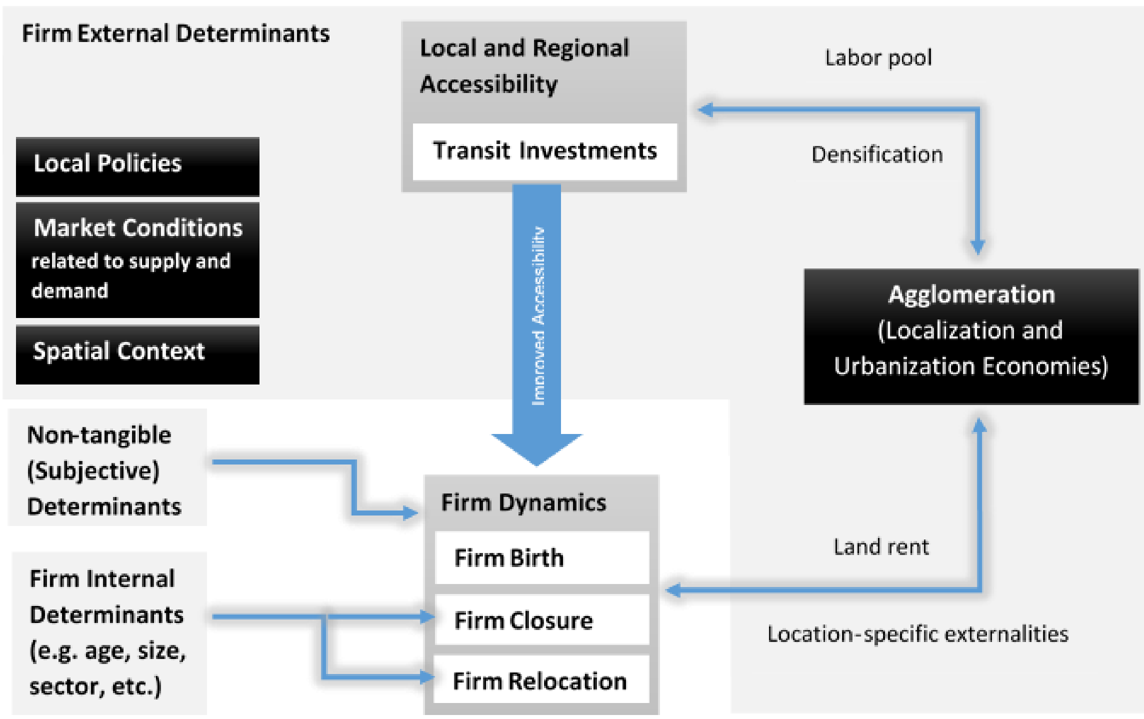
Studies on the determinants of firm dynamics (i.e. firm birth, firm survival and closure, and firm relocation) draw theoretical basis mainly from theories of economic geography or more specifically, location theories. Two classical pieces of literature by Krugman (1991) and Hayter (1997) have renewed the spatial dimension in mainstream economics, the widely accepted economics as taught across prominent universities. As agreed by many economists, Krugman's work (1991) was the beginning of New Economic Geography, the field that tries to explain "what the spatial dimension of the economy had to say about the nature of economic forces" (Krugman, 2011). Hayter's book (1997) titled "Industrial Location Theory" also attracted wide attention from scholars to the spatial implications of urban economics. The underlying assumption of both location theories is that firms seek to minimize production and transportation expenditures and maximize returns.

In recent decades, an increasing number of economists, geographers, and urban planners have directed attention toward examining the determinants of firm dynamics. As discussed in section 2.1, firm birth is the most empirically examined pattern of firm dynamics, while the determinants of firm relocation decisions are the least examined. There are no direct theories that link transportation investments to the patterns of firm dynamics (i.e. firm birth, closure, and relocation). The theoretical framework of this research is, therefore, driven mainly by the existing theories on industrial location as well as the general theories and empirical research on the determinants of firm birth, closure, survival, and relocation patterns.

The literature on the subject provides an extensive list of factors (external factors) that influence firm dynamics. Most empirical studies tend to group the determinants of firm birth into two groups: (1) market conditions, and (2) localization and urbanization economies (Borwning, 1980; Reynolds, 1994; Ace, Armington, and Zhang, 2007; Brixy and Grotz, 2007; Strotmann, 2007, Wennberg and Lindqvist, 2010). Market conditions normally include variables on socio-economic characteristics such as population, income, race, and level of education. Localization and urbanization economies (agglomeration economies) include variables on population and employment densities, firm density, and density of firms in similar industry sector. A few studies, however, go beyond these traditional factors to include other relevant determinants of firm dynamics. Reynolds (1994) goes further to include measures of local policies in his analysis of firm birth. Armington and Acs (2002) and Kronenberg (2012) take the variation in the regional (macro) context into account when examining the determinants of firm birth and firm relocation, respectively, by including distance to regional center in their analyses. There is a dearth of studies that include firm-specific (internal) variables in their analyses, mainly due to data limitation (Sleutjes and Beckers, 2013).

The determinants of firm dynamics (i.e. firm birth, closure, and relocation) can be categorized into five main groups of factors: (1) market conditions related to supply and demand, (2) agglomeration economies (urbanization and localization economies), (3) policy environment, (4) regional context, and (5) firm-specific (internal) factors. The conceptual framework in Figure 1 shows hypothetical relationships between these groups of factors and the patterns of firm dynamics. The hypothesis is that rail investments influence changes in firm birth, closure, and relocation patterns though improved accessibility. Firm agglomerations may also play an intermediary role between transit investments and the patterns of firm dynamics. That is, transit investments influence changes in firm agglomeration (density), which consequently influence changes in firm dynamics due to localization and urbanization externalities (see Figure 1).

Figure 1. A conceptual framework of the determinants of firm dynamics



The following subsections provide a discussion of each group of factors in detail along with a list of variables considered by empirical studies to operationalize these five groups of factors, as well as their impact on firm dynamics.

2.2.1. Market conditions related to supply and demand

Across regions and urban areas, the socio-economic structure of the population is fundamental to the supply and demand aspects of the economy. Markets and demand for goods and services change over time and across locations. Entrepreneurs and firms constantly react to changes in the patterns of local labor and consumerism, which in turn leads to various potential changes in firm birth, survival, and relocation patterns. Therefore, explanatory variables related to population growth, income growth, unemployment, and education level are the most commonly used in empirical studies to operationalize market conditions when examining the patterns of firm birth (Reynolds, 1994; Armington and Acs, 2002; Sutaria and Hicks, 2004; Cheng and Li, 2010; Cheng and Li, 2011), the patterns of firm survival (Wennberg and Lindqvist, 2010; Ace, Armington, and Zhang, 2007), or the patterns of firm relocation (Risselada, Schutjens, and Van Oort, 2013).

Changes in population sizes of urban areas lead to changes in the size of the labor force and in the demand for products and services. Such changes may consequently lead to different patterns of firm dynamics. A high population growth rate, for instance, has a positive influence on the rates of firm birth, firm survival, and net firm relocation (i.e. inward firm relocations minus outward firm relocations). Empirical studies on the determinants of firm birth conclude that population growth positively influences firm birth (Audretch and Fritsch, 1994b; Reynolds, 1994; Guesnier, 1994; Armington and Acs, 2002; Sutaria and Hicks, 2004; and Cheng and Li, 2010). The positive relationship between population and firm birth remains when studies use either the ecological or the labor market method to standardize firm birth.

Household income is another key factor that influences demand in a local market. There is unambiguous evidence that income growth stimulates firm birth (e.g. Reynolds, 1994; Smith and Florida, 1994; Armington and Acs, 2002; Sutaria and Hicks, 2004), and attracts firms to relocate (Kronenberg, 2012). Income growth in an area increases the demand for goods and services hence stimulating the birth of new firms, or the attraction of firms from other locations. Chatman et al. (2016) examined firm birth of two-digit NAICS industrial sectors within the Census blocks of Portland, Oregon and Dallas, Texas metropolitan areas; Surprisingly, they found that median household income has a negative association with firm birth for firms that have more than five employees. They found the association between household income and firm birth to be positive, however, for smaller firms (i.e. firms with five or fewer employees). It is not clear from findings in the literature how income growth influences firm closure. Manzato, et al. (2010) examined rates of firm closure of fifteen office industry sectors within Netherlands municipalities. For twelve out of fifteen sectors examined, they found that the higher the average population income, the higher the rates of firm closure. The authors do not provide any explanation on why this positive association exists between income and firm closure. One likely explanation is that high income level in an area indicates inflated costs of property and labor which may force some existing firms that are less competitive out of business.

The literature on the subject of firm birth provides inconsistent conclusions about the influence of unemployment on the rates of firm birth. Unemployment rate was found to have both positive and negative relationships with the rates of firm birth. For instance, a few empirical studies found higher unemployment rates to have a positive influence on the rates of newly formed firms (Reynolds, 1994; Armington and Acs, 2002). Another study suggests that unemployed workers are more likely to start their own businesses compared to employed workers. Therefore, unemployment rates may consequently decrease within regions with high unemployment rates due to new entrepreneurship and the potential employment opportunities they create (Sutaria and Hicks, 2004). Armington and Acs (2002), using ordinary least square (OLS) regression, attribute this positive[1] relationship to the exceptionally low level of unemployment within the US Labor Market Areas[2] (LMAs) in 1990s. The implication in these studies for the positive association between unemployment and firm birth is that when individuals become unemployed in a region, the rates of firm birth tend to go up.

On the other hand, other empirical studies have found a negative relationship between unemployment and firm birth (Audretch and Fritsch, 1994b; Sutaria and Hicks, 2004). That is, the higher the unemployment rate, the lower the rates of firm birth. Using a cross-sectional OLS regression for seventy-five large regions in West Germany, Audretch and Fritsch (1994b) found a negative association between unemployment and firm birth, suggesting that higher unemployment rates lead to lower rates of firm birth. Sutaria and Hicks (2004) examined the rates of firm birth in the manufacturing industry between 1976 and 1991 for the twenty-seven Metropolitan Statistical Areas[3] (MSAs) in Texas, Unites States. Using fixed-effect panel regression models, they found a negative relationship between changes in unemployment rate and the rate of firm birth.

---

[1] Their finding was small in magnitude and statistically insignificant for all firms but statistically significant for five out of six examined industry sectors.
[2] According to the U.S. Department of Labor, LMAs are sub-state geographic areas that consist of one or more counties or county equivalents.
[3] According to the U.S. Census bureau, MSAs consist of one or more counties that contain a city of 50,000 or more inhabitants and/or urbanized areas (UA) of 100,000 or more inhabitants.

The use of cross-sectional methods and/or relatively large units of analysis limit the validity of the findings of many of these reviewed empirical studies that found either positive or negative associations between unemployment and firm birth. Many micro-level differences exist within a region over time regarding the association between unemployment (as well as other determinants) and the patterns of firm dynamics. Section 2.4 discusses limitations that the past studies of the determinants of firm dynamics had regarding methodologies, geographic units of analysis, and industrial aggregation.

Empirical studies often use education level as a proxy for the level of technical skills needed in the economy, such as those of engineers and scientists, and the level of entrepreneurial skills needed to start a business, such as in the sectors of finance and marketing (Armington and Acs, 2002). To measure the level of educated or skilled population, studies often use the share of total population or adult population with college or higher degrees in geographic units of analysis. The literature on the determinants of firm birth unanimously reveals that the higher the share of population with college or higher degrees, the higher the rates of firm birth (Audretch and Fritsch, 1994b; Reynolds, 1994; Guesnier, 1994; Armington and Acs, 2002; Sutaria and Hicks, 2004; and Cheng and Li, 2010). Additionally, Armington and Acs (2002) found that higher shares of unskilled workers (measured by the share of adults without high school degrees) also have a positive influence on firm birth after controlling for the share of adults with college or higher degrees. The positive influence of population with the lower level of education on firm birth is attributed to the fact that nearly all firms need unskilled workers. Therefore, in addition to the availability of highly skilled workers, a greater share of cheap labor facilitates the process of firm birth (i.e. share of workers without high school degrees).

Among researchers, less popular factors influencing market conditions are related to quality of life, such as living costs, local amenities, property ownership, opportunities for cultural experiences, landscape, social capital, and political and administrative climate (Płaziak and Szymańska, 2014; An et al., 2014). Factors related to quality of life can be relevant to the location decision of firms and firm birth because they indicate how innovative and inviting the business (or investment) climate is in a region. A limited number of empirical studies, however, control for factors related to quality of life when examining the determinants of firm dynamics. An et al. (2014) include residential location factors (such as density of schools and density of large grocery stores) in addition to traditional factors to examine firm relocation patterns within the service and manufacturing industrial sectors in the Seoul Metropolitan Area. Their results suggest that the density of schools and large grocery stores have no statistically significant impact on firm relocation in both manufacturing and service sectors. Examining firm survival, Wennberg and Lindqvist (2010) use a time-variant measure of mean housing prices as a proxy for the cost of living in a region. Their analysis, however, shows positive but statistically insignificant relationship between housing prices and firm survival rates.

Table 1 provides a summary of market condition variables influencing the patterns of firm dynamics that are considered by the empirical studies on firm birth, closure, or relocation. The columns in this summary table provide information on: (1) the explanatory variables, (2) the empirical study that controlled for the stated explanatory variable, (3) the measurement used to operationalize the stated explanatory variable, (4) the firm dynamic under examination in each respective study (B=firm birth, C=firm closure, and R=firm relocation), and (5) the direction of impact each explanatory variable has on firm birth, closure, or relocation decisions(i.e. whether positive, negative, or statistically insignificant).

Table 1. Market conditions variables influencing the patterns of firm dynamics

No.	Explanatory Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	(-) Effect	Statistically Insignificant Effect
1	Population	Audretch and Fritsch (1994b)	Population change	B	All sectors, Manufacturing, Services		
		Reynolds (1994)	Population change	B	All sectors, Manufacturing, FIRE		
		Smith and Florida (1994)	Total population	B	Auto-related manufacturing		
		Coughlin and Segev (2000)	Total population	B	Foreign-owned manufacturing		
		Armington and Acs (2002)	Population change	B	All sectors, Business services, Distribution, Extraction, Local Market, Manufacturing, Retail		
		Sutaria and Hicks (2004)[2]	Population change	B	Manufacturing		
		Holl (2004a, 2004b, 2004c)	Total population	B	Manufacturing		
		Manzato, et al (2010)	Total population	C		Office firms	
2	Race	Smith and Florida (1994)	% of minority population	B	Auto-related manufacturing		
		Coughlin and Segev (2000)	% African American	B	Foreign-owned manufacturing		
		Chatman et al. (2016)	% African American	B		7 sectors (NAICS 2-digits code)	
3	Unemployment	Audretch and Fritsch (1994b)	Unemployment rate	B	(All sectors, Manufacturing, Services)	Manufacturing, Services	All sectors
		Reynolds (1994)	Unemployment rate	B	All sectors		Manufacturing, FIRE
		Coughlin and Segev (2000)	Unemployment rate	B			Foreign-owned manufacturing
		Armington and Acs (2002)	Unemployment rate	B	Business services, Distribution, Local Market, Manufacturing, Retail		All sectors, Extraction
		Sutaria and Hicks (2004)[2]	Unemployment rate	B			Manufacturing
		Audretch and Fritsch (1994b)	Change in Unemployment	B	All sectors, Manufacturing, Services		
		Reynolds (1994)	Change in Unemployment	B	Finance, insurance, and real estate (FIRE)	Manufacturing	All sectors
		Sutaria and Hicks (2004)[2]	Change in Unemployment	B		Manufacturing	

Table 1. Continued…

No.	Explanatory Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	(-) Effect	Statistically Insignificant Effect
5	Income	Reynolds (1994)	Per capita income change	B	All sectors, Manufacturing		FIRE
		Smith and Florida (1994)	Mean employee wage	B	Auto-related manufacturing		
		Coughlin and Segev (2000)	Total personal income	B	Foreign-owned manufacturing		
		Armington and Acs (2002)	Per capita income change	B	All sectors, Business services, Local Market		Distribution, Extraction, Manufacturing, Retail
		Sutaria and Hicks (2004)[1]	Per capita income change	B			Manufacturing
		Holl (2004a, 2004b, 2004c)	Mean manufacturing wage	B			Manufacturing
		Chatman et al. (2016)	Median HH income	B		7 sectors (2-digits NAICS[2] code)	
		Manzato, et al (2010)	Mean population income	C	Office firms		
		Kronenberg (2012)	Average daily salary	R	All sectors, Manufacturing, Services		
6	Education	Reynolds (1994)	% of population over 23 with college degree	B	FIRE	All sectors, Manufacturing	
		Smith and Florida (1994)	% of total population with HS or higher degree	B	Auto-related manufacturing		
		Coughlin and Segev (2000)	% of population over 25 with HS or higher degree	B	Foreign-owned manufacturing		
		Armington and Acs (2002)	% of adults w/o HS degree	B	All sectors, Distribution, Extraction, Local Market, Manufacturing, Retail		Business services
		Armington and Acs (2002)	% of adults with college degree	B	All sectors, Distribution, Extraction, Local Market, Retail		Business services, Manufacturing
		Holl (2004a, 2004b, 2004c)	% of labor force with higher education	B	Manufacturing		

[1] Sutaria and Hicks (2004) study yields different effect when using the ecological approach to standardize firm birth instead of the labor market approach.
[2] The North American Industry Classification System

Table 1. Continued…

No.	Explanatory Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	(-) Effect	Statistically Insignificant Effect
8	Property value	Reynolds (1994)	Median dwelling value	B	All sectors, Manufacturing, FIRE		
		Chatman et al. (2016)	Median rent	B	7 sectors (NAICS 2-digits code)		
9	Property ownership	Reynolds (1994)	% of owner-occupied dwellings	B	FIRE		All sectors, Manufacturing
		Armington and Acs (2002)	Proprietors / labor force	B	Local Market, Manufacturing, Retail		All sectors, Business services, Distribution, Extraction
11	Market productivity	Audretch and Fritsch (1994b)	Per capita value added	B	All sectors, Manufacturing		Services
		Coughlin and Segev (2000)	Per capita value added	B			
12	Financial capital	Sutaria and Hicks (2004)[1]	Per capita local bank deposits	B	Manufacturing		
13	Unionization	Smith and Florida (1994)	Count of auto-related unions	B		Auto-related manufacturing	
		Coughlin and Segev (2000)	% of unionized employee	B			Foreign-owned manufacturing

[1] The study used the ecological approach only to standardize firm birth.

2.2.2. Agglomeration economies (urbanization and localization economies)

Agglomeration economies play an important role in the analysis of firm dynamics. The underlying assumption of the theories of agglomeration economies is that the clustering of firms brings about higher cost saving or higher economic return to individual firms than if they were otherwise spatially segregated (Marshal, 1964; Porter, 1990; Anas et al., 1998). Localization and urbanization economies are the two distinct manifestations of agglomeration economies. Agglomeration externalities of the clustering of firms of the same industry are called localization economies (or specialization externalities), and agglomeration externalities of the clustering of firms of various industry sectors are called urbanization economies or diversity externalities (Anas et al., 1998; Beaudry and Schiffauerova, 2009). Existing research considers population density, industry specialization, average firm size, and firm age to analyze the impact of agglomeration economies on firm dynamics.

Firms may gain monetary or technological benefits when they agglomerate. Monetary agglomeration economies are the result of a reduction in the cost of inputs without a decrease in productivity when firms cluster. For example, the search for workers with specific skills is less costly for firms located in large cities or within large employment clusters; larger populations contain a greater share of skilled workers in the labor force compared to small cities with small populations. In contrast, since the labor force of a small city or small employment cluster may contain few workers with desired skills, the search for workers with a certain specialization may become more costly.

Firms can also reduce shipping costs by locating near markets and suppliers. Technological gains, on the other hand, are the result of a rise in the productivity of firm inputs because of clustering of firms without a corresponding reduction in input costs. This is a unique characteristic of high-tech and research and development (R&D) firms. The rise in productivity is attributed to higher knowledge spillover and an increase in competitiveness across firms and workers. Information on new or innovative technologies and know-hows may be shared in informal face-to-face meetings between employees of different firms. The larger or more localized the employment clusters, the higher these externalities tend to be (Brueckner, 2011).

Despite the large body of theoretical literature on agglomeration economies and productivity, there is little empirical evidence establishing the link between agglomeration economies and the patterns of firm dynamics combined (i.e. firm birth, closure, and relocation). Past empirical studies have generally focused on only one of these dynamics at a time to examine association with agglomeration.

The field of urban economics asserts the positive relationship among industrial agglomeration, economic externalities, and firm productivity (Reynolds, 1994; Armington and Acs, 2002; Cheng and Li, 2010; Cheng and Li, 2011). However, the influence of transportation investments on local agglomerations is ambiguous because of the limited empirical evidence. In addition, a distinction must be made between rail transit investment and highway investments when examining the influence of transportation investments on local agglomeration. Rail transit investments have the potential, through modal shift, to reduce diseconomies of agglomeration caused by road congestion, as greater traffic congestion consequently leads to the decentralization of firms (Wheaton, 2004).

Population density is the most commonly used explanatory variable to operationalize urbanization economies in empirical studies that examine firm dynamics. Across geographic units of analysis, higher population density implies higher levels of urbanization (i.e. population densities separate urban areas from suburban and rural areas). The empirical evidence is abundant regarding the positive association between population density and the rates of firm birth (Audretch and Fritsch, 1994b; Reynolds, 1994; Coughlin and Segev, 2000; Wennberg and Lindqvist, 2010; Chatman et al., 2016). Areas with high population densities have higher overall supply of labor and human capital, as well as higher demand for goods and services, consequently leading to more firm births.

Reynolds (1994), however, found instances of negative association between population density and firm birth. Using cross-sectional OLS regression, he examined firm birth rates of manufacturing and finance, information, and real estate (FIRE) industries within US labor market areas between 1986 and 1988. He found population density to have a positive influence on the rates of firm birth in the FIRE sector, whereas the influence was negative for manufacturing firms, arguing that low manufacturing firm births in densely populated areas reflect a continuing displacement of manufacturing companies toward lower cost (low density) regions away from urban centers. This finding is inconsistent with the findings of other empirical studies on firm birth, and also inconsistent with Krugman's (1991) theory that links population density to a higher concentration of economic activity.

The influence of industry specialization on firm birth is not straightforward. The literature on industrial location defines industry specialization as the concentration of firms of a certain industry sector, which is measured by either employment density or the number of firms in the industry sector. The findings of empirical studies on firm dynamics indicate that industry specialization in an area can have either a negative influence on firm birth (Holl, 2004a; Chatman et al., 2016) or a positive one (Smith and Florida, 1994; Armington and Acs, 2002). Smith and Florida (1994) examined the rates of firm birth of Japanese auto-related manufacturing firms within US counties and found that the percentage of labor force in auto-related manufacturing is positively associated with the number of births of Japanese auto-related manufacturing firms.

On the other hand, Chatman et al. (2016) examined firm birth within the census blocks of two U.S. cities, and found that the number of firms within own-industry has a negative relationship with firm birth. In other words, they found that the higher industry specialization leads to the lower rate of firm birth for any given industry sector. In general, a negative association between industry specialization and firm birth at the local level suggests that "inter-industry economies"[1] are more important than "within-industry economies,"[2] whereas a positive association would suggest otherwise. The geographic

[1] Also called 'Jacobsian economies' since the theoretical argument of interindustry economies inspired by Jacobs (1969) who identifies diversification externalities, highlighting the knowledge-spillover across firms in complementary sectors.
[2] Also called 'Marshallian economies' since the theoretical argument of within-industry economies was inspired by the seminal work of Alfred Marshall (1964).

scale of measurement (unit of analysis) often affects which of these two mechanisms is more likely to be found in empirical studies examining firm dynamics (the geographic scale of measurement is discussed in detail in section 2.4).

Agglomerations with firms of different sizes (in terms of number of employees) can have a different influence on firm birth. Smaller firms may benefit more from clustering than larger firms. For a given employment size, a smaller average firm size indicates the presence of a high number of firms (i.e. more small firms than large firms). A region with a large share of small firms indicates a higher presence of business owners, consequently stimulating entrepreneurship and contributing to information spillover (Reynolds, 1994; Chatman et al., 2016). Average firm size in an area is often calculated by dividing the total employment by the total number of firms (Audretsch and Fritsch, 1994b; Armington and Acs, 2002; Sutaria and Hicks, 2004).

Findings of empirical studies are not consistent about the association between average firm size and firm birth. Sutaria and Hicks (2004) found that average firm size within the Metropolitan Statistical Areas (MSAs) in Texas is positively associated with firm birth (i.e. large firms stimulate firm birth). There are two possible mechanisms that explain the positive association between average firm size and firm birth; either larger firms spin-off[1] entrepreneurs or new firms are established to service larger firms. On the other hand, Armington and Acs (2002) found a negative association between average firm size and the rates of firm birth within U.S. labor market areas (LMAs). That is, the larger the average firm size in an area, the lower the rate of firm birth.

[1] Spin-offs are new firms founded by employees of firms in the same industry (parent firms).

Two key standardization methods in these studies (i.e. the ecological or the labor market approach, as shown in section 2.1) generate different results regarding the direction of influence that average firm size has on firm birth. For instance, Audretsch and Fritsch (1994b) used a cross-sectional OLS regression model to examine the geography of firm birth within 75 regions in West Germany, and several of their control variables had opposing directions of influence on firm birth across the two standardization approaches. The poor explanatory power of linear regression models is another reason behind contradictory results of empirical studies on the determinants of firm dynamics (see section 2.4).

Besides the average size of firm, the average age of firms within a geographic unit of analysis can also influence the propensity of firms to thrive, fail, or relocate (Anas, et al., 1998). Smaller and younger firms are more likely to fail or relocate than bigger and more mature firms (Hayter, 1997). Kronenberg (2013), using logit regression, examined Dutch manufacturing and services firms within the 485 Netherlands municipalities in 2003. He found that higher average age of firms within a municipality negatively influences firm relocation patterns. Moreover, examining industry sector restructuring, Sutaria and Hicks (2004) also found no statistically significant relationship between the total earnings of service sector (representing the dominance of service industry within Texas MSAs) and the birth of manufacturing firms.

The presence of local universities and research institutions can have a positive influence on agglomeration economies. Academic institutions may attract firms to cluster because of the knowledge spillover benefits they provide (Audretsch and Feldman, 1996). Examining the birth and survival rates of newly formed firms, Wennberg and Lindqvist (2010) controlled for the number of medical and educational institutions present within Sweden's 87 labor market areas. Using time-series OLS method, they found no statistically significance influence for this control variable on either firm birth or firm survival. This control variable is, however, rarely used in the studies on the determinant of firm dynamics.

Table 2 provides a summary of the explanatory variables used in the reviewed empirical research to operationalize agglomeration economies. The columns in this summary table provide information on: (1) the explanatory variables, (2) the empirical study, (3) the variable measurement, (4) the firm dynamic under examination in each respective study (B=firm birth, C=firm closure, and R=firm relocation), and (5) the direction of influence (i.e. whether positive, negative, or statistically insignificant).

Table 2. Urbanization and Localization Economies' variables influencing the patterns of firm Dynamics

No.	Explanatory Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	(-) Effect	Statistically Insignificant Effect
1	Level of urbanization	Audretsch and Fritsch (1994b)	Population density	B	All sectors, Manufacturing, Services		
		Smith and Florida (1994)	Population density	B			Auto-related manufacturing
		Reynolds (1994)	Population density	B	FIRE	Manufacturing	All sectors
		Coughlin and Segev (2000)	Population density (9 dummy variables)	B	Foreign-owned manufacturing		
		Wennberg and Lindqvist (2010)	Population density	B	Manufacturing (4 sub-sectors)		
		Chatman et al. (2016)	Population and emp. densities	B	7 sectors (NAICS 2-digits code)		
		Kronenberg (2012)	Population density	R	All sectors, Manufacturing, Services		
2	Industry specialization or intensity	Reynolds (1994)	Index (industry workers/total workers)	B			All sectors, Manufacturing, FIRE
		Smith and Florida (1994)	% of labor force in manufacturing	B	Auto-related manufacturing[1]		Auto-related manufacturing
		Armington and Acs (2002)	Industry establishments/ population	B	All sectors, Business services, Distribution, Extraction, Local Market, Manufacturing, Retail		
		Holl (2004a)	LQ (share of manufacturing emp.)	B		Manufacturing	
		Chatman et al. (2016)	Number of firms in industry category	B		7 sectors (NAICS 2-digits code)	
		Kronenberg (2012)	An employment specialization index	R		All sectors, Services	Manufacturing
		Kronenberg (2012)	LQ (share of sector emp.)	R	All sectors, Services		Manufacturing

[1] Performing a Poisson model generated the positive effect in this study.

Table 2 Continued…

No.	Explanatory Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	**(-) Effect**	**Statistically Insignificant Effect**
4	Industry size structure	Audretsch and Fritsch (1994b)	Mean firm size (employment/firms)	B	(All sectors, Manufacturing, Services)[1]	All sectors, Manufacturing, Services	
		Reynolds (1994)	Share of small firms	B	All sectors, Manufacturing		FIRE
		Armington and Acs (2002)	Mean firm size (employment/firms)	B		All sectors, Distribution, Local Market, Manufacturing, Retail	Business services, Extraction
		Sutaria and Hicks (2004)	Mean firm size (employment/firms)	B	Manufacturing		
		Kronenberg (2012)	Mean firm size (employment/firms)	R		All sectors, Manufacturing, Services	
5	Industry age structure	Kronenberg (2012)	Average firm age	R		All sectors, Manufacturing, Services	
6	Industrial restructure	Sutaria and Hicks (2004)	Change in service share of total earnings	B			Manufacturing
7	Firm birth and closure	Sutaria and Hicks (2004)	Prior year firm birth rate	B	Manufacturing		
		Sutaria and Hicks (2004)	Prior year firm closure rate	B	Manufacturing		
		Chatman et al. (2016)	Number of firm closures	B	7 sectors (NAICS 2-digits code)		

[1] The sign of the effect changes when the study uses the ecological approach to standardize firm birth instead of the labor market approach.

2.2.3. Policy environment

Government policies and actions can directly (or indirectly) influence the rates of firm birth, firm closure, and firm relocation. Government policies that may directly influence the patterns of firm dynamics are interventions to guide urban development patterns through land use and zoning regulations. Urban development policies vary across US jurisdictions. Some jurisdictions are more proactive than others in pursuing denser and more compact urban development. For example, as early as 1960s, Montgomery County, Maryland, adopted a number of planning tools to promote a compact and dense urban development pattern. These policies impact the county's ability to attract business investment and economic growth (Knaap, et al., 2015), and subsequently influence firm birth, closure, and relocation patterns.

Indirect government policies can also influence the patterns of firm dynamics. Based on data availability, a few empirical studies use tax policy (or government spending policy) as a proxy for the policy environment within a region (Smith and Florida, 1994; Reynolds, 1994; Coughlin and Segev, 2000; Sutaria and Hicks, 2004). The relevant literature is ambiguous, however, about the influence of government taxation on the patterns of firm dynamics. For instance, Smith and Florida (1994) found property tax to have both positive and negative influence on firm birth depending on the used regression model. Their negative binomial model yields a negative association between property tax and firm birth, whereas the Poisson model yielded a positive association.

2.2.4. Spatial context

Several studies examining the patterns of firm dynamics use distance measures to represent the spatial context of firms within a region or urban area. The most common measure to depict spatial context, and the most familiar to urban planners, is the distance to central business district (CBD). Distance to CBD is often used as a proxy for the level of urbanization or accessibility. Theoretically, proximity to the city or regional center indicates higher urban density and accessibility, compared to less accessible suburban or rural areas. One can therefore assume that distance to CBD is related to agglomeration, and including distance to CBS as a control variable in the analysis can be considered double-counting. Unlike agglomeration economies, however, these distance measures are invariable across time. Therefore, in a polycentric region with several employment centers, distance to CBD as an accessibility measure is different from measures capturing accessibility to employment centers (agglomerations) over time. In addition to distance to CBD, a good proxy to capture the level of accessibility at the micro-level is the ratio of transit to auto accessibility (Chatman et al., 2016). High ratios of transit to auto accessibility are normally characteristic of denser urban areas, since they are often equipped with better public transport service compared to low density suburban areas.

Locations proximate to urban centers can be attractive for business investment and firm relocation, and also conducive to firm longevity. A limited number of studies have controlled for factors related to spatial context when analyzing firm birth, firm closure, or firm relocation. Chatman et al. (2016) found that proximity to CBD in Portland, Oregon and Dallas, Texas has a positive influence on firm birth. They also found a positive association between the ratio of transit to auto accessibility and the rates of firm birth. Manzato, et al. (2010) on the other hand, found that distance to major shopping areas within Netherlands municipalities has a negative relationship with the rates of firm closure; greater distances from shopping areas, they found, are highly correlated with firm closure. Similarly, Kronenberg (2012) found that municipalities that are farther away from the Netherland's economic center, Randstad, (a megalopolis in the central-western Netherlands and the economic center of the Netherlands) are less attractive for firms to relocate to. These studies, and others, highlight the importance of distance to city or urban centers in discussions on economic development and firm dynamics.

Table 3 and Table 4 provide summaries of the explanatory variables used in the reviewed empirical research to operationalize policy environment and spatial context. The columns in these summary tables provide information on: (1) the explanatory variables, (2) the empirical study, (3) the variable measurement, (4) the firm dynamic under examination (i.e. whether firm birth (B), firm closure (C), or firm relocation (R)), and (5) the direction of influence (i.e. whether positive, negative, or statistically insignificant).

Table 3. Government Policy Environment variables influencing the patterns of firm dynamics

No.	Exp. Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
					(+) Effect	(-) Effect	Statistically Insignificant Effect
1	Local taxation	Smith and Florida (1994)	Per capita personal and property tax	B	Auto-related manufacturing[1]	Auto-related manufacturing	
		Coughlin and Segev (2000)	Per capita property tax	B			Foreign-owned manufacturing
2	Local government spending	Reynolds (1994)	Gov. spending per capita	B	All sectors		Manufacturing, FIRE
		Sutaria and Hicks (2004)[2]	Gov. spending per capita	B			Manufacturing

Table 4. Spatial Context variables influencing the patterns of firm dynamics

Exp. Variables	Empirical Study	Variable Measurement	Effect on Firm Dynamics by Industrial Sector (B=Birth, C=Closure, R= Relocation Inward)			
				(+) Effect	(-) Effect	Statistically Insignificant Effect
Spatial context	Smith and Florida (1994)	Distance to the closest Auto-assembler; Distance to the biggest 3 Auto assemblers	B		Auto-related manufacturing	
	Chatman et al. (2016)	Distance to CBD	B		Seven industry sectors (NAICS 2-digits code)	
		Transit to Auto accessibility ratio	B	Seven industry sectors (NAICS 2-digits code)		
	Kronenberg (2012)	Distance to the center of the Randstad[3]	R		All sectors, Services	Manufacturing
	Manzato, et al. (2010)	Distance to the shopping areas	C		Office firms	

[1] Poisson model generated positive effect.
[2] The study used the ecological approach to standardize firm birth.
[3] Randstad is a central area in Netherlands, located between Amsterdam, Rotterdam, The Hague, and Utrecht.

2.2.5. Non-tangible determinants

Market conditions, accessibility, and agglomeration economies are not the only factors that may influence the patterns of firm dynamics. There are non-tangible factors (also known as soft factors) that can also influence the patterns of firm dynamics but can hardly be quantified. The soft factors are subjective in nature since they are related to the emotional and cultural preferences of entrepreneurs and business owners. For instance, an entrepreneur may choose a certain location to start a business mostly because of already established social capital or business ties. Data on soft factors and personal preferences are nearly impossible to objectively quantify since they are not necessarily supported by rational arguments (Risselada et al. 2012).

This research, therefore, focuses on tangible and quantifiable factors – highlighted throughout this chapter – to examine the determinants of firm birth, closure, and relocation patterns. Table 5 summarizes the explanatory variables considered in this research and their expected direction (sign) of impact on firm birth, closure, and relocation patterns. The following section provides a detailed account of transportation-related determinants of firm dynamics that are overlooked in the relevant literature and are thus the main focus of this research.

Table 5. Variables explaining the rates of firm birth, closure, and relocation

Determinants of Firm Dynamics (Birth, Closure, and Relocation)			Expected direction of impact		
Categories	**No.**	**Explanatory Variables**	**Birth**	**Closure**	**Net Relocation**
Market conditions	1	Unemployment rate	+/-	+/-	+/-
	2	Income	+	-	+
	3	Race (percentage African-American)	+/-	+/-	+/-
	4	Education level	+	-	+
	5	Property value (rent)	+	-	+
Urbanization and localization economies	6	Population density	+	-	+
	7	Industry specialization (share of firms per sector)	+	-	+
	8	Employment density	+/-	-	-
	9	Average firm size	+/-	-	-
	10	Average age of firms	N/A[1]	-	-
	11	Firm closure rates	+	N/A[2]	+
	12	Firm birth rates	N/A[3]	+	+
Policy environment	13	Property tax	+/-	+/-	+/-
Spatial context	14	Distance to CBD	+	-	+

Notes: 1. Firm closure and relocation analyses include the average age of firms as a control variable.
2. Firm birth and relocation analyses include the number of firm closures as a control variable.
3. Firm closure and relocation analyses include the number of firm births as a control variable.

2.3. Transportation infrastructure and firm dynamics

Transportation Infrastructure is vital to the economic growth of cities and regions (Chatman and Noland, 2011). Investments in rail transit, in particular, could increase the size, density, and diversity of cities, which could substantially contribute to increased economic productivity (Chatman, et al., 2016). The question of how transit investments influence the spatial dynamics of economic activities, however, remains vaguely and partially answered. Delineating the influence of transit investments on the patterns of firm dynamics is key to answering this question more comprehensively, and is the main objective of this research.

Research shows that transit accessibility plays a key role in linking firms to markets. Proximity to consumers and suppliers is a significant determinant of the location decisions of firms (Giuliano, 2004), and the spatial separation of producers (origins) and consumers (destinations) drives the demand for investment in transportation infrastructure. The earliest theoretical records of accessibility as a significant factor in the location decisions of firms emerged in the agricultural land rent theory by Von Thunen (1826). More than a century later, the monocentric urban theory by Alonso (1964) also shows that accessibility, which is measured by the distance to the central business district (CBD), plays a major role in shaping the urban spatial structure. According to new economic geography models, higher costs of transportation lead to more dispersed economic activities as firms need to supply dispersed markets (consumers) locally (Puga, 1999; Fujita, Krugman, and Venables, 1999; Ottaviano, Tabuchi, and Thisse, 2002). Reduced transportation costs, on the other hand, improves accessibility between markets leading subsequently to the centralization of economic activities due to: (1) access to larger and more diverse inputs (e.g. raw materials and labor), and (2) access to larger pool of consumers. The concentration of markets through improved transit accessibility offers advantages to firms due to agglomeration economies, as discussed in section 2.2.2.

Since improved transportation accessibility facilitates the centralization of economic activities, it can also influence changes in 1) the patterns of firm dynamics, 2) agglomeration economies and productivity, and 3) property values. Among these, firm dynamics are the least examined, whereas changes in property values are the most thoroughly studied aspect of transportation investments. At least two meta-analyses have been conducted to date that summarize the finding from dozens of empirical studies on the link between transportation and property values. Debrezion, Pels, and Rietveld (2007) and Mohammad et al. (2013) conducted a meta-analysis of studies that examine the connection between rail transit and changes in property values. They found that a consensus exists across empirical studies about the positive association between the presence of rail stations and commercial and residential property values. Market concentration and dispersion forces can impact transportation investments by reducing transportation costs for firms. The level of transportation infrastructure affects the costs that are incurred for firms to transport inputs and outputs. Transportation investments, therefore, influence the geographic extent of the market area that firms can access. Similarly, firms can realize agglomeration benefits from clustering through labor and input sharing and knowledge spillover (Armington and Acs, 2002; Beaudry and Schiffauerova, 2009).

These concentration and dispersion forces are not straightforward, however. Firms across different industry sectors may respond differently to changes in transportation costs. Firms with relatively high transportation costs may concentrate in response to a reduction in transportation costs. On the other hand, firms with relatively low transportation costs may disperse to take advantage of lower costs of labor and land rent (kilkenny, 1998; Holl, 2004b). Therefore, examining transportation impact across different industrial sectors is imperative to the analysis of firm dynamics.

The external factors must be accounted for when trying to determine the net influence of transportation investments on the patterns of firm dynamics, as discussed earlier in this chapter. These factors include the nature of pre-existing market conditions, urbanization and localization economies (agglomeration economies), and existing local regulations and policies, such as zoning and land-use plans and local taxation.

There are no coherent accounts of how rail transit infrastructure influences the patterns of firm dynamics. Most empirical studies on firm dynamics overlook control variables related to rail transit due to the difficulty of accounting for complex and interlinked transportation-related factors, such as type, scale, location, and the operating characteristics of rail transit infrastructure. The few existing studies that include transportation variables have examined only one aspect of firm dynamics, mostly firm birth (see Table 6).

Jointly examining the patterns of firm dynamics (i.e. firm birth, closure, and relocation patterns) is key to understanding the overall impact of rail transit systems on firm dynamics because the patterns of firm dynamics may interact differently across the units of analysis. Moreover, understanding the patterns of firm dynamics in urban area is key to understanding the overall growth or decline of its urban economy. For instance, high rates of firm birth in a certain area (e.g. an area near a rail station) does not necessarily indicate that this area is experiencing a net growth in firm agglomeration if the rate of firm birth is equal or less than the rates of firm closure. Rail transit variables are not considered in the studies that linked transportation and firm dynamics, where most of the emphasis is given to highway accessibility. Table 6 and Table 7 provide summaries of the studies that take into account transport variables when examining the patterns of firm dynamics. The tables summarize these studies by the following factors: (1) the dependent variable and how it is measured, (2) the transportation-related variables used as control, (3) the units of analysis (geographic resolution), (4) the analyzed industry sectors and the period, (5) the statistical method, and 6) the impact of the transportation-related variable.

The inclusion and operationalization of transportation infrastructure factors in the analyses of the patterns of firm dynamics varies considerably in the literature. Previous studies used a variation of binary and continuous measures of transportation, often related to the availability of highways, to examine the influence of transportation accessibility on the patterns of firm dynamics. Other studies used a binary variable to indicate whether or not highways are present within the examined units of analysis, such as Smith and Florida (1994) and Coughlin and Segev (2000). To improve on this crude measurement, later studies used the distance to nearest highways (Holl, 2004a, 2004b, 2004c), as well as the distance to nearest rail stations, to account for the availability of transportation networks (Manzato, et al., 2010; Nguyen et al., 2013; Risselada, Schutjens, and Van Oort, 2013 ; An, Kang, and Lee, 2014; Chatman, Noland, and Klein, 2016).

Table 6. Summary of past studies on firm birth that include transportation-related variables in their analysis

Empirical Study	Dependent Variable	Transport Variables	Geographic Resolution	Industrial Sectors (Time Period)	Method	Impact of Transport Variables
Firm Birth						
Smith and Florida (1994)	Number of firm births	Presence of an interstate highway (dummy variable)	US counties	Japanese-affiliated manufacturing automotive-related industries (1990)	Cross-sectional, using Tobit, Poisson, and Negative binominal models	Positive and significant
Coughlin and Segev (2000)	Number of firm births	Presence of an interstate highway (dummy variable)	US counties	Two-digit SIC foreign-owned manufacturing (1989-1994)	Cross-sectional, using negative binominal models (8 regions dummies)	Positive and significant
Chatman, Noland, and Klein (2016)	Number of firm births	**Distance to nearest rail station** (includes dummies for 0.25, 0.25-0.5, 0.5-1 mile thresholds); distance to nearest highway exit	Census blocks of Dallas and Texas metropolitan areas	Two-digit NAICS industry sectors (seven broad categories by authors) (1991-2008)	Cross-sectional, Negative binominal count model	Positive and significant (coefficient sizes are larger in Portland compare to Dallas for most industry sectors)
Holl (2004a)	Number of firm births	Distance to nearest motorway (10km intervals up to 50km and >50km)	Spain municipalities	Manufacturing: 10 sub-sectors (1980-1994)	Fixed-effect Poisson with time dummy	Positive and significant
Holl (2004b)	Number of firm births	Distance to nearest motorway (10km intervals up to 50km and >50km)	All 275 Portugal municipalities	Manufacturing: 13 sub-sectors; Services: 9 sub-sectors (1986-1997)	Fixed-effect Negative Binominal model with time dummy	Positive and significant
Holl (2004c)	Number of firm births	Distance to nearest motorway (10km intervals up to 50km and >50km)	All 275 Portugal municipalities	Manufacturing: 13 sub-sectors; Services: 9 sub-sectors (1986-1997)	Fixed-effect Negative Binominal Model and Poisson with time dummy	Positive and significant
Melo, Graham,and Noland (2010)	Number of firm births	**Density of railway network**; density of motorway network	Portuguese municipalities	Five industrial sectors: 1) primary industries; 2) manufacturing; 3) electricity, gas, and water; 4) construction; 5) wholesale and retail (1995 and 2003)	Negative Binomial Models	Positive and significant
Bacher and Brulhart (2013)	Number of firm births	Distance to the nearest highway access, distance to the nearest airport	Swiss municipalities	Forty six sectors based on two-digit level of Eurostat's NACE classification	Fixed effects Poisson regression	Positive and significant

Table 7. Summary of past studies on firm closure, survival, and firm relocation that include transportation-related variables

Empirical Study	Dependent Variable	Transport Variables	Geographic Resolution	Industrial Sectors (Time Period)	Method	Impact of Transport Variables
Firm Closure and Survival						
Manzato, et al (2010)	Firm closure rates	Distance to nearest airport, **nearest rail station, nearest high-speed train station**, nearest highway exit	Netherlands municipalities and provinces	15 Dutch office industry sectors (1996 - 2006)	Parametric duration models	Negative and significant except for distance to nearest high-speed train station
Cader and Leatherman (2011)	Firm survival	Presence of interstate highway	All 105 Kansas counties, U.S.	IT-producing, Goods-producing, and Service-producing industries (1990-2003)	Two-step OLS and proportional hazard model	*Two-step OLS*: Negative and significant for IT-producing industry. Positive but insignificant for the other two industries. *Proportional hazard model*: Negative and significant for all industries.
Firm Relocation						
An, Kang, and Lee (2014)	Probability of firm to relocate in (binary)	Distance to nearest expressway; **distance to nearest rail station; distance to nearest subway station**; distance to airport; distance to harbor; distance to main road; density of bus-line	Seoul Metropolitan Area divided into 300-meter-wide hexagons (158,453 hexagons total)	Manufacturing and service firms (2006 - 2011)	Binary logit model	Subway station: negative and significant
Nguyen et al. (2013)	Probability of firm to relocate out (binary)	Distance to highway, **distance to train station**	17 regions and 335 zones in Tokyo Metropolitan Area	Manufacturing and retail firms (1994)	Binary logit mode	Highway: positive and significant for manufacturing but insignificant for retail; Train station: insignificant for manufacturing but positive and significant for retail
Risselada, Schutjens, and Van Oort (2013)	Probability of firm to relocate out (binary)	**Distance to train station**; distance to freeway	Urban residential neighborhoods of the municipality of Amsterdam, Netherland - 6-digits postal codes	All industrial sectors: (2005 - 2008)	Binary logistic models	Marginal to insignificant for both train station and freeway.

This section provides a comprehensive review of the empirical studies that consider transportation-related factors when examining the patterns of firm dynamics. It reviews existing empirical studies with respect to their methodologies, the considered industrial sectors, and their findings. Most of the limited number of studies indirectly examine transport infrastructure in relation to the patterns of firm dynamics. Only a few studies have transportation-related factors as the focal point of the analysis when examining the patterns of firm dynamics (i.e. firm birth, firm relocation, or firm closure or survival). Firm birth, firm survival, and firm closure are discussed jointly in the following section because they are interconnected, and because there are no empirical studies on firm closure that account for transportation-related factors.

2.3.1. Rail Transit and Firm Birth, Closure, and Survival

Firm birth has a positive influence on the economic growth of a region. Job creation and changes in economic structure are the most notable positive externalities of firm birth. While the empirical research on the determinants of firm birth is abundant, a limited number of studies have examined the link between proximity to transportation infrastructure and the number of firm births (Smith and Florida, 1994; Coughlin and Segev, 2000; Holl, 2004a, 2004b, 2004c; Melo, Graham, and Noland, 2010; Chatman, Noland, and Klein, 2016). Most of these studies account for proximity to highway infrastructure but fail to account for proximity to rail transit infrastructure (see Table 6). There are only two studies that take into account proximity to rail station as an explanatory variable in their analysis of firm birth (Melo, Graham, and Noland, 2010; and Chatman, Noland, and Klein, 2016). Relevant empirical research indicates a positive connection between access to rail transit and firm birth but fails to consider other aspects of firm dynamics such as firm closure or survival.

These limited number of studies predominantly found a positive relationship between the availability of transportation infrastructure and firm birth. The proximity to, or the presence of, highways are examined in all these studies and typically in relation to firm birth in the manufacturing sector (see Table 6). Studies that examined firm birth in relation to highway systems found that the closer the distance to highway exits, the higher the rates of firm birth (Smith and Florida, 1994; Coughlin and Segev, 2000; Holl, 2004a, 2004b, 2004c). As stated earlier, there is a lack of empirical evidence on the association between the availability of passenger rail infrastructure and firm birth.

In general, the distance to rail station is the most commonly used method to operationalize the availability of passenger rail infrastructure (Manzato, et al., 2010; Chatman, Noland, and Klein, 2016). Nonetheless, only one study directly examines the association between areas within short distances to passenger rail stations and firm birth. Chatman, Noland, and Klein (2016) used dummy variables to indicate the presence or non-presence of higher numbers of firm birth within specific distance intervals from rail stations in Portland, Oregon and Dallas, Texas metropolitan areas. Another less popular method is the level of density of rail networks within certain geographic units of analysis. Melo, Graham, and Noland (2010) used the density of rail networks within Portuguese municipalities to examine its impact on firm births. In both studies, the finding suggests that there is a positive and statistically significant connection between the availability of rail transit and the number of firm births.

There is more to the dynamic process of urban economics than firm birth. In a particular area, number of firm births is not a sufficient indication, per se, of a net economic gain or loss. Firm births can merely be the result of the closure of existing firms in a process called "creative destruction," coined by the classical economist Joseph Schumpeter (Schumpeter, 1934). Moreover, the process of firm birth and firm closure can be considerably heterogeneous across different geographical areas. In other words, existing firms go out of business (i.e. fail to make profit or to compete with existing firms) and new firms emerge disproportionally across different geographical areas. Therefore, it is important to examine both firm birth and firm closure to determine the net influence of transportation infrastructure (such as passenger rail system) on firm birth relative to firm closure. However, studies that include transportation variables in firm birth analysis only consider the number of firm births as an outcome variable (Holl, 2004; Smith and Florida, 1994; Coughlin and Segev, 2000; Chatman, Noland, and Klein, 2016). These studies examined whether the presence of (or the proximity to) transportation infrastructure (mostly highways) increases the probability of firm birth. They consistently found the association to be positive and significant between proximity to transport infrastructure and the number of firm births. These findings, however, are inadequate to indicate the net influence of transportation infrastructures on firm birth without examining their influence on firm closure.

2.3.2. Rail Transit and Firm Relocation

Despite the importance of information on firm relocation to providing accurate analyses of firm dynamics, there is very limited research on the subject, and even less research on the influence of rail transit on firm relocation decisions. Unlike firm birth and closure, firm relocation explicitly accounts for the decision of firms to substitute one location with another. Certain external and internal factors may influence firms to relocate. As discussed in section 2.2, firms constantly adjust to new circumstances due to changes in market conditions, urbanization and localization economies, government policies and regulations, or other non-tangible factors. Other changes in firm-specific characteristics, such as size, age, sector, and growth patterns, may also lead to changes in the locational preferences of firms.

In the relevant literature, firm relocation decisions remain the most understudied filed of research. The reason for this gap is likely the lack of datasets that accurately capture firm relocation patterns. In recent years, only three studies have examined the probability of firms to relocate in relation to firms' proximity to rail station (An, Kang, and Lee, 2014; Nguyen et al., 2013; Risselada, Schutjens, and Van Oort, 2013). Table 7 provides a structured summary of these studies including dependent variables; transportation variables; geographic resolutions; industrial sectors; methodologies; and predicted impacts of transportation variables. None of these studies are focused on the U.S. economy, however. Moreover, these studies specifically examine the probability of firms to either relocate out or to relocate in. Using binary logit model, Nguyen et al. (2013) examined factors influencing the probability of firms to relocate out (push factors) or stay (pull factors[1]) within 17 regions and 335 zones (following the zone system used in their dataset) in Tokyo Metropolitan Area. They found that the distance to the nearest

[1] Pull factors and push factors are not mutually exclusive factors influencing firm relocation. covariates can be both a push and a pull factor at the same time. For instance, population density can negatively influence some firms (push factor) and at the same time positively influence others (pull factors).

train station is negatively associated with the decision of retail and manufacturing firms to not relocate out (i.e. areas within shorter distances to the rail stations were more likely to deter outward firm relocation than areas farther away from the stations), but the results were statistically insignificant for manufacturing firms. Their analysis suggests that retailers are more likely to take advantage of rail transit systems than manufacturers. However, their analysis of whether or not firms are attracted to these zones (inward relocation) did not include the distance to the nearest station variable.

Risselada, Schutjens, and Van Oort (2013) also examined firm relocation push factors within urban residential neighborhoods of the municipality of Amsterdam, Netherlands (6-digits postal codes) using binary logit model. They found proximity to rail stations had no statistically significate influence on firms' relocation decisions. Again, their study focuses on only one aspect of firm relocation dynamics, that is the probability of firms to move or stay within the geographic units of analysis. On the other hand, An, Kang, and Lee (2014) examined the location choice factors of relocating manufacturing and service firms within Seoul Metropolitan Area divided into 300-meter-wide hexagons, using binary logit model once again. The authors found that areas within short distance to subway stations had a positive association with the probability of inward firm relocation for service firms, but the same areas had a negative association with the probability of inward firm relocation for manufacturing firms. In contrast to the Nguyen et al. (2013) study, An, Kang, and Lee (2014) analysis is focused only on pull factors (i.e. the probability of inward firm relocation).

As shown throughout this subsection, studies on firm relocation have focused on the impact of certain factors (only a few considering the proximity to rail stations as a factor) on either inward firm relocation or outward firm relocation. It is important, however, to examine both of these factors within firm dynamics (inward and outward firm relocation) at the same time in order to understand the net firm relocation effects within a given unit of analysis. More importantly (to the purpose of this research), both inward and outward firm relocation patterns must be examined to determine the influence of rail stations on net firm relocation (by comparing inward firm relocation effects to outward firm relocations effects). This research fills the gap in literature by examining the impact of areas within short walking distances to passenger rail stations on net firm relocation.

2.4. Methodological aspects

Empirical studies take different approaches to examine the determinants of firm dynamics. The findings are often inconsistent and sometimes even contradictory regarding the influence rail transit accessibility have on the patterns of firm dynamics, even among the few studies that account for transportation-related factors. The inconsistencies are attributed to differences in: (1) included explanatory variables, (2) model of analysis, (3) geographic scope and units of analysis, and (4) industrial sectors considered. The previous sections of this chapter covered in detail a review on the explanatory variables that studies consider for the analysis of firm dynamics. The following sections review models, units of analysis, and industrial aggregation analyses of the previous studies on the patterns of firm dynamics.

2.4.1. Models

Earlier studies on the determinants of firm birth modeled the dependent variable as the rate of firm births relative to labor force or relative to existing firms. To generate a continuous dependent variable suitable for OLS or fixed-effect regressions, many of the past studies calculate the rate of firm birth using either firm births per population (or size of labor force) or firm births per existing firms (see Table 8). The standardization of the dependent variable allowed simple OLS or fixed effect panel regressions to be applied (Audretsch and Fritsch, 1994b; Reynolds, 1994; Armington and Acs, 2002; Sutaria and Hicks, 2004).

Empirical research on firm survival is abundant in the literature, whereas empirical evidence on firm closure is scarce. In particular, little empirical evidence exists regarding the association between transportation-related factors and firm closure or survival (see Table 7). Firm survival requires a different modeling technique compared to the occurrence of firm birth, closure, and relocation. Duration modeling, therefore, is the most common method for examining firm survival. Hazard models[1] are the common mathematical models used to analyze survival events (Table 9 provides a summary of a few studies on firm survival). Manzato et al (2010) use a duration hazard model to examine office firms within municipalities and provinces of Netherlands. They found that the proximity to intercity rail stations increases the probability of office firm survival. No

[1] Hazard models are statistical models used to examine the association between the survival time of certain events/objects and one or more explanatory variables. An example is Cox proportional-hazards model.

studies, however, are found to have examined firm closure in relation to transportation-related factors.

As a dependent variable, firm birth, closure, and relocation are better modeled as count measure while firm survival is more suitably modeled as the duration of survival of firms. Since firm births are discrete events, more recent studies assume that the process of firm birth follows a Poisson or a negative binominal distribution (see Chapter 3 for detailed discussion), which makes count models more suitable for firm birth analysis (Holl, 2004a, 2004b, 2004c; Melo et al., 2010; Chatman et al., 2016).

Table 8. Methodological aspects of past studies that examine the determinants of firm birth

Author	Dependent Variable	Geographic Unit of Analysis	Industrial Sectors (Time Period)	Method
Audretsch and Fritsch (1994b)	1) Rate of firm births relative to existing firms; 2) Rate of new firms relative to labor force	75 regions in west Germany	All industries, manufacturing industry, and services (1986)	Ordinary least-squares regression (OLS)
Reynolds (1994)	1) Rate of firm births relative to existing firms; 2) Rate of new firms relative to labor force	US metropolitan areas; manufacturing rural areas; and traditional rural areas	All sectors, manufacturing, and FIRE (1986-1988)	Cross-sectional OLS regression
Armington and Acs (2002)	Rate of firm births relative to labor force	US labor market areas (LMAs) (394)	4-digit SIC industries1 (1994-1996)	Cross-sectional OLS regression
Sutaria and Hicks (2004)	Annual rate of new firms relative to existing firms	All 27 Metropolitan Statistical Areas (MSAs) in Texas, US.	Manufacturing (1976-1991)	Cross-sectional, OLS and Fixed-effect regression models
Cheng and Li (2010)	Rate of firm births relative to labor force	All U.S. counties	Two-digit NAICS industry sectors (2001-2003)	Geographically weighted regression (GWR)
Cheng and Li (2011)	Rate of firm births relative to labor force	All U.S. counties	Ten industrial categories based on the 2-digit NAICS code (2001-2003)	OLS, GWR, and spatial error model (SEM)
Lasch, Robert and Le Roy (2013)	Number of firm births	The 348 labor market areas (LMAs)	Information and communication technologies (ICT) sector (1993-2001)	Multivariate regression model
Brixy and Grotz (2007)	Rate of firm births relative to labor force	West German regions	All industries, manufacturing industry, and business services (1987-1997)	Fixed effects panel regression
Bosma, van Stel, and Suddle (2008)	Number of new independent start-ups and the number of new subsidiaries	Forty regions in the Netherlands	Manufacturing and service industries (1988-2002)	Linear regression SUR (seemingly unrelated regression)

Empirical research on the patterns of firm relocation is sparse. Most of the previous studies on firm relocation have used binary logit models (Kronenberg, 2012; An, Kang, and Lee, 2014; Nguyen et al., 2013; Risselada, Schutjens, and Van Oort, 2013) to examine the likelihood that firms will relocate (see Table 9). Among these studies, only three account for proximity to rail transit stations. Studies on firm relocation examine the likelihood of firms to either relocate inward (An, Kang, and Lee, 2014) or to relocate outward (Risselada, Schutjens, and Van Oort, 201; Nguyen et al., 2013) but not both (see Table 7). The dependent variable in relocation studies is modeled as a binary variable (i.e. the dependent variable equals "1" if firms relocate and "0" if they do not). Both relocation patterns must be examined, however, to find out the net impact of proximity to rail transit on firm relocation patterns (i.e. the patterns of inward firm relocation relative to outward firm relocation). If proximity to rail stations negatively influences inward relocation (as found by An, Kang, and Lee, 2014), this negative association does not necessarily indicate that subway stations are repulsive to firms (push factor), but it is likely that areas near rail stations retain more firms from relocating outward than attract relocating firms to locate within. This research, therefore, examines inward and outward relocation patterns to accurately determine the net relocation effects of areas near passenger rail stations.

Table 9. Methodological aspects of past studies that examine the determinants of firm relocation or firm closure and survival

Author	Dependent Variable	Geographic Unit of Analysis	Industrial Sectors (Time Period)	Method
Firm Closure and Survival				
Wennberg and Lindqvist (2010)	Duration of new firm survival	Sweden's 87 labor market areas	Telecom and consumer electronics, financial services, IT, medical equipment, and biopharmaceutical industries (1993 - 2002)	Event history analysis.
Ace, Armington, and Zhang (2007)	New-firm survival rate	The U.S. Labor Market Areas (LMAs)	Service sector (1990 -1998)	Cross-sectional, ordinary least square linear (OLS) regression
Brixy and Grotz (2007)	Survival rate	West German regions	Manufacturing, business services, and all industries (1981 - 1997)	Panel regression with fixed effects
Firm Relocation				
Sleutjes and Volker (2012)	Probability of firm to relocate	145 neighborhoods within 40 Dutch municipalities	Commercial firms (2008)	Poisson regression model
Kronenberg (2012)	firm relocation (binary)	The 485 Netherlands municipalities in 2003	Dutch manufacturing and services firms (2002-2003)	Two-stage nested logit regression
Sleutjes and Beckers (2013)	Probability of firm to relocate	Five neighborhoods in three Dutch cities	Chamber of Commerce's classification of 11 industrial sectors (2005 - 2007)	Qualitative and descriptive analysis (50 in-depth interviews with entrepreneurs)

Moreover, several cross-sectional studies have generated results that differ substantially from time-series studies. For example, Sutaria and Hicks (2004) examined birth rates of manufacturing firms by using four fixed effect models to assess the contribution of time-related effects (i.e. between 1976 and 1990) and/or location-related effects (i.e. across 27 Metropolitan Statistical Areas in Texas, US). They found substantial differences in their models regarding the overall explanatory power and the direction of impact of several of the explanatory variables depending on the inclusion or exclusion of one or both of time and location fixed effects[1] (i.e. unspecified year and region dummy variables). Cross-sectional models, therefore, tend to yield unreliable results because they fail to account for changes in the patterns of firm dynamics over time. A panel model structure is essential to capture over time changes in firm dynamics. Moreover, none of the past studies account for the endogeneity of the placement of rail stations when examining firm birth, closure, or firm relocation. Chapter 3 provides details on why it is important to account for the endogeneity of rail systems and shows how this study accounts for it.

2.4.2. Units of analysis

As mentioned earlier, a great deal of spatial variation tends to occur in areas within close proximity to transportation infrastructure, especially in the case of rail transit. The use of macro geographic units of analysis in regression models generates unreliable findings when examining the connection between transportation networks and the patterns of firm

[1] The authors do not specify what fixed effect variables were used in their regression models.

dynamics, especially in the case of rail transit network. Unlike road networks, rail transit networks tend to be spatially scattered and often accessed by walking. Therefore, the use of smaller geographic units of analysis in regression models is necessary to accurately determine the association between areas within short walking distance of rail transit stations and the patterns of firm dynamics.

In the reviewed empirical studies, the geographical units of analysis are often too large to account for micro-level spatial variation in the number of firm births (see Table 8). The analyses are often conducted at the county level in the U.S. (Smith and Florida, 1994; Coughlin and Segev, 2000) or at the municipal level in European countries (Holl, 2004a, 2004b, 2004c; Melo, Graham, and Noland, 2010; Bacher and Brulhart, 2013). Only three empirical studies have used geographic units of analysis small enough to adequately capture spatial variation of a pattern of firm dynamics in relation to rail transit infrastructure (Risselada, Schutjens, and Van Oort, 2013; An, Kang, and Lee, 2014; Chatman, Noland, and Klein, 2016).

Chatman, Noland, and Klein (2016) use census blocks within the metropolitan regions of Dallas, Texas, and Portland, Oregon, to examine the connection between proximity to rail stations and firm birth. They found proximity to rail transit stations to have a positive influence on firm birth, and the influence was stronger in Portland than it was in Dallas. The authors attribute the difference in the influence of rail transit on firm birth between the two cities to Dallas's lower transit usage, higher off-street parking requirements, and poor policies toward densification near rail stations. An, Kang, and Lee (2014) divided Seoul Metropolitan Area into 300-meter wide hexagons to examine the probability of relocating firms to locate within these hexagons. They found proximity to rail stations to have a positive influence on the probability of firms to relocate within.

2.4.3. Industry sector

Across industry sectors, several control variables may have different influence on firm birth, closure, and relocation patterns, as discussed in the second section of this chapter. For instance, firms in service industry are expected to be drawn to densely populated areas to take advantage of local market conditions, whereas manufacturing firms may be deterred by high density due to the associated higher costs of labor and property. It is, therefore, imperative to examine the patterns of firm dynamics across multiple industry sectors to strengthen the body of knowledge and improve the level of understanding of how firms across different industry sectors are influenced by rail transit investments.

2.5. Chapter summary

Based on the literature review, this chapter identified factors that potentially influence the patterns of firm birth, firm closure, and firm relocation. The influential factors of firm dynamics were divided into five categories: (1) market conditions, (2) agglomeration economies, (3) policy environment, (4) spatial context, and (5) transportation-related factors. This chapter identified several key gaps in the past empirical studies on the determinants of firm dynamics with special attention to the studies that included in their analysis factors related to rail transit investments. First, most of the empirical research is focused on the determinants of firm birth, and there is limited empirical evidence on the subjects of firm closure and firm relocation patterns. Second, a considerable number of the previous studies used regression models that are not appropriate for the analysis of the patterns of firm dynamics. The regression models of these studies were either cross-sectional or the dependent variable (e.g. number of firm births) were standardized to allow for simple OLS or fixed effect panel regressions. Third, most of the empirical studies fail to capture the patterns of firm dynamics at the micro-level since the units of analysis used in the regression models of these studies are too large (e.g. counties, cities, or regions). Finally, there is little empirical research that examines the association between transportation related factors and the patterns of firm dynamics across different industry sectors because the previous research mostly focused on all sectors combined or only manufacturing sector.

The following chapter (Chapter 3) discusses the statistical methods that are most appropriate to analyze the count of firm births, closures, inward relocations, and outward relocations (outcome variables) at a micro level. Two regression methods are discussed in Chapter 3 that are used for the analysis: (1) a standard negative binomial regression, and (2) a negative binomial regression that controls for the endogeneity of the placement of rail stations. Chapter 3 also provides a description of the influential factors (control variables) that are considered in the analysis of this dissertation, which were identified throughout this chapter.

CHAPTER 3: DATA, VARIABLES, AND METHODOLOGY

In empirical research, the selection of one method of analysis over another has led to different conclusions over the determinants of firm birth, firm closure, or firm relocation patterns. Further, there is a lack of empirical evidence on the patterns of firm dynamics other than firm birth. As explained in Chapter 2, similar explanatory variables have offered inconsistent results about the influences on the patterns of firm dynamics across empirical studies, leading to different conclusions. Studies on the determinants of firm dynamics have often applied ordinary least squares (OLS) regression models to examine a set of explanatory variables; data used for these methods are either cross-sectional or measured over a relatively brief period. Additionally, limitations in the quality of data has restricted the number of variables considered in the past empirical studies that examine the patterns of firm dynamics. Limited data quality has restricted how these studies operationalized the dependent and independent variables, as discussed throughout Chapter 2. The distinction between firm birth and relocation has not been made in most of these studies. These shortcomings caused ambiguity in the literature on the subject regarding the magnitude and the direction of impact that relevant explanatory variables have on firm birth, closure, or relocation patterns.

This research contributes to the literature in the field of firm dynamics by conducting a comprehensive examination of the determinants of firm birth, closure, and relocation patterns using a large and more detailed dataset. This chapter is divided into three sections. The first section provides a detailed description of the dataset and units of analysis used in this study to examine the determinants of firm dynamics. The second section provides a discussion on how the outcome and control variables are structured to operationalize firm dynamics and their determinants. The third section provides a detailed description of the statistically controlled methods used in this study to examine micro-level firm dynamics as a function of proximity to passenger rail stations, agglomeration, socio-economics, and spatial context.

3.1. Data and units of analysis

This study uses National Establishment Time-Series (NETS) dataset to construct the dependent variables. NETS database offers the advantage of a detailed account of dynamics of the U.S. economy. It was made available when Walls & Associates teamed up with Dun and Bradstreet (D&B) to convert their archival establishment data into a time-series database of establishment information (Walls, 2008). NETS microdata is a reliable data source for studying static business activity in high detail (Barnatchez, Crane and Decker, 2017). Relevant to the purpose of this research, NETS database contains information on the first and last year when each firm existed and the industry sector to which it belongs (NAICS classification).

NETS database also distinguishes between firm relocation and firm birth by providing information on the previous location of establishments (longitude and latitude) in addition to information on the current location to capture relocation. NETS database provides the latitude and longitude of firm locations at multiple geographic levels, ranging from Census block to Zip Code levels. However, the longitude and latitude of most firms are provided at the block level (see Appendix B for more details). Another important advantage of Dun and Bradstreet database is that it assigns a unique identifier to each establishment (called a DUNS-number) that is retained over time even if an establishment relocates. Additionally, the U.S. census data and GIS shapefiles are obtained from the National Historical Geographic Information System (NHGIS). NHGIS provides U.S. census socio-demographic data along with GIS-compatible boundary files from year 1790 to the present (Manson et al., 2017). The GIS shapefiles are originally from the U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) shapefiles, which do not include demographic data but do contain geographic identification codes (GEOIDs) that can be linked to the Census Bureau’s demographic data.

With respect to the geographic unit of analysis, this research uses the U.S. Census blocks from the year 2000 to conduct the micro-level analyses. The 2010 Census block shapefiles were initially considered for the analysis, but ultimately replaced by the 2000 Census block shapefile because of the considerable number of unqualified Census blocks in the 2010 shapefile. The U.S. Census Bureau’s 2010 Census block shapefile includes road and railway polygons as Census blocks but does not identify them as such, since they were automatically generated from satellite imagery. Using the 2010 Census block shapefile causes the problem of including a considerable number of unqualified polygons in the analysis. Therefore, this research uses the 2000 Census blocks since they do not have the problem of unqualified polygons. Socio-economic data are joined to the 2000 census blocks using the ArcGIS spatial join function.

This study uses U.S. Census data at the smallest available geographic units since the purpose of the study is to examine micro patterns of firm dynamic. Census block-group is the smallest geographic unit at which the U.S. census bureau collects most socio-economic data. Census block group data are only available at the Decennial Census. This study, therefore, uses socio-economic data from 1990, 2000, and 2010 Decennial Census to conduct a panel analysis of the patterns of firm dynamics. Census blocks (the units of analysis) within the study area obtain their socio-demographic data from the Census block group that contains their centroid.

Table 10 provides information on the source and geographic level of the control variables used in this study. The next section of this chapter provides a detailed discussion on the structure of outcome and control variables, followed by a detailed discussion on the panel methods used in this study to examine the patterns of firm dynamics.

Table 10. Variable description, source, and geographic level

Determinant	Variable	Geographic level	Source
Firm dynamics (outcome variables)	Firm births	Block	NETS data (1991 – 2009)
	Firm closures	Block	NETS data (1991 – 2009)
	Inward firm relocations	Block	NETS data (1991 – 2009)
	Outward firm relocations	Block	NETS data (1991 – 2009)
Agglomeration economies	Population density	Block Group	U.S. Census (1990, 2000, 2010)
	Employment density	Block Group	U.S. Census (1990, 2000, 2010)
	Number of firms	Block	NETS data (1991 – 2009)
	Average age of firms	Block	NETS data (1991 – 2009)
Socio-demographics	Median household income (in U.S. dollars)	Block Group	U.S. Census (1990, 2000, 2010)
	Unemployment rate	Block Group	U.S. Census (1990, 2000, 2010)
	Percent college educated (Persons 25 years and over)	Block Group	U.S. Census (1990, 2000, 2010)
	Percent African-American	Block Group	U.S. Census (1990, 2000, 2010)
	Median housing rent (in U.S. dollars)	Block Group	U.S. Census (1990, 2000, 2010)
	Property tax (in U.S. dollars)	County/Municipality	The Maryland Department of Assessment (2010)
Spatial context	Transit to auto accessibility ratio[1]	Block Group	EPA's Smart Location Database (SLD) (2010)
	Distance to highway	Block	Calculated using TIGER GIS shapefiles
	Distance to CBD	Block	Calculated using TIGER GIS shapefiles

[1] The transit to auto accessibility ratio is calculated using the SLD accessibility index variables on job accessibility by transit (d5dei) and job accessibly by auto (d5cei). The SLD calculates transit and auto accessibility using origin and destination (OD) matrices for each Census block group within 45 minutes travel time.

3.2. Variable structure

This research examines the influence of heavy and light passenger rail stations located within the State of Maryland on the patterns of firm dynamics. The Maryland Area Regional Commuter (MARC) stations are not included in the analysis because the commuter rail predominantly serves dispersed areas with low residential density and rural development patterns (Liu et al., 2016). The rail stations considered for the analysis belong to three rail transit systems: the Washington Metrorail transit service, the Baltimore Metro Subway, and the Baltimore light rail system. Map 1 highlights the study area and the location of the passenger rail stations. The rail stations were opened in different years over a 26-year span.

The Washington Metrorail system was opened in 1978 but most of the stations were opened after 1984. The most recent stations were opened in 2004. Meanwhile, the Baltimore Metro Subway has a total of fourteen stations operating along a 15.5-mile long route that crosses Baltimore County and the city of Baltimore. The system went through three phases of construction. The first nine stations were opened in 1983 along eight-mile route within the city of Baltimore. In 1987, three more stations were added to the metro system along six-mile route within the suburbs of Baltimore County located northwest of the city of Baltimore. In the last phase, two more stations were opened to the public within the city of Baltimore in 1995. The following subsections define proximity to rail station and present the way in which related variables are constructed. Control variables related to agglomeration economies, the characteristics of local population, and other relevant determinants of firm dynamics considered in this research are also described in this section.

3.2.1. Proximity to rail station

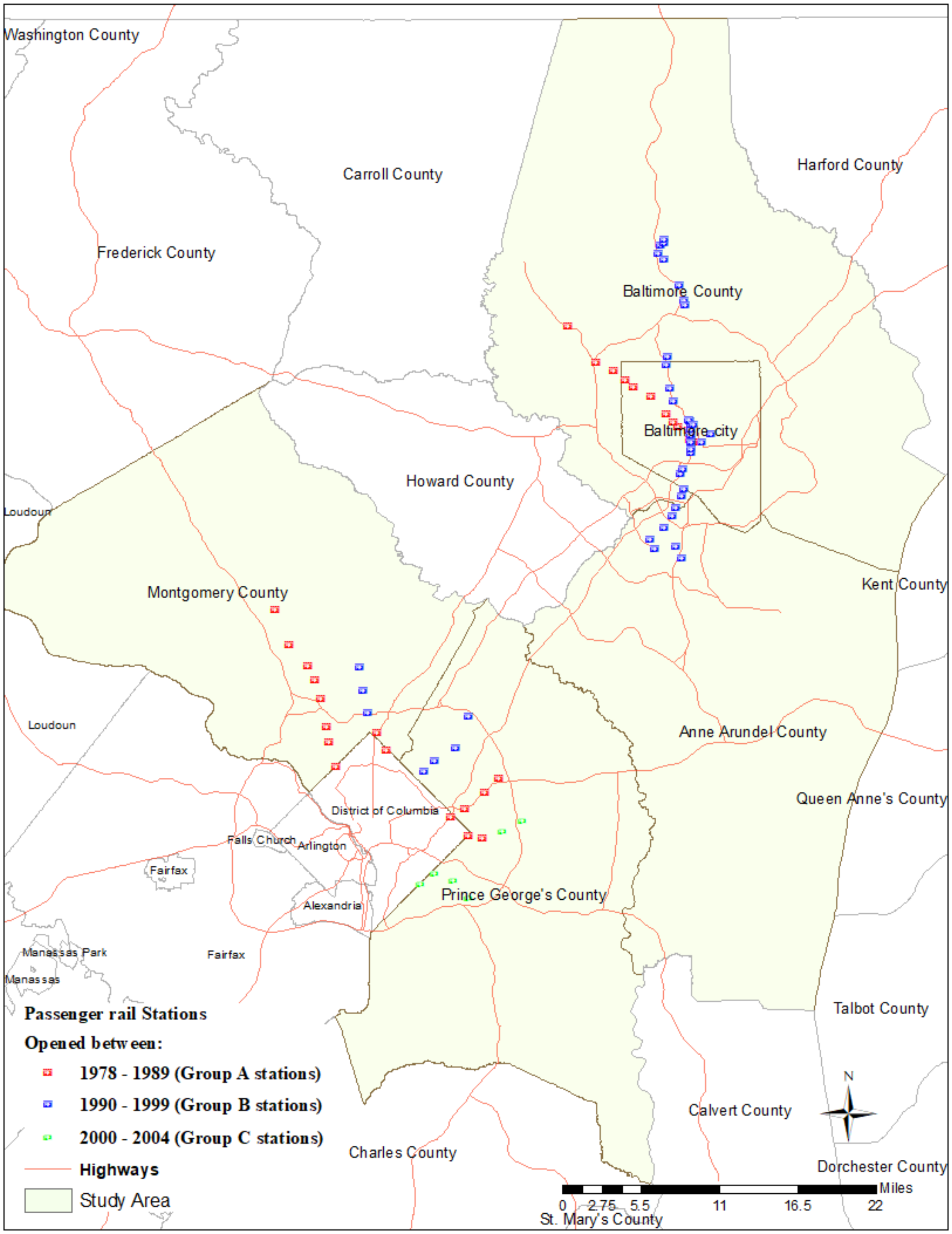
Urban planners typically define areas designated as suitable for transit-oriented development as those within a half-mile radius from rail stations (Hess et al., 2007). The half mile designation is often justified as being the walking distance that people on average are willing to take to reach a station (i.e. about a 10-minute walking distance). Studies that examine property values in relation to rail stations often assign a binary variable to indicate whether or not properties are located within one-quarter or a half-mile of a station (Cervero and Duncan, 2002; Pan, 2013). However, a few more recent studies have presented evidence that rail stations have impacts that extend beyond the conventional half-mile buffer to reach up to one mile away. For instance, Nelson et al. (2015) examined office rental rates in relation to proximity to rail stations in metropolitan Dallas, Texas and Denver, Colorado. Their findings show that a quarter of the rent premiums have extended to locations approximately a mile away from the rail stations. Examining firm birth, Chatman and Noland (2016) also found that areas within a mile of rail stations in Portland, Oregon, and Dallas, Texas are associated with significant positive change in occurrences of firm birth.

Therefore, this research examines the patterns of firm dynamics within three consecutive buffer zones (rings) that extend up to one mile from the passenger rail stations. These three buffers from each rail station are: (1) a *quarter mile* buffer, (2) a *quarter to half mile* buffer, and (3) a *half to one mile* buffer. A Census block is considered to be within one of the three buffers if the buffer contains the block centroid. Map 2 shows an example of the three straight-line buffers used in this study to identify the blocks within proximity to the passenger rail stations.

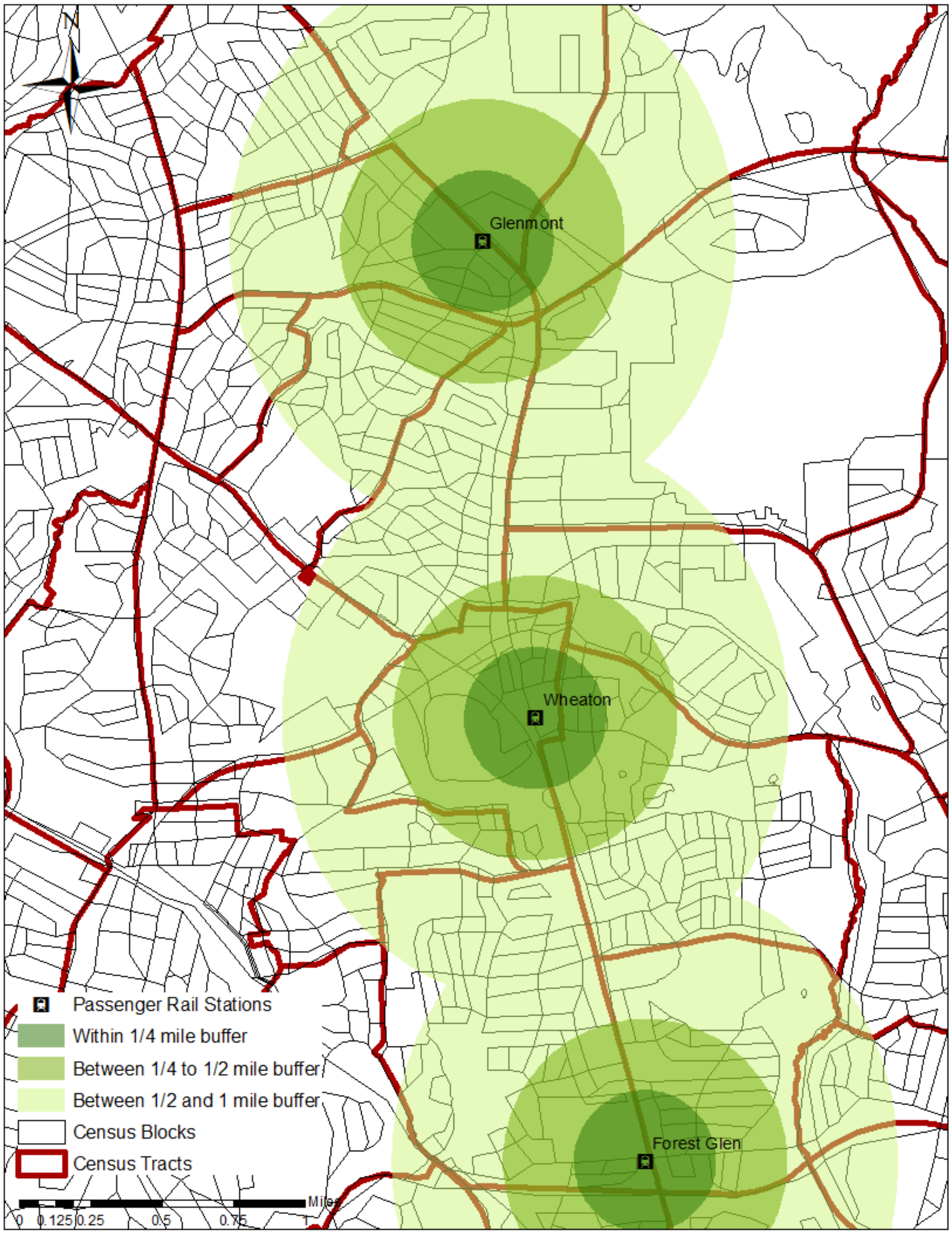
Map 3 and Map 4 identify the Census blocks within proximity to the rail stations belonging to the three examined rail systems within the study area. In addition to the three distance-to-station dummy variables, this research includes a continuous control variable that measures the distance from Census block centroids to a nearest rail station to capture the impact of proximity to stations beyond the one-mile buffer. The continuous distance variable and the dummy buffer variables are calculated for each Census block by measuring the straight-line from each block centroid to a nearest rail station for the three analysis periods (1990, 2000, and 2010) based on the opening date of stations.

To accurately examine the impact of rail stations on the pattern of firm dynamics, a distinction must be made between stations opened more recently and those with a longer time period since their opening. The more mature a rail station, the higher the likelihood that the area around the station has already reached a development saturation point leaving limited or no potential for additional growth. Therefore, the patterns of firm dynamics may demonstrate different trends across rail stations with different level of maturity. The areas around rail transit stations are therefore categorized into three groups in the analyses of this study to account for the variation in the opening year of the stations.

Map 1. The passenger rail stations within the study area categorized by their opening year

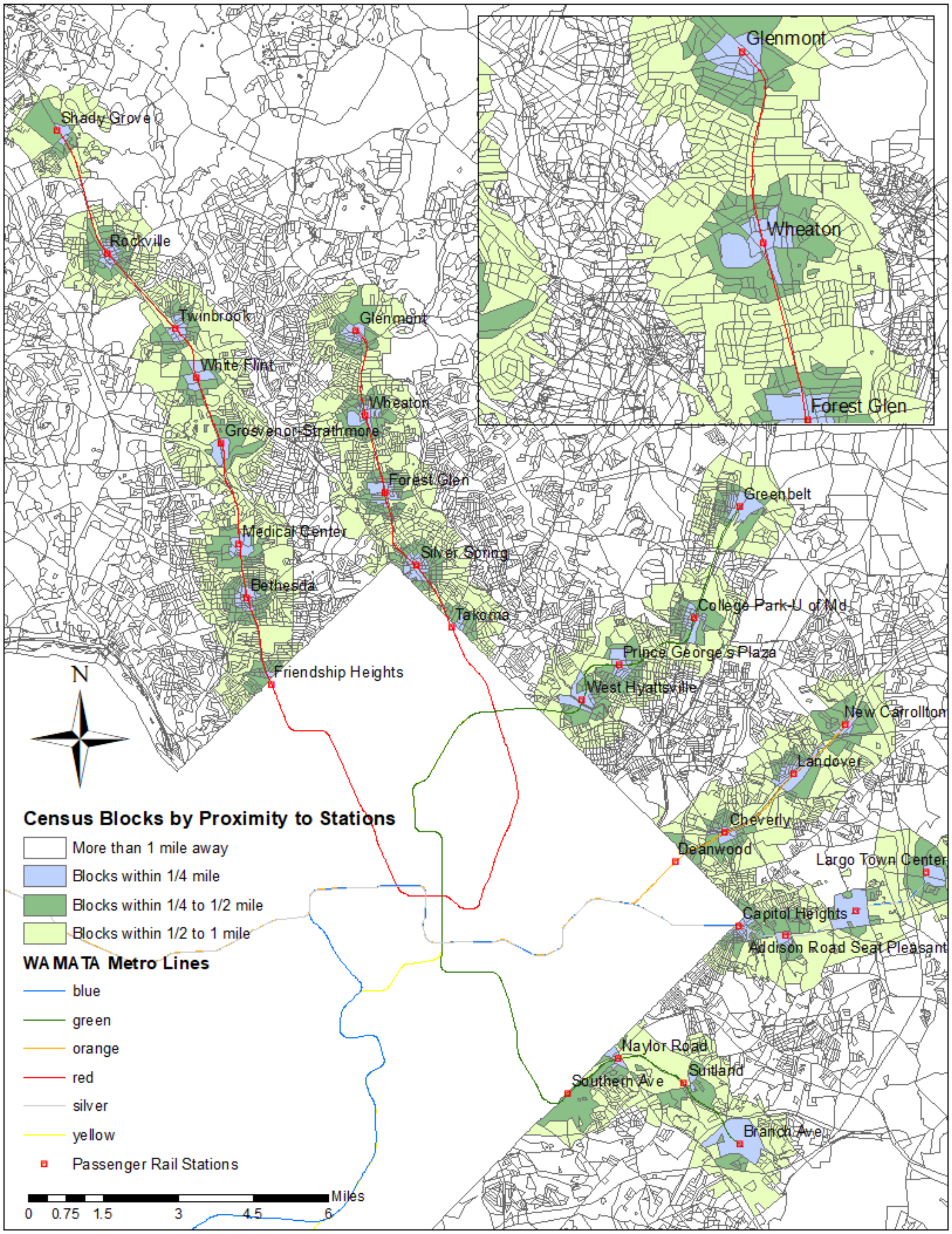


Map 2. An example of the three buffers (rings) used to identify the Census blocks within quarter mile, a quarter to half mile, and a half to one mile straight-line distance of passenger rail stations

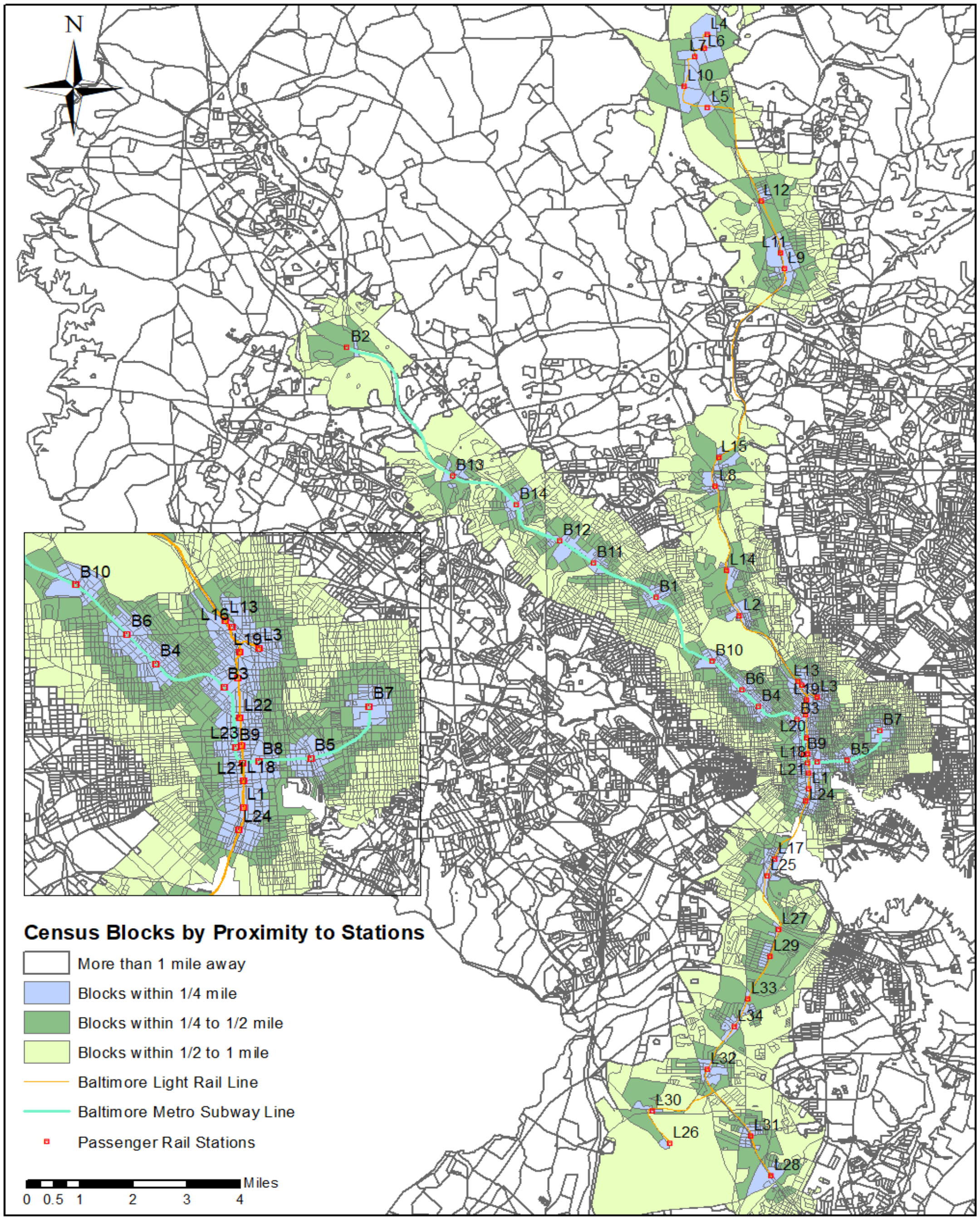


Note: Census tracts boundaries are highlighted in this map (red polygons) as a reference to show that if they are considered in the analysis instead of Census blocks (the unit of analysis of this study) many station buffers will end up without any units of analysis representing them (when considering the centroids for selection).

Map 3. The identified Census blocks within proximity to WMATA metro stations for the examination of the patterns of firm dynamics



Map 4. The identified Census blocks within proximity to Baltimore heavy and light rail stations for the examination of the patterns of firm dynamics



Station ID	Station Name	Station ID	Station Name	Station ID	Station Name
B1	WEST COLDSPRING STATION (METRO) N/B	L3	PENN STATION LIGHT RAIL	L19	MT ROYAL & LT RAIL STAT nb
B2	OWINGS MILLS STATION (METRO)	L4	HUNT VALLEY & LT RAIL STAT nb	L20	CULTURAL CTR & LT RAIL nb
B3	STATE CENTER (METRO) N/B	L5	WARREN RD LIT RAIL nb	L21	PRATT ST LIGHT RAIL STAT nb
B4	UPTON STATION (METRO) N/B	L6	PEPPER RD & LT RAIL STAT nb	L22	CENTRE ST & LT RAIL STAT nb
B5	SHOT TOWER STATION (METRO) N/B	L7	McCORMICK RD LIGHT RAIL nb	L23	LEXINGTON ST & LIGHT RAIL STAT nb
B6	PENN NORTH STAT (METRO) N/B	L8	MT WASHINGTON & LT RAIL STAT nb	L24	HAMBURG & LIGHT RAIL STAT nb
B7	JOHNS HOPKINS STATION (METRO)	L9	LUTHERVILLE & LT RAIL STAT nb	L25	CHERRY HILL LIGHT RAIL STAT nb
B8	CHARLES CENTER (METRO) N/B	L10	GILROY RD LIGHT RAIL nb	L26	BWI AIRPORT LT RAIL nb
B9	LEXINGTON MARKET (METRO) nb	L11	TIMONIUM BUSINESS PARK	L27	PATAPSCO & LIGHT RAIL STAT nb
B10	MONDAWMIN STATION (METRO) nb	L12	TIMONIUM LT RAIL STAT nb	L28	CROMWELL STATION / GLEN BURNIE nb
B11	ROGERS STATION (METRO) N/B	L13	RELIEF PT/DIVISION	L29	BATIMORE HIGHLANDS & LT RAIL nb
B12	REIST PLAZA STATION (METRO) N/B	L14	COLDSPRING & LIGHT RAIL STAT nb	L30	BWI BUSINESS PARK STATION
B13	OLD COURT STATION (METRO) N/B	L15	FALLS RD & LT RAIL STAT nb	L31	FERNDALE LT RAIL nb
B14	MILFORD MILL STATION (METRO) N/B	L16	NORTH AVE LIGHT RAIL nb	L32	LINTHICUM LT RAIL nb
L1	CAMDEN/LT RAIL	L17	WESTPORT LIGHT RAIL STAT nb	L33	NURSERY RD & LIGHT RAIL nb
L2	WOODBERRY/LT RAIL nb	L18	BALTIMORE ST & LIGHT RAIL STAT nb	L34	NORTH LINTHICUM LT RAIL nb

Table 11 provides an inventory of the number of passenger rail stations in the study area by group (A, B, and C), based on the year in which the stations were opened. Table 12, on the other hand, describes the dummy variables used to operationalize proximity to rail station. Group A consists of the most mature rail stations within the study area, defined as stations opened prior to year 1990. Group B consists of stations opened between 1990 and 1999 (43 stations out of the total 77), the largest group within the study area. The last group, group C, consists of stations opened in or after year 2000. Map 2 highlights the three categories of the passenger rail stations based on their opening year. There are several gap years in the analysis of rail stations that were opened before 1990 because the analysis of this study starts in 1990. Therefore, in this study, the distinction must be made across the stations opened prior to the first year of the analysis (before 1990 – group A), around the start year of the analysis (around 1990 – group B), and those opened years after start year of the analysis (after 2000 – group C).

Table 11. Number of rail stations by group (based on the opening year of stations)

Station category	**Year opened**	**No. of stations**	**Total per group**
Group A	1978	6	28
	1980	2	
	1983	9	
	1984	8	
	1987	3	
Group B	1990	2	43
	1992	20	
	1993	10	
	1995	2	
	1997	8	
	1998	1	
Group C	2001	4	6
	2004	2	
Total number of rail stations		**77**	

Table 12. Rail transit variables, source, and geographic level

Determinant	Variable		Geographic level	Source
Proximity to passenger rail	Distance to Rail station (in mile)		Block	TIGER shapefile at Census block (2000) level, and MTA (Mass Transit Administration) for information on the station opening dates.
	Mature stations	Group A stations: area within <=1/4 mile	Block	
		Group A stations: area within 1/4 to 1/2 mile	Block	
		Group A stations: area within 1/2 to 1 mile	Block	
	Less mature stations	Group B stations: area within <=1/4 mile	Block	
		Group B stations: area within 1/4 to 1/2 mile	Block	
		Group B stations: area within 1/2 to 1 mile	Block	
	More recent stations	Group C stations: area within <=1/4 mile	Block	
		Group C stations: area within 1/4 to 1/2 mile	Block	
		Group C stations: area within 1/2 to 1 mile	Block	

3.2.2. Local agglomeration, demographics, and other determinants

To account for the local agglomeration within each Census block, the existing number of firms is an important variable to consider. The presence of a larger number of firms in a certain location may influence the pattern of firm dynamics differently compared to locations with fewer existing firms. NETS firm-level data are summed up within each Census block at each of the three study periods to construct the control variable representing firm agglomeration. Additionally, the presence of a larger number of firms in the same industry sector in a certain location may have a greater influence on the pattern of firm dynamics in that industry sector (i.e. localization economies). The location decision processes of large and small firms may also differ. For instance, the economies of urbanization may play a more important role in the location decisions of smaller firms, whereas larger firms may benefit more from improved access to the labor force. Therefore, this research calculates the total number of firms within each Census block (1) by different firm-size categories and (2) by various industry sectors to construct the control variables that operationalize urbanization and localization (specialization) economies discussed in Chapter 2.

In addition to measures on the number of firms by size and industry sector, this study also controls for the socio-economic effects of population and employment densities on the pattern of firm dynamics (see Chapter 2 for the discussion on these variables). As mentioned previously, the units of analysis within this study are at the Census block level, the smallest unit of analysis within the Census data. However, socio-economic data is only available at the block group level, which is comprised of Census blocks. Socio-economic data for each Census block in the dataset is drawn from and therefore identical to the block group to which it belongs.

The study controls for the characteristics of local population by using data at the Census block group level on median household income, median housing rent, share of the population that is African American, share of the population that is college educated, and share of the population that is unemployed. Analyses also include time-invariant measures as control variables: transit-to-auto accessibility ratio (at peak time from year 2010); distance to nearest highway ramp; and distance to the nearest central business district (either Baltimore City or Washington DC CBD). The transit-to-auto accessibility ratio variable is calculated using transit and auto accessibility measures from the Smart Location Database (SLD), a database developed by the Environmental Protection Agency (EPA) for every Census block group in the United States. The SLD transit and auto accessibility measures are generated by EPA using demographic and travel data from 2010 U.S. Census, as shown in Chapter 2.

3.3. Research methodology

This research uses a series of regression models to examine firm birth, firm closure, and firm relocation patterns with subsets of firm size categories and industry sectors. The first set of models are carried out using a random effect negative binomial regression. Negative binomial regression is a type of generalized linear model in which the outcome variable is a count of the number of occurrences of an event. The second set of models are carried out using a propensity score method to adjust the negative binomial regression. The following two subsections provide more details on the two regression methods used in this research to determine the causal effects of proximity to rail station on the patterns of firm dynamics.

3.3.1. Negative binominal regression model

As discussed in Chapter 2, most of the previous studies on firm birth operationalized the dependent variable as the rate of newly-formed firms using the ecological approach (firms per population) or the labor market approach (firms per employment). The standardization of firm birth allowed these studies to use simple OLS or fixed-effect regression methods. Rate based dependent variables, however, cause considerable illusory correlations that can be ruled out by a count data model. Several empirical studies assume that the firm birth process follows a Poisson or negative binomial distribution since firm births are discrete events, properly analyzed using count models (Holl, 2004a, 2004b, 2004c; Melo et al., 2010; Chatman et al., 2016).

A Poisson distribution assumes its variance is equal to its mean, which is often not realistic. In many cases, the distribution of count variables has a variance that is not equal to its mean (Hilbe, 2011). Overdispersion, where the variance of a distribution is larger than its mean, is a common characteristic of real datasets and of firm birth, closure, and relocation events as well. Under these circumstances, Poisson regression models are not a good fit for count variables. Negative Binomial model (NB) estimates the over-dispersion parameter alpha (α), which makes the model a better fit for count data than Poisson model. Therefore, this research applies a random effects negative-binomial panel model specification (e.g. Hausman et al., 1984) to analyze the relationship of the counts of firm dynamics to rail station proximity at the Census block level. Random effects model is preferred over fixed effect model because some of the explanatory variables are time-invariant, impeding the use of fixed-effects models (Bell and Jones, 2015; Chatman et al., 2016).

The negative binomial regression model is implemented using maximum likelihood estimation. It is a generalized linear model in which the dependent variable (Y) is a count of how many times an event occurs. The dependent variables in the analyses of this study are therefore the count of firm birth, the count of firm closure, and the count of firm relocation for both outward and inward firm relocations. The parametrization of the negative binomial regression (also termed NB2 due to the quadratic nature of its variance function, $\mu + \alpha\mu^2$) takes the following form (Hilbe, 2011):

$$P(y \mid X) = \frac{\Gamma(y+\alpha^{-1})}{y!\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu}\right)^{\alpha^{-1}}\left(\frac{\mu}{\alpha^{-1}+\mu}\right)^{y} \qquad (3.1)$$

where α is the overdispersion parameter. Therefore, if α = 0 in the equation, the model reduces to a simple Poisson regression. While μ (> 0) is the mean of the dependent variable (y). Hilbe (2011) derives this parametrization as a Poisson-gamma mixture, or alternatively as the number of failures before the (1/α) success. The standard negative binomial regression model (NB2) is expressed as follows:

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (3.2)$$

where the explanatory variables x_1, x_2, …, x_n are the covariates considered in the analysis, and $\beta 1$, $\beta 2$, …, βn are coefficients to be estimated.

Unlike most of previous studies, this analysis considers various firm sizes and six 2-digit NAICS industrial sectors. As presented previously in Table 10, the explanatory variables include spatial location measures and demographic data for year 1990, 2000, and 2010 to cover changes in socio-economics, agglomeration, and spatial context of the study area. Because demographic data at the micro level (i.e. census block group) are only available within the U.S. decennial census, the three periods of time are selected to carry out a panel negative binomial regression. The spatial location measures that are time-invariant (i.e. distance to the nearest highway exit and to the nearest CBD) also make the use of random effects desirable because variables that do not differ across time become collinear with dummy variables when using fixed-effect models (Hilbe and William, 2007; Bell and Jones, 2015). The socio-economic variables include measures of agglomeration such as population and firm densities, and measures of market conditions such as household income and housing rent.

It is challenging, however, to accurately examine the impact of rail transit systems on the spatial development patterns due to their endogenous nature. In regions with rail transit systems, rail lines and stations were not randomly placed. They are rather placed in areas with pre-existing location-specific conditions to meet certain objectives, such as: (1) attracting higher ridership, (2) serving existing residential and job locations, and (3) stimulating economic development. These characteristics also change over time. This site selection in actual planning creates greater challenges to accurately measuring economic impacts associated with proximity to rail stations. Therefore, this research uses a Propensity Score (PS) technique to control for the endogeneity of rail transit investments. The following subsection describes the theory behind the propensity score matching (PSM) method and explains how the PS weighting method is used to adjust the standard negative binomial regression model. This study uses both methods (Standard NB model and PS-weighted NB model) to analyze the association between proximity to rail stations and firm birth, closure, and relocation patterns. The use of both methods in this research provides the mean to find out whether controlling for the endogeneity of the placement of rail stations leads to a change in results magnitude and significance.

3.3.2. Propensity score matching and weighted regression

Propensity score matching techniques are frequently used in impact evaluation studies focused on causal effects. There are a limited number of studies, however, that apply the method to measure economic impacts of provision of transportation infrastructure. The general idea behind the method is to measure the impact of a treatment on the treated groups by building counterfactuals for the treated groups using information from non-treatment observations. The method therefore allows observational (or non-randomized) studies to mimic the characteristics of a randomized research design (Austin, 2011). No study prior to this research has used this method in any form to study the impact of rail transit stations on the patterns of firm dynamics.

In the case of rail transit investment, the decision as where to locate rail stations is nonrandom but rather based on pre-determined spatial attributes. The systematic difference between station and non-stations areas sways traditional regression methods to misestimate the impact rail stations have on the patterns of firm dynamics by attributing already existing differences between rail areas and control areas to rail stations. In general, traditional regression methods assume that treated and non-treated groups come from the same distribution. That is, groups that receive the treatment do not differ systematically from non-treated groups. In the case of "introduction of rail stations" as a treatment, consider the following:

Rail station areas (treatment group) = T1

Non-station areas (control group) = T0

Outcome (rail station areas) = M1

Outcome (non-station areas) = M0

In the randomized experiment, the average treatment effect (ATE) can be calculated as:

$$ATE = E\ (M1 \mid T1) - E\ (M0 \mid T0) \tag{3.3}$$

The assumption is that the treatment group would have had the same outcome as the control group if they received no treatment, i.e. E (M0 | T1) = E (M0 | T0).

The treatment is non-randomized, however, in the case of rail investment. The treatment groups (i.e. blocks within a mile of rail stations) and control groups (i.e. blocks more than one mile away from stations) may differ systematically, and the above assumption may not hold true. Therefore, the term E (M0 | T0) does not constitute a valid counterfactual for the treatment areas, i.e. the average effect of non-treated areas does not hold as good proxy to measure unobserved effects of the treated areas. To account for the non-randomized aspect of rail stations, a new mechanism is needed to establish specific control areas that are as similar as possible to the treatment areas prior to the treatment, according to a set of covariates.

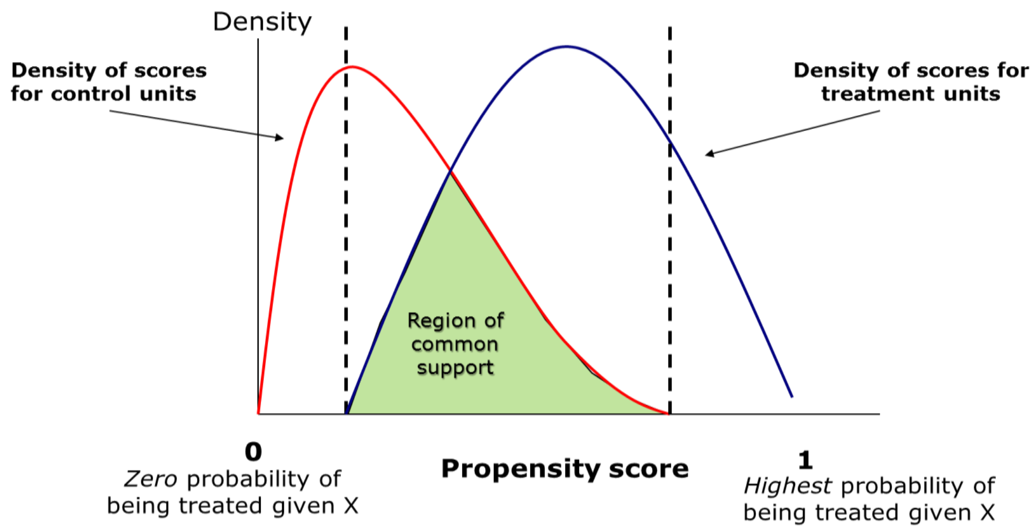
Propensity score matching techniques, first introduced by Rosenbaum and Rubin (1983), suggest that matching individuals on balancing scores (such as propensity score) is more accurate than matching them based on a vector of observable characteristics. Propensity score (PS) is the probability that individuals (or units of analysis) will be assigned to the treatment group given their covariates. The PS is therefore a method that uses stratification[1], full matching[2], or weighting to remove confounding. It is calculated using variables that influence the location decision of rail stations as well as variables relevant to the outcome (i.e. firm dynamics). Ranging from 0 to 1, PS provides the probability of a geographic unit of analysis to have a rail station located within based on the characteristics of Census blocks. Variables influencing the location decision of rail stations include information on population, employment, and firm densities. These variables are also the factors that generally influence subsequent development patterns. In addition to variables related to the treatment (the opening of rail stations in the case of this research), several scholars recommend the inclusion of variables that may not be related to the treatment but are relevant to the outcome in the PS calculation (Jacovidis et al., 2016). Therefore, the variables selected for PS calculation in this research are population and employment densities, household income, unemployment rate, percentage of college graduate population, the percentage of African American population, housing rent, and distance to the nearest highway exit.

Traditionally, the PSM method consists of four steps: (1) propensity score estimation, (2) matching units based on propensity score, (3) matching quality evaluation, and (4) outcome analysis (Pan and Bai, 2015). Several studies, however, have used propensity score estimation as a weight adjustment in regression models (Leuven and Sianesi 2003; Freedman and Berk, 2008; Posner and Ash, 2012). The PS adjusted regression method was also proposed in the initial paper by Rosenbaum and Rubin

[1] Stratification involves stratifying (subclassifying) units of analysis into mutually exclusive subsets based on their estimated PSs, using previously defined thresholds of the estimated PSs.
[2] Full matching is sophisticated form of subclassification that creates a series of matched sets of units of analysis (based on PSs) that are not mutually exclusive. That is, each matched set may have many treated and control units of analysis.

(1983). Rather than match, the PS weights allow the analysis to use all the data by up-weighting some observations and down-weighting others (unless weights are set to 0). For any propensity score dependent analysis, PSM or PS weighted regression, a sufficient overlap should exist between the propensity-score distributions of both the treatment and control groups. The overlap in propensity scores between treatment and control groups is called *the region of common support*. The shaded area in Figure 2 illustrates *the region of common support*.

Figure 2. Propensity score region of common support illustrated



In PS weighted regression (also known as doubly robust estimation), the first step is to estimate propensity scores of the units of analysis by using a binary Probit or Logit regression model (see equation 3.4). The dependent variable (*D*) of the binary model is a dummy variable on whether or not the observation (e.g. Census block) is assigned to the treatment group (e.g. qualify to have a rail station). Theoretically, *D* equals 1 for treated observations, and *D* equals zero for control observations. There are debates about the number of independent variables (covariates) to include in the PS calculation. Nonetheless, researchers generally agree that covariates (X_i) should include all the variables influencing the selection of treatment group (e.g. the location of rail stations), variables associated with the outcome (e.g. the patterns of firm dynamics), or both. The inclusion of both sets of variables increases the precision of the estimated treatment effect without increasing bias (Brookhart et al., 2006; Jacovidis et al., 2016). In the case of rail stations, the propensity score $P(X)$ is the probability that a location will be selected $D = 1$ with characteristics $X = X_i$ for the placement of a station, or

$$P(X) = P\ (D = 1|X = X_i) \tag{3.4}$$

The second step in PS-weighted regression is to use estimated probabilities from the first step (i.e. from the binary regression model) to construct weights. The weights are then used to fit the regression model, which can take a variety of forms: linear, logistic, Poisson, hierarchical Poisson, or proportional hazards regression (Freedman and Berk, 2008). There are different approaches to using the calculated propensity scores to adjust regression models, but only two approaches are common. The first approach is to include the calculated propensity score as a covariate in the standard regression model. This approach was proposed initially by Rosenbaum and Rubin (1983) and later applied by several scholars in the fields of medical research and sociology (Posner and Ash, 2012). Researchers have more recently criticized the use of estimated propensity scores in observational studies as a regression covariate, and recommended using them to weigh the data instead (Bang and Robins, 2005; Hade and Lu, 2014). The second approach in PS-adjusted regression is to calculate the inverse probability of treatment weight (IPTW) and include it as a sampling weight in the regression model. This research uses the second approach to adjust the first set of negative binomial regression models (the standard NB method) because including PSs as a covariate can bias the regression results. The next paragraph explains the problem associated with the first approach.

In the first set of models in this study, the formula for the standard negative binomial regression model as derived from equation (3.2) is:

$$\log(y) = \beta_0 + \beta t\, T + \beta_{\mathrm{x}}\, X \tag{3.5}$$

where (T) represents the treatment dummy variables (i.e. dummy variables for being within ¼ mile, ¼ to ½ mile, and ½ to 1 mile buffer from the rail stations for each of the three station groups). While (X) represents the control variables considered for the analysis. If propensity score is included as a covariate in the regression, the model will take the following form:

$$\log(y) = \beta_0 + \beta t\, T + \beta_{\mathrm{x}}\, X + \beta ps\, P(V) \tag{3.6}$$

where P is the calculated propensity score using a vector of covariates (V) that may or may not include all *X* covariates in the standard negative binomial regression model. βt is the treatment coefficient, and βps is the propensity score coefficient. The problem with model (3.6) is that the effect of βt will be diluted by the existence of P(V) in the model. Specifically, the P(V) value will be high when treatment T=1, so the effect of βt will be much less than in the first model (3.5). Therefore, the treatment coefficient (βt) will yield an underestimated treatment effect (y) of being in the treated group if used as an estimator of the treatment effect (Posner and Ash, 2012).

In this research, the weight that is added to the negative binomial regression model is the inverse probability of treatment weight (IPTW), which is the inverse of the (estimated) PS for treated subjects (D = 1) and the inverse of "1 minus the PS" for untreated subjects (D = 0). The IPTW weighting was first introduced by Rosenbaum (1987) as a form of model-based direct standardization.

Aside from the first set of negative binomial models (standard NB method), this study carries out a second set of models using the IPTWs to adjust the standard binomial negative regression models. The endogeneity of the location of rail station is a factor not to be ignored in the analysis of firm dynamics because pre-existing characteristics of areas within short distance to rail stations can be confounders of the patterns of firm dynamics. The PS-weighted model ensures that the distribution of covariates is similar for the treated and untreated groups, so they are no longer confounders.

In the context of regression adjustment, IPTW is part of causal methods known as *marginal structural models*. Marginal structural models estimate, from observational data, the causal effect of a time-dependent treatment in the presence of time-dependent covariates that may be simultaneously confounders and intermediate variables (Robins et al., 2000). Imbens (2000) proposed the use of IPTW to adjust regression models for estimating causal effects of treatments. Joffe et al. (2004) also provide detailed discussion on PS-weighted regression using the inverse probability of treatment. The adjustment of regression models using IPTW is equivalent to the process of weighting survey samples to ensure they are representative of specific population groups (Morgan and Todd, 2008; Austin, 2011). As mentioned earlier, let T_i be an indicator variable representing whether or not a Census block *i* is treated (i.e. within a one-mile radius from a rail station). Also, let P_i represent the propensity score for the *i*th Census block calculated as a function of a vector of covariates. A simple form of the inverse probability of treatment weights (W_i) can be calculated using the following equation:

$$W_i = \frac{T_i}{P_i} + \frac{(1-T_i)}{1-P_i} \tag{3.7}$$

The equation indicates that a subject matter's weight is equal to the inverse probability of its treatment status. In the context of this study, a Census block's weight is equal to the inverse probability that the block will have a rail station located within a mile radius of the block centroid. Imbens (2000) showed that IPTW regression adjustment produces unbiased estimates of the true treatment effect. The objective of the method is to estimate the average treatment effect between treated and control observations conditional on their observed covariates:

$$\text{ATE} = \text{E}[\, Y_T \mid X \,] - \text{E}[\, Y_C \mid X \,] \tag{3.8}$$

where T is an indicator of treatment (1=treatment, 0=control), X is the vector of independent variables such as population and employment densities, and Y is the outcome (i.e. firm birth, closure, and relocation events). The following equation shows that weighting by the inverse of the propensity score, (p (*x*, T)), produces an unbiased estimate of treatment effect (Imbens, 2000; Posner and Ash, 2012).

$$E\left[\frac{Y\times T}{p(x,T)}\right]=E\left[E\left[\frac{Y\times T}{p(x,T)}\middle|X\right]\right]=E\left[E\left[\frac{Y\times T}{p(x,T)}\middle|X,T=1\right]P(T=1|X)\right]=$$
$$E\left[E\left[\frac{Y_t}{p(x,t)}\middle|X\right]p(x,t)\right]=E[E[Y_t|X]]=E[Y_t] \tag{3.9}$$

where T is an indicator of treatment, X is the vector of independent variables, Y is the outcome, and Y_t is the outcome for each value of the treatment (i.e. blocks within a mile of the stations in group B). Investigators differ on the procedure used for choosing covariates in the PS calculation, as mentioned earlier in this section. Some researchers use all available covariates while others carry out a screening process, so that only variables identified as important or out-of-balance are included in the weight estimation (Freedman and Berk, 2008). This research follows the PS and IPTW calculation documented by Lunt (2014). There are two main steps to generate the IPTWs:

1. the PSs are calculated using a logistic regression model. In this study, Census blocks that have their centroid within a mile from rail stations are considered as treated (*D* = 1) while Census blocks outside of the one-mile buffer are considered untreated (*D* = 0).
2. the calculated propensity scores are then diagnosed for goodness of fit of the covariates using Hosmer-Lemeshow test. In general, a statistically significant Hosmer-Lemeshow test shows that the logistic regression model does not fit the dataset well. This result suggests one of two potential problems in the PS estimation: either non-linearity in the relationships between the covariates and

the log odds of being treated, or an interaction between two of the covariates (Appendix B shows the calculation of PSs and IPTWs).

The covariates selected in this research to calculate the PSs include 1990 data on population and employment densities, income, unemployment rate, rent, education level, and the distance to nearest highway[1] (Appendix B shows the list of covariates used in the PS calculation). Some of these variables may have influenced the decision to locate a rail station while others may have influence on the examined outcome (i.e. firm birth, closure, or relocation). The Hosmer-Lemeshow (HL) test[2] was also carried out to ensure the goodness-of-fit of the calculated PSs. The calculated PSs are then used to estimate the IPTW which theoretically compares what we would expect to see if every unit of analysis received treatment to what we would expect to see if none received treatment. In this research, the PS-weighted regression model limits the analysis to the passenger rail stations in group B because these stations were opened after 1990 and before year 2000. The 1990 covariates used to calculate the PSs are, therefore, at a time prior to the treatment (i.e. the opening of rail stations).

Both regression methods (the standard NB and the PS-weighted NB) are relevant for the analysis of the patterns of firm dynamics. The PS-weighted NB method provide unbiased estimates for the dummy variables of the three rail station buffers (i.e. the *quarter mile*, the *quarter to half mile*, and the *half to one mile* buffer of group B stations)

---

[1] The highway GIS shapefile used to calculate the distance to nearest highway exist is from 2010 but the highway system in the study area has predominantly remained the same between 1990 and 2010 (i.e. time-invariant variable).

[2] HL is a goodness of fit test for logistic regression, which shows how well the data fits the logistic model. Specifically, the HL test calculates if the observed event rates match the expected event rates in population subgroups (A more detailed explanation is provided in Appendix B).

since it controls for the endogeneity of the placement of group B rail stations. The PS-weighted NB method controls for the endogeneity concern by giving more weight to Census blocks similar in characteristics to the blocks near the rail stations prior to their opening (see Appendix B for more details). The blocks located within one-mile of group A and group C stations are dropped from the PS-weighted NB method because there is more than 10-year-gap between their opening date and 1990 (the first year of the analysis), as shown earlier in this chapter (see subsection 3.2.1).

The standard NB method is a better fit for the interpretation of the control variables on agglomeration, socio-economics, and spatial context because: (1) unlike the PS-weighted NB, all the Census blocks within the study area are included in the standard NB models, and (2) the standard NB method does not assign any weights to the Census blocks. The following two chapters present the results of firm birth, firm closure, and firm relocation patterns using both regression methods. The main focus of these chapters is on the patterns of firm dynamics within areas of close proximity to rail stations. Chapter 4 and Chapter 5 also provide discussions on the predicted effects of other control variables (agglomeration, socio-economics, and spatial context variables) on firm birth, closure, and relocation patterns.

3.4. Chapter summary

This research examines the patterns of firm dynamics in relation to rail transit investments using a series of negative binomial (NB) regression models. Three methodological aspects make this study unique. First, in addition to a standard negative binomial regression method, this research applies a second regression method that accounts for the endogeneity of the placement of rail stations (namely propensity score weighted NB regression) to examine the connection between areas within proximity to passenger rail stations and the patterns of firm dynamics. The proximity to rail stations are operationalized by one continuous distance-to-station variable, as well as three dummy variables indicating whether or not the units of analysis are within a *quarter mile*, *quarter to half mile*, or *half to one mile* of a nearest rail station.

Second, unlike previous studies, this research analyzes all the three patterns of firm dynamics: firm birth, firm closure, and firm relocation patterns (both inward and outward firm relocation), which are all relevant to the overall spatial dynamics of the economy. U.S. Census blocks are used as the unit of analysis in the regression models to capture differences in the patterns of firm dynamics at the micro-level. The dependent variables in the regression models are, therefore, the number of firm births, the number of firm closures, the number of inward firm relocations, and the number of outward firm relocations within each Census block.

Finally, this research examines the patterns of firm dynamics across four firm size categories and six industrial sectors. The six industrial sectors are selected following the North American Industry Classification system (NAICS). U.S. statistical agencies use NAICS's classification of business establishments since it was first adopted in 1997 to replace the old Standard Industrial Classification (SIC) system. This research specifically examines the influence investments in three passenger rail systems in Washington-Baltimore metropolitan area have had on the patterns of firm dynamics across the twenty 2-digit NAICS industrial sectors. The examined passenger rail systems consist of the

Washington Metro rail system, administered by the Washington Metropolitan Area Transit Authority (WMATA), as well as the Baltimore Metro Subway and Light Rail systems, administered by the Maryland Transit Administration (MTA).

CHAPTER 4: ANALYSIS RESULTS: FIRM BIRTH AND CLOSURE

The objective of this chapter is to examine whether areas within short walking distance to passenger rail stations influence changes in firm birth and firm closure after accounting for other determinants of firm dynamics. The research hypothesizes that areas near passenger rail stations provide advantages to firms through improved transport accessibility. Areas near passenger rail stations are assumed to experience, on average, an increase in the probability of firm birth and a reduction in the probability of firm closure, compared to control areas that do not have the benefit of proximity to a rail station.

This research further hypothesizes that the magnitude of these effects varies across different industrial sectors and by firm size. To test these hypotheses, this research uses the national establishment time series (NETS) dataset within the case study area in the state of Maryland. The case study area consists of five jurisdictions (i.e. Anne Arundel County, Baltimore County, Montgomery County, Prince George's County, and the City of Baltimore). The units of analysis are the U.S. Census bureau's Census-blocks, which are the smallest U.S. Census units of analysis. The analysis at the Census block level allows this research to capture changes in the patterns of firm dynamics at the micro level. The examination of various firm dynamics for the same study area (i.e. birth, closure, inward relocation, and outward relocation) is unprecedented in the relevant literature.

This chapter is divided into two main sections. The first section focuses on the impact of rail transit on firm birth and the second section focuses on firm closure. To validate robustness of results, each section carries out two regression methods to examine the predictability of influence that the proximity to rail station has on each pattern of firm dynamics. The two regression methods used for modeling are: (1) a standard negative binomial regression model, and (2) an adjusted negative binomial regression model, weighted through a propensity score technique (see section 3.3 in Chapter 3).

The analysis uses data from 1990, 2000 and 2010 since data at the Census block group level are only available at the Decennial Census, as discussed in Chapter 3. The analysis uses 1990, 2000, and 2010 NETS data for the variables of the patterns of firm dynamics.[1] Table 13 provides a summary of the main variables included in the regression models. The summary statistics table provides the mean and standard deviation of data acquired from the U.S. Census bureau at the Census block group level, as well as the mean and standard deviation of NETS data calculated at the Census block level. The summary statistics provided in Table 13 are the average values of the three periods (1990, 2000, and 2010). As explained in Chapter 3, Census blocks within the study area obtain their socio-demographic data from the Census block group that contains their centroid. More detailed descriptive statistics on the patterns of firm birth and firm closure are presented consecutively in the opening of each section.

[1] For NETS related variables, the sum of three years is calculated for each the three periods of the analysis (1990, 2000, and 2010). For example, in 2000, the number of firm births is the sum of firm births in year 1999, 2000, and 2001. The same calculation applies to firm closure and firm relocation variables.

Table 13. Summary statistics for regression variables

Variable	Mean	Std. Dev.	Min	Max
Firm dynamic variables				
Birth counts	1.96	8.31	0	552
Closure counts	1.23	8.15	0	497
Inward Relocation counts	0.22	1.60	0	130
Outward Relocation counts	0.23	2.15	0	165
Proximity to rail station variables				
Distance to Rail station (in mi)	4.85	4.85	0.01	26.39
Group A stations: Blocks within <=1/4 mile	0.01	0.11	0	1
Group A stations: Blocks within 1/4 to 1/2 mi	0.03	0.17	0	1
Group A stations: Blocks within 1/2 to 1 mi	0.07	0.25	0	1
Group B stations: Blocks within <=1/4 mile	0.01	0.11	0	1
Group B stations: Blocks within 1/4 to 1/2 mi	0.02	0.15	0	1
Group B stations: Blocks within 1/2 to 1 mi	0.05	0.22	0	1
Group C stations: Blocks within <=1/4 mile	0.0001	0.01	0	1
Group C stations: Blocks within 1/4 to 1/2 mi	0.0004	0.02	0	1
Group C stations: Blocks within 1/2 to 1 mi	0.0026	0.05	0	1
Agglomeration variables				
Population per sq. mi. (in 1000s)	6.18	7.21	0	165.78
Employee per sq. mi. (in 1000s)	2.89	3.15	0	92.29
Number of firms	4.10	17.53	0	894
Average age of firms	11.33	15.69	0	400
Socio-economic variables				
Median HH Income (in $1000s)	61.57	35.7	0	250
Unemployment rate	5.85	5.95	0	100
Percent college educated	32.78	22.92	0	100
Percent African-American	23.34	34.21	0	100
Median housing rent (in $1000s)	0.84	0.49	0	2.00
Property tax (in $1000)	1.26	0.60	0.12	2.76
Spatial context variables				
Transit to auto accessibility ratio	0.11	0.11	0	1
Distance to highway (in mi)	1.67	1.64	0.003	16.89
Distance to CBD (in mi)	10.28	6.67	0.02	32.4

4.1. Rail transit impact on firm birth

The descriptive and analytical results by firm size and by industry sector demonstrate that areas in close proximity to passenger rail stations exhibit an overall positive sum of the probability of firm birth and firm closure, compared to areas further from the stations. From year 1991 through 2009, the five jurisdictions within the study area experienced the birth of 393,609 firms. Undoubtedly, these firms are not distributed evenly throughout the study area. The density of firm births varied substantially across the Census blocks. As shown in Table 14, among the study area's 39,288 Census blocks, 10,083 blocks (around 26%) had no firm births during the 20 years period of the analysis. To spatially highlights firm birth in the study area, Map 5 shows the spatial variation in the standard deviation of the number of firm births per square mile across Census blocks between 1991 and 2009.

Table 14. Number of units of analysis (Census blocks) within the study area per jurisdiction

Jurisdiction	Number of blocks	Number of blocks with at least 1 birth	Number of blocks without any births
Anne Arundel County	6,446	4,456	1,990
Baltimore City	8,967	6,775	2,192
Baltimore County	7,992	5,692	2,300
Montgomery County	8,212	6,499	1,713
Prince George's County	7,671	5,783	1,888
Total	39,288	29,205	10,083

Source: NETS data, birth densities computed using ArcMap.

Map 5. Firm births per square mile (firm density) within each Census block of the study area (for the period between 1990 and 2009)

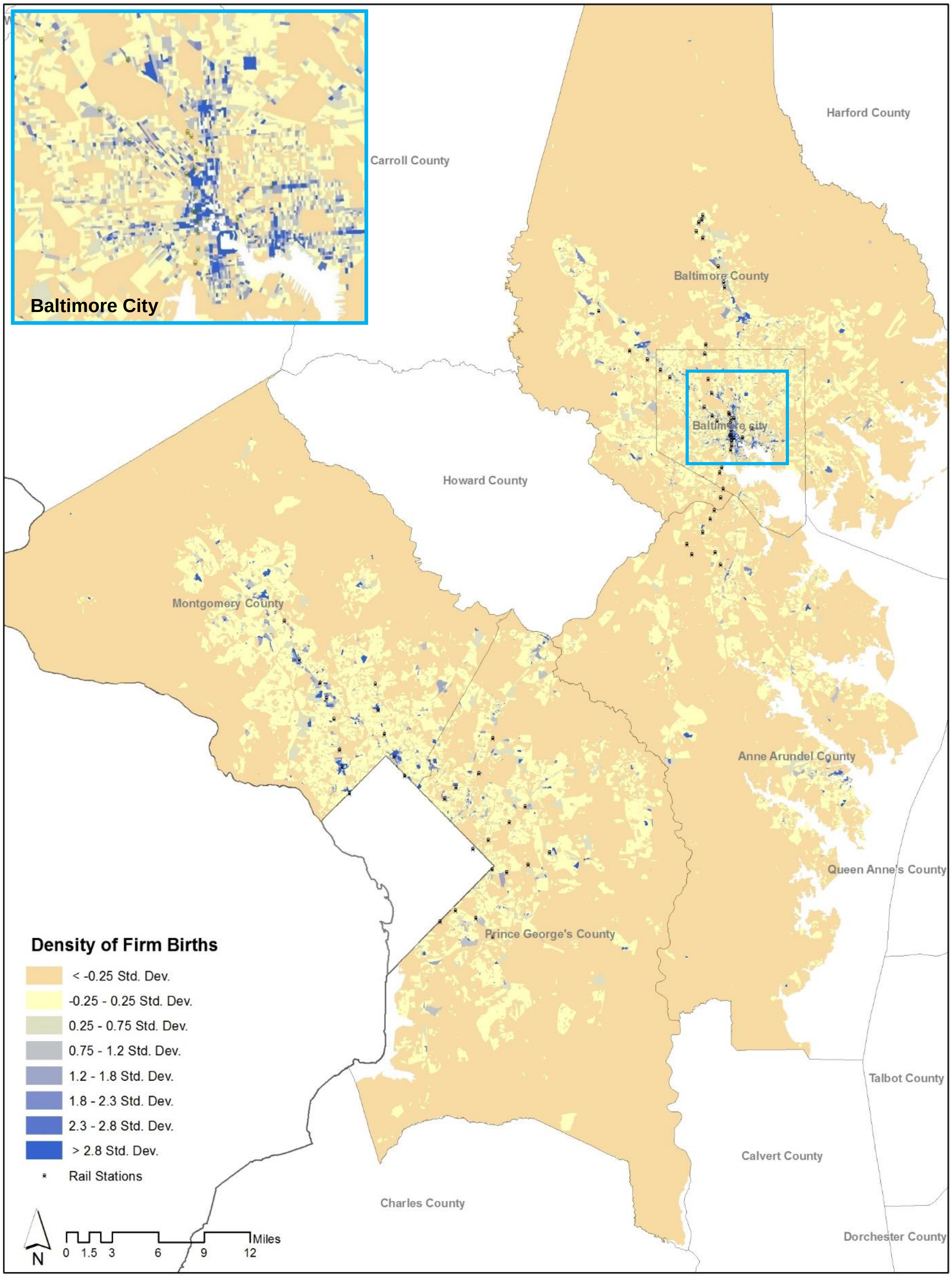
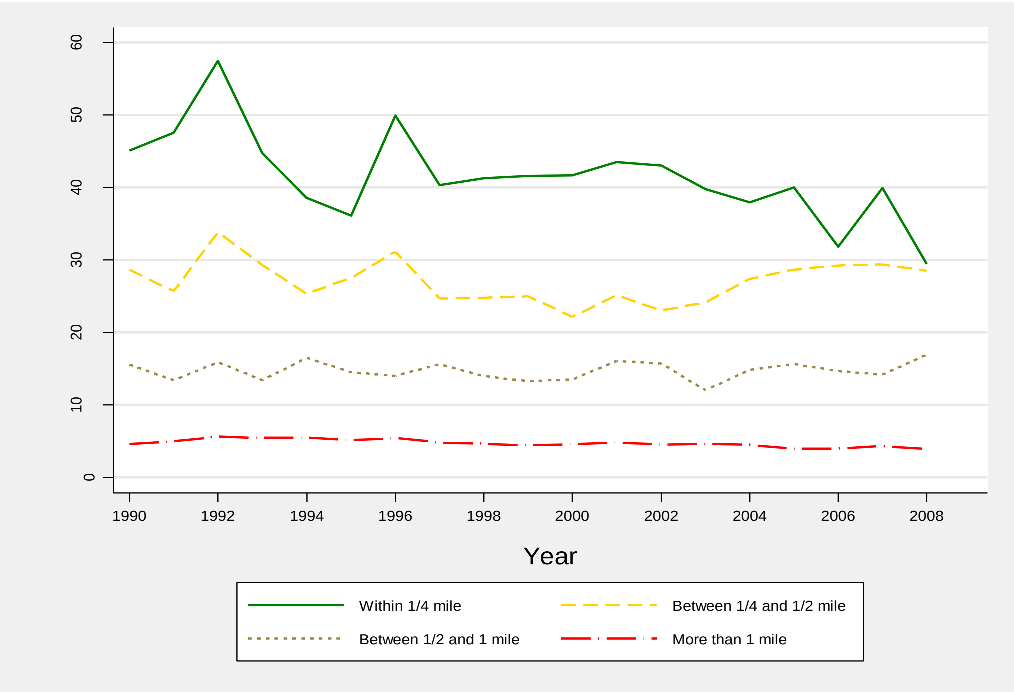


Figure 3 shows the changes in the density[1] of firm birth over time for areas within the three station buffers (i.e. within the *quarter mile* buffer, the *quarter to half mile* buffer, and the *half to one mile* buffer), as well as areas within the study area that are more than one-mile away from the stations (*control areas*). Throughout the period of the study, the number of firm births per square mile in each Census block remained the highest in areas within a *quarter mile* distance from the rail stations followed by areas within a *quarter to half mile* buffer, in comparison to the rest of the study area. However, the high association between proximity to rail station and firm birth can potentially be due to other (confounding) factors, and therefore a controlled statistical analysis is needed to examine whether or not there is indeed a positive and statistically significant association.

Figure 3. Change in the density of firm birth (births per square mile) within the study area over time by distance from rail stations (all firms)



[1] Density of firm birth is only used for descriptive representation of the data. The number of firm births is the outcome measure in the regression models.

The relationship between areas near transit rail stations and firm birth varies by the size of firms. Table 15 summarizes the density of firm births (count of firm births per square kilometer) by size and distance from the passenger rail stations within the study area. Further, Table 15 shows that the number of firm births per square mile (of all size categories) in areas near rail stations is higher compared to control areas located more than a mile from rail stations. Firms with more than five employees had a higher number of firm births within a *quarter mile* distance from the rail stations, relative to firm birth of the same size category in the control areas (248.5 to 13.2 births per square mile). This finding is consistent with the study by Chatman et al. (2016) in Portland, Oregon, and Dallas, Texas.

Table 15. Firm births per square mile by distance from Station and firm size: in the study area (for the period between 1991 and 2009)

Distance to Station	All Firm Births	Sole Proprietor	Five or Less Employees	More than 5 Employees
Within 1/4 mile	1694.9	394.4	1446.5	248.5
1/4 to 1/2 mile	967.0	268.3	838.4	128.6
1/2 to 1 mile	525.3	173.0	472.6	52.7
More than 1 mile	149.6	53.5	136.4	13.2
Average for all blocks	186.7	63.8	168.8	17.9

Source: NETS data. Birth densities computed using ArcGIS.

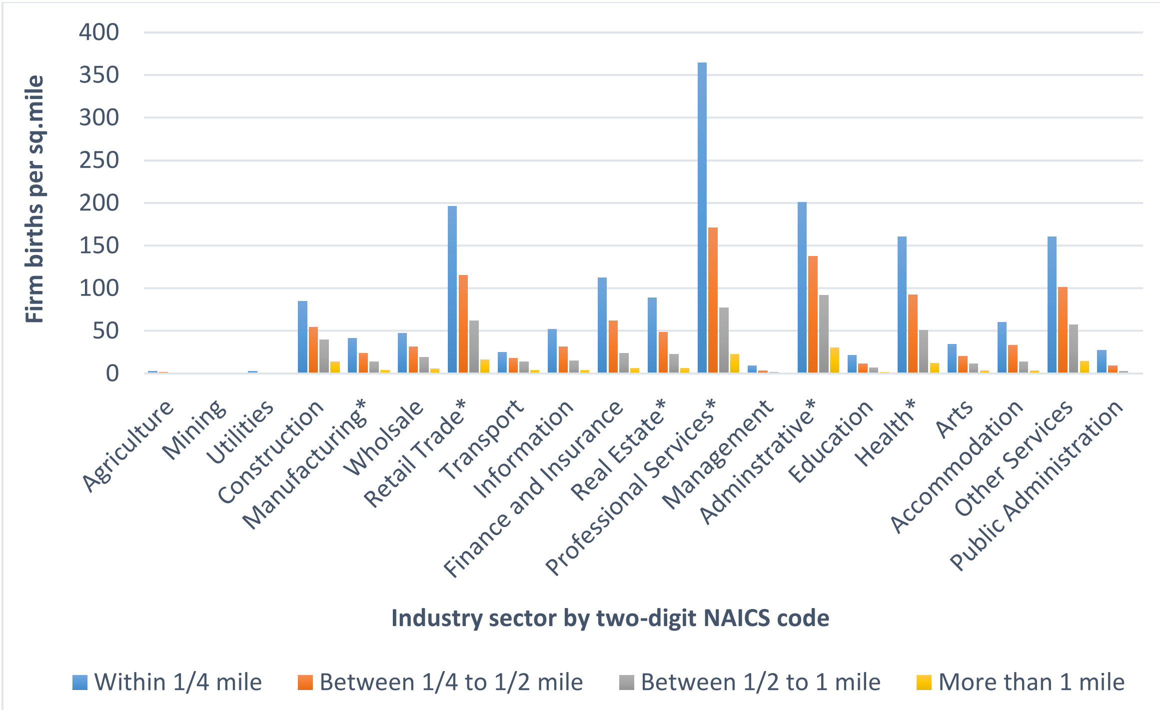
The number of firm births per square mile is disproportional across Census blocks and varies across different industry sectors. Certain areas can be more attractive (or less attractive) to certain industry sectors than others. For example, areas within short distance to passenger rail stations can be more attractive to retail or service firms than they are to manufacturing firms because rail stations generate foot traffic that is normally more beneficial to retail and service firms than it is to manufacturers. Therefore, part of the analysis examines whether the patterns of firm dynamics vary across different industrial sectors in relation to proximity to rail transit stations, and, if it does, in what magnitude these patterns differ across industry sectors.

This research examines five industry sectors that are most dominantly present within the study area (in term of number of firms per square mile), as well as the manufacturing sector for its importance in literature. Reasons for the selection of these six sectors are discussed in further detail in Chapter 3. Based on the two-digit NAICS code, the six industry sectors are:

1. Professional, Scientific, and Technical Services (NAICS code=54);
2. Retail Trade (NAICS code=44 and 45);
3. Finance and Insurance and Real Estate and Rental and Leasing (FIRE) (NAICS code=52 and 53);
4. Administrative and Support and Waste Management and Remediation Services (NAICS code=56);
5. Health Care and Social Assistance (NAICS code=62);
6. Manufacturing (NAICS code=31, 32, and 33).

Figure 4 highlights the number of firm births per square mile (firm density) relative to the distance proximity to the passenger rail stations for the period between 1991 and 2009 and for all the two-digit NACIS industry sectors.

Figure 4. Number of firm births per square mile by two-digit NAICS code and distance from rail stations (1991-2009)



Note: The asterisks (*) indicates the six industry sectors selected for the analysis.

The following two sections present the results of the two regression methods (i.e. the standard NB and the PS-weighted NB methods) to examine the impact of rail transit on firm birth. The first section focuses on regression results by firm size, and the second section analyzes results by industry sector. The PS-weighted NB is carried out by adding the IPTW to the negative binomial model (see section 3.2.2. in Chapter 3). This IPTW gives higher weights to *control* Census blocks that are similar in characteristics to *treated* Census blocks. For the PS calculation, the treated Census blocks are those within a one-mile buffer from the rail stations, taking into account the opening dates of stations. For each regression method, two tables of regression results are provided. One table is for the analysis by firm size (all firms as well as four firm-size categories), and the second table is for the analysis by six industry sectors.

4.1.1. Firm birth by size category: regression results

Firms of different sizes are likely to respond differently to proximity to passenger rail stations as well as other determinants of firm birth because larger firms are inherently different in structure from smaller firms (e.g. larger firms are normally more well established than smaller ones). Compared to smaller firms, firms with higher number of employees might also benefit more from improved accessibility to the labor force that rail systems provide. This section, therefore, analyzes the impact of passenger rail stations on firm birth, considering four firm size categories (i.e. firms with sole proprietor; firms with more than one employee; firms with five or less employees; and firms with more than five employees). As discussed in Chapter 3, rail station maturity or age is essential to the discussion of how firms of varied sizes are impacted by proximity to stations.

The number of firm births per Census block is estimated as a function of distance from Census block to the nearest station in miles, three distance-to-station buffers, and other control variables. As discussed earlier, data are obtained either at the Census block or Census block group. At the Census block level, variables include distance to the nearest highway exit, distance to the nearest central business district (either in Washington, DC or Baltimore City), the total number of firms in all categories, and the number of firm closures[1] measured for the three study periods, 1990, 2000, and 2010. At

[1] As discussed in Chapter 3, firm closure is included as a control variable in the firm birth analysis because the number of firm closures may influence the probability of firm birth within a Census block.

the Census block group level, the control variables are those capturing the socio-economic characteristics of the local population, such as population and employment densities, income, and education (see Table 13 presented earlier in this chapter).

This section starts with an extended discussion of the regression estimates of proximity to rail stations variables (i.e. station buffer variables). The results of other control variables are discussed at the end of the section. The estimated coefficients of distance to station buffers indicate that the proximity to passenger rail station has a positive influence on the probability of firm birth. The Census blocks in the closer proximity to passenger rail stations have experienced higher number of firm births than in the control Census blocks (blocks located more than a mile from the stations). However, as this research hypothesized, the influence of proximity to rail stations on firm birth is heterogeneous across different firm size categories and across the six industry sectors. There are also substantial differences in the magnitude of influence across different station categories based on their level of maturity (i.e. group A, B, and C). As shown in Chapter 3, rail stations in group A are the most mature stations, opened before 1990. Rail stations in group B are those opened between 1990 and 1998. Rail stations in group C are the most recent stations within the study area, opened between 2000 and 2004.

Table 16 shows the regression results of the standard negative binomial method. For rail stations that were opened after year 1990 (group B and C), there are positive and significant effects on firm births associated with whether a Census block is within a *quarter mile* buffer, a *quarter to half mile* buffer, and a *half to one mile* buffer. Remarkably, for all firm births (model 1), the magnitude of the coefficient is much larger for rail stations in group C within the *quarter mile* buffer (0.792) than the other two group of stations, as well as other buffers (ranging between -0.291 and 0.357). That is, more recently opened stations experience higher number of firm births within a *quarter mile* buffer than more mature stations. However, there is a mixed relationship between proximity to rail stations of different maturity groups (A, B, and C) and the magnitude of the probability of firm birth across the four firm size categories.

Table 16. The count of firm birth as a function of proximity to rail stations, agglomeration, and socio-economic characteristics

Dependent variable: firm births	(1) All Firms	(2) Sole Proprietor	(3) Firms > 1 employee	(4) Firms <=5 employees	(5) Firms >5 employees
Distance to Rail station (in mi)	-0.054***	-0.055***	-0.049***	-0.057***	-0.025***
Group A stations: within <=1/4 mile	0.087	0.011	0.241***	0.047	0.811***
Group A stations: within 1/4 to 1/2 mi	-0.007	-0.008	0.075	-0.019	0.420***
Group A stations: within 1/2 to 1 mi	-0.291***	-0.350***	-0.250***	-0.324***	-0.049
Group B stations: within <=1/4 mile	0.229***	0.339***	0.220***	0.323***	0.089
Group B stations: within 1/4 to 1/2 mi	0.217***	0.266***	0.157***	0.266***	-0.035
Group B stations: within 1/2 to 1 mi	0.279***	0.239***	0.218***	0.312***	-0.086
Group C stations: within <=1/4 mile	0.792***	1.069**	0.849***	1.081***	0.284
Group C stations: within 1/4 to 1/2 mi	0.357**	0.836***	0.142	0.393**	0.190
Group C stations: within 1/2 to 1 mi	0.235***	0.544***	0.038	0.264***	0.062
Accessibility ratio	0.847***	0.609***	0.821***	0.775***	1.136***
Population per sq. mi. (in 1000s)	-0.015***	-0.046***	-0.003	-0.014***	-0.043***
Employee per sq. mi. (in 1000s)	0.037***	0.119***	0.001	0.042***	0.014
Number of firms	0.007***	0.009***	0.007***	0.007***	0.011***
Firm closures	0.000	0.001***	0.001**	0.001***	-0.002***
Median HH Income (in $1000s)	0.013***	0.018***	0.007***	0.014***	-0.011***
Unemployment rate	2.423***	4.455***	0.741***	2.656***	-1.126***
Percent college educated	0.004	-0.177***	0.214***	0.008	0.509***
Percent African-American	0.625***	0.744***	0.581***	0.663***	0.194***
Median housing rent (in $1000s)	0.230***	0.420***	0.096***	0.268***	-0.274***
Distance to highway (in mi)	-0.043***	-0.028***	-0.056***	-0.036***	-0.153***
Distance to CBD (in mi)	0.034***	0.042***	0.034***	0.036***	0.036***
Property tax (in $1000)	-0.032*	0.004	-0.090***	-0.020	-0.294***
Constant	-0.685***	-2.284***	-0.051	-0.934***	1.301***
ln_r Constant	0.761***	0.928***	0.913***	0.777***	1.223***
ln_s Constant	-0.253***	-0.110***	-0.504***	-0.230***	-1.558***
N. of cases	116820	116820	116820	116820	116820
Log Likelihood	-160389.805	-102335.347	-126097.983	-155377.573	-35504.633
chi2	27245.735	19500.111	13257.299	28630.265	3584.268

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

To understand the magnitude of the probability of firm birth clearly, the coefficients (βs) from the standard NB regression model in Table 16 can be converted to the percentage of the probability of effect by the equation [e(β) – 1]. For instance, the coefficient β=0.792 for the variable on the *quarter mile* buffer of group C stations means that the Census blocks located within the buffer have experienced 121% more firm births ($e^{0.792} - 1$= 2.21−1= 1.21) compared to the control Census blocks, all else held equal. The control Census blocks are those located within the study area but are more than one-mile away from the rail stations. As stated earlier, the predicted probability of firm birth differs substantially across rail stations areas with different level of station maturity.

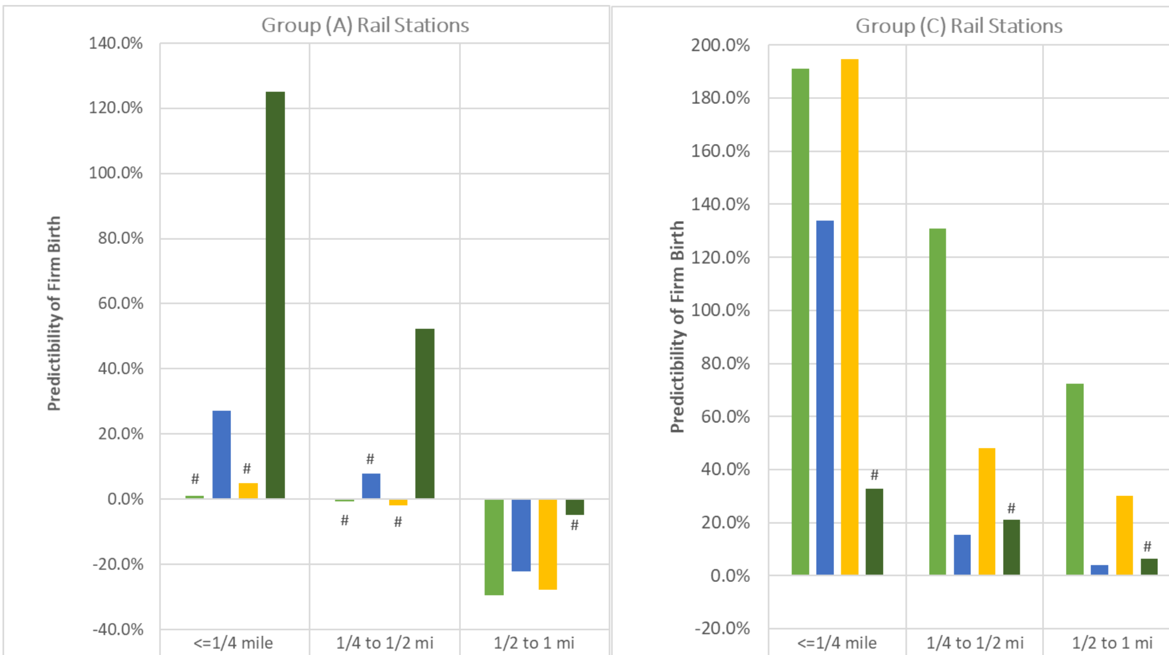
On one hand, the predicted effects of areas near the less mature rail stations (group B and C) on firm birth of all firms (model 1 in Table 16) are positive and statistically significant across all the three station buffers (coefficients range between 0.217 and 0.792). On the other hand, for the mature rail stations that were opened before 1990 (group A stations), the small influence on firm birth of all-firms is statistically insignificant for Census blocks located within the *quarter mile* buffer (β=0.087) and the *quarter to half mile* buffer (β=-0.007), as shown in Table 16. However, Census blocks located within a *half to one mile* of group A stations have a negative and statistically-significant influence on the number of firm births of all firms (β=-0.291). Two potential reasons can explain the negative associations between the blocks within a *half to one mile* buffer of the mature rail stations and firm birth. The first explanation is that the area within a *half to one mile* distance of the mature stations have reached near the saturation point in business establishments, consequently leading to a reduction in the probability of firm birth. The second more plausible explanation is that areas within a *half to one mile* distance of the mature stations have attracted very limited or no commercial urban development and remained dominantly residential over time, even though nearly three decades have passed since their opening. The first explanation is not applicable to several rail stations in group A because areas near these stations remained predominantly residential over time. Examples include Cheverly and Capital Heights stations located in Prince George’s County. The results are not only mixed across the stations with different level of maturity but also across the firm size categories.

The firm birth regression results from the standard negative binomial model (Table 16) show that the magnitude of effect varies considerably across the firm-size categories within areas near the passenger rail stations. For instance, in the *quarter mile* buffer of group B rail stations, the coefficients range in magnitude between 0.089 for firms larger than five employees to 0.339 for firms with sole proprietor. This is a clear indication that the size of firm is an important factor in the association between proximity to rail station and firm birth. The results presented in Table 16 show that smaller firms (i.e. firms with sole proprietor or less than five employees) are the ones benefiting the most from better accessibility to the passenger rail stations, especially for the less mature rail stations in group B and C (i.e. stations opened after 1990) since the coefficients are larger in magnitude for smaller firms. For example, the coefficients for the *quarter mile* buffer of group B station are $\beta=0.339$ for firms with sole proprietor and $\beta=0.323$ for firms with five or fewer employees, which are much larger in magnitude than the coefficient ($\beta=0.089$) for firms larger than five employees. For larger firms, the results are mixed across the station buffers and levels of maturity.

If access to the labor force is the main benefit provided by rail systems, one would expect births of larger firms to be strongly correlated with station proximity. The regression results suggest that this is true only in the case of mature rail stations. Blocks within proximity to the mature rail stations (group A stations) have experienced significantly higher incidents of firm birth of firms with more than five employees compared to areas near less mature stations (group B and C stations). In Table 16, for firms with more than five employees, the coefficients are positive and statistically significant for the *quarter mile* buffer (β =0.811) and the *quarter to half mile* buffer (β=0.420) for the mature rail stations, whereas the coefficients are statistically insignificant for the three buffers of group B and group C stations. In other words, larger firms are more likely to locate in areas within short walking distance of mature rail stations than less mature stations. This result suggest that larger firms benefit more from better labor access via rail. Figure 5 highlights the differences between the influence of the mature rail stations (group A) and the more recently opened stations (group C) on firm birth, across the four firm-size categories.

The analysis by firm size, therefore, suggests that areas near more mature rail stations are more attractive to larger firms (firms with more than five employees) than smaller ones (firms with less than five employees). Figure 5 shows the predicted effects of firm birth within each of the three distance-to-station buffers relative to control Census blocks, all else held equal. The percentages in Figure 5 are calculated from the regression coefficients in Table 16 using the equation ($e(\beta i) - 1$), where βi is the coefficient for a respective distance-to-station buffer. Figure 5 shows that blocks within a *quarter mile* and *quarter to half mile* of the mature rail stations (group A stations) have experienced significantly higher incidents of firm birth compared to areas within a half mile buffer of the more recently opened stations (group C stations). On the other hand, areas near more recently opened stations appear to be more attractive for smaller firms to locate their business startup.

Figure 5. Predicted probability of firm birth by firm size for the three distance-to-station buffers, comparing the difference in outcome between areas near mature stations (group A) to areas near more recent stations (group C)



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	1.1%	-0.8%	-29.5%
> 1 employee	27.3%	7.8%	-22.1%
<=5 employees	4.8%	-1.9%	-27.7%
>5 employees	125.0%	52.2%	-4.8%

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	191.2%	130.7%	72.3%
> 1 employee	133.7%	15.3%	3.9%
<=5 employees	194.8%	48.1%	30.2%
>5 employees	32.8%	20.9%	6.4%

Note: The y-axis shows the predicted effect of firm birth relative to control Census blocks, all else held equal. The percentages (predicted effects) are calculated from the estimated coefficients using (e(β*i*) – 1), where β*i* is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989, and group C stations are those opened between 2000 and 2004. The symbol (#) refers to the statistically insignificant values.

Next, the focus is on the stations in group B that were opened between 1990 and 1998. This study is able to control for their endogeneity. There are no previous studies that explicitly control for endogeneity of the placement of rail stations when analyzing the micro patterns of firm dynamics in relation to proximity to the rail stations. As discussed in Chapter 2, the decision to place rail stations is often not arbitrary, but rather calculated according to a set of pre-conditions such as the pre-existing population and employment densities. Therefore, this study analyzes the patterns of firm dynamics using a second regression method that adjust the standard NB regression using IPTW (this research calls this method the PS-weighted NB regression). In the PS-weighted NB regression method, the analysis is restricted to the stations in group B (opened between 1990 and 1998) for reasons discussed in Chapter 3. Also discussed in Chapter 3, this method gives more weight to Census blocks that are similar, in terms of a number of covariates,[1] to treatment areas prior to the opening of stations (i.e. the 1990 covariates of the Census blocks located within a mile of group B rail stations).

The results from the PS-weighted NB method suggest that proximity to rail stations has a positive impact on the probability of firm birth even after controlling for the endogeneity of the placement of rail stations. Table 17 shows the results of the PS-weighted negative binomial regression models across the firm-size categories. One difference between the two regression methods is that the PS-weighted NB models produce lower magnitudes of influence on firm birth (Table 17) compared to the magnitudes of influence generated by the standard NB models (Table 16). For example, in all-firms model in Table 17 (model 1), the coefficient of the *half to one mile* buffer of group B stations is (0.197) in the PS-weighted NB method, which is lower than the coefficient of the same station buffer in the standard NB model (0.279); both coefficients

[1] As shown in appendix B, the covariates considered in the PS calculation include population and employment densities, household income, level of education, unemployment rent, distance to highway, and housing rent.

are positive and statistically significant. Regression coefficients are translated into predicted probabilities to clearly explain the differences in the magnitude of influence across the two methods.

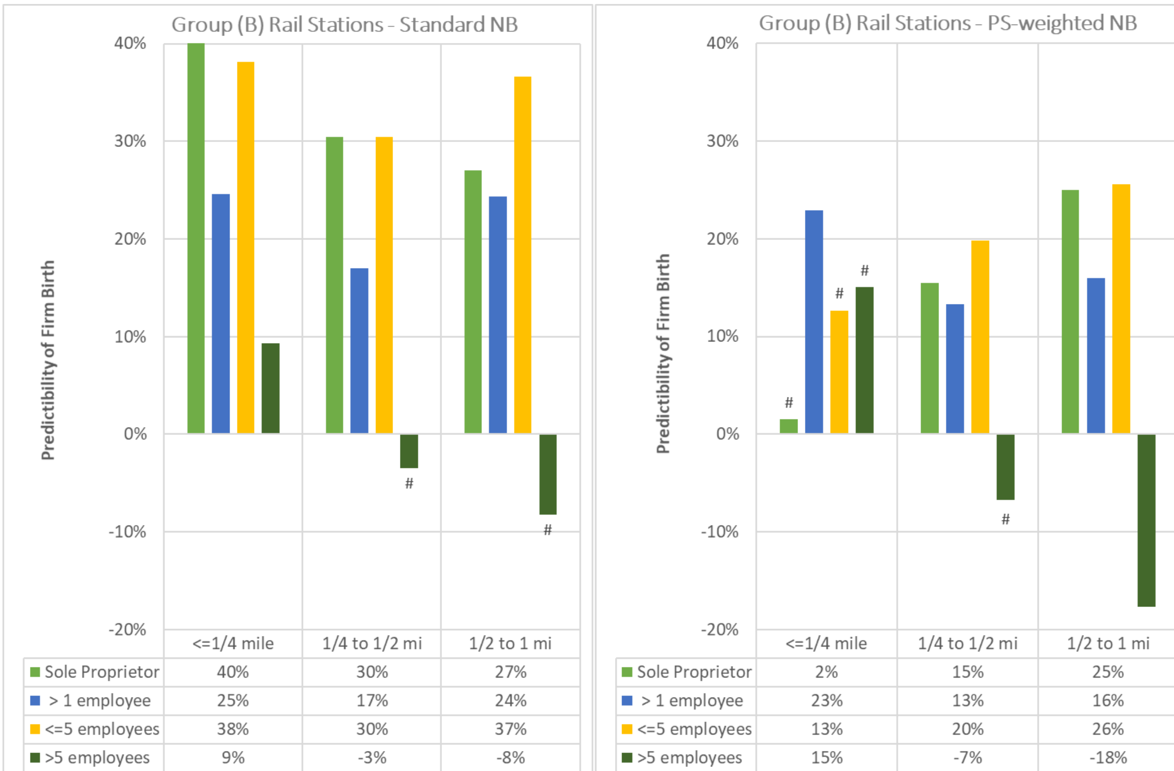
Figure 6 illustrates the differences in the predicted probabilities of firm birth in areas near group B stations across the two regression methods. For example, the PS-weighted NB regression predicts that blocks within a *quarter to half mile* of group B stations are, on average, 15% more likely to experience a birth of a firm with sole proprietor compared to control areas, all else held equal. On the other hand, the standard NB model predicts that the same blocks (within a *quarter to half mile* buffer of group B stations) are, on average, 30% more likely to experience a birth of firm with sole proprietor, all else held equal. Note that the comparison between the two methods is only possible for group B rail stations because the PS-weighted NB models are restricted to these stations, for reasons discussed earlier in Chapter 3.

Table 17. The count of firm birth by firm-size as a function of proximity to rail stations, agglomeration, and socio-economic characteristics, using PS-weighted negative binomial regression

Dependent variable: firm births	(1) All Firms	(2) Sole Proprietor	(3) Firms > 1 employee	(4) Firms <=5 employees	(5) Firms >5 employees
Distance to Rail station (in mi)	-0.034***	-0.042***	-0.029***	-0.036***	-0.014*
Within <=1/4 mile (Group B)	0.100	0.015	0.206*	0.119	0.140
Within 1/4 to 1/2 mi (Group B)	0.155***	0.144**	0.125**	0.181***	-0.070
Within 1/2 to 1 mi (Group B)	0.197***	0.223***	0.148***	0.228***	-0.194*
Accessibility ratio	0.230*	-0.111	0.242*	0.135	0.587**
Population per sq. mi. (in 1000s)	-0.022***	-0.044***	-0.011***	-0.021***	-0.045***
Employee per sq. mi. (in 1000s)	0.057***	0.121***	0.019**	0.061***	0.011
Number of firms	0.073***	0.051***	0.066***	0.074***	0.057***
Firm closures	-0.023***	0.033***	-0.044***	-0.014***	-0.053***
Median HH Income (in $1000s)	0.010***	0.014***	0.005***	0.011***	-0.013***
Unemployment rate	2.150***	4.534***	0.247	2.369***	-1.917***
Percent college educated	-0.059	-0.258***	0.124*	-0.083	0.377***
Percent African-American	0.637***	0.670***	0.625***	0.654***	0.338***
Median housing rent (in $1000s)	0.162***	0.241***	0.077***	0.184***	-0.158***
Distance to highway (in mi)	-0.037***	-0.015**	-0.052***	-0.025***	-0.169***
Distance to CBD (in mi)	0.025***	0.034***	0.024***	0.026***	0.029***
Property tax (in $1000)	-0.029	-0.015	-0.072***	-0.021	-0.228***
Constant	-1.363***	-2.816***	-1.214***	-1.580***	-1.374***
chi2	40743.336	70059.049	11621.58	62245.608	8516.205
N. of cases	101859	101859	101859	101859	101859

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

Figure 6. Predicted probability of firm birth by distance from rail stations in group B, comparing results from two regression methods: the standard NB (left side) and the PS-weighted NB (right side)



Note: The y-axis shows the predicted effect of firm birth relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients using ($e(\beta i) - 1$), where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B stations are those opened between 1990 and 1999. The symbol (#) refers to the statistically insignificant values.

Figure 6 also shows that areas within a *quarter mile* of group B stations have higher probability of firm birth for firms with more than five employees compared to control areas (the coefficient 0.089 translates to 9% probability in the standard NB model in Table 16, whereas the coefficient 0.140 translate to 13% probability in the PS-weighted model in Table 17). The predicted probability in the PS-weighted regression is statistically insignificant, however. As clearly shown in Figure 9, the predicted probabilities of firm birth across the two methods differ in magnitude of effect (coefficient) but not direction (sign of the coefficient). The endogeneity of the location of rail station, therefore, is a factor which should not be ignored in the analysis of firm dynamics. The tendency of the standard NB method to overstate the predicted effects compared to the PS-weighted indicates that some of the predicted effect in the standard NB method is due to pre-existing characteristics of the areas within short distance to the rail stations.

Focusing on the other transport-related variables in Table 16, the standard NB method[1], the coefficients of transit-to-auto accessibility ratio suggest that greater transit access matters more for larger firms than smaller ones (i.e. β=1.136 for firms with more than five employees compared to β=0.609 for firms with sole proprietor), which is similar to the finding by Chatman and Noland (2016). The regression results suggest that distance to highway exit also has a negative and statistically significant association with firm birth across all firm size categories. Considering the magnitude of effect, the access to highway similarly appears to be a more important factor for larger firms (β=-0.153 for firms with more than five employees) than smaller ones (β=-0.028 for firms with sole proprietor and β=-0.036 for firms with five or less employees). For every mile away from a nearest highway exit, Census blocks are 3% less likely to experience a firm birth of firms with sole proprietor ($e^{-0.028} - 1$= -0.03), all else held equal. Evidently, for small firms, shorter distance to the nearest rail station matters more than shorter distance to the nearest highway. For every mile away from a nearest rail station, Census blocks are 5% less likely to experience a firm birth of firms with sole proprietor ($e^{-0.054} - 1$= -0.05), all else held equal. For the location-decisions of larger firms, however, proximity to highway

[1] For reasons discussed in Chapter 3, the standard NB method is used to interpret and discuss the results of the other explanatory variables (i.e. variables other than the three distance-to-station buffers).

appears to be a more important factor (14% higher likelihood of birth for every mile closer to a nearest highway exit) compared to proximity to distance to a rail station (2% higher likelihood of birth for every mile closer to a nearest station).

For the agglomeration related variables in Table 16, the total number of existing firms is a significant determinant of firm birth across all firm size categories. The regression results show a positive and statistically significant association between the number of existing firms and firm birth in all size categories, although the positive effects are small in magnitude (ranging between β=0.007 and β=0.011 which translate to 0.9% and 1.1% probability of effect, consecutively). One unexpected finding is the result on the association between population density and the rate of firm birth. The association is negative and statistically significant, however small in magnitude, across all firm size categories (ranging between β=-0.003 and β=-0.046). This finding suggests that population density is not an important factor influencing firm birth within the study area.[1] The insignificant effect of population density on firm birth signals the vast suburbanization of the study area, where extensive residential areas exist in isolation of commercial and employment zones. Many Census blocks with relatively high population density experienced zero firm birth within the study area. Employment density, on the other hand, is an important factor influencing firm birth for smaller firms, that is, firms with sole proprietor (β=0.119) followed by firms with five or less employees (β=0.042).

Turning to the socio-economic variables in Table 16, the association is positive and statistically significant between median household income and firm birth for smaller

[1] The negative association between population density and firm birth remained even when the models use data on population density at the Census tract level instead of Census block-group level.

firms (i.e. firms with sole proprietor, β= 0.018, as well as firms with five or fewer employees, β= 0.014). The same association is true for the median housing rent. The positive associations between these income-related variables and firm birth of smaller firms suggest that smaller firms are more likely to locate in areas with high income levels. On the other hand, the association between income-related variables (both median household income and median housing rent) and firm birth is negative and statistically significant for larger firms with more than five employees (β=-0.011 and β=-0.274, consecutively), which suggests that larger firms are more likely to be attracted to areas with lower property and labor costs.

The regression results in Table 16 show that education level (represented by percentage of population with a college or higher degree) is a positive and statistically significant determinant of firm birth for firms with more than five employees (β=0.509), suggesting human capital is also an important factor for larger firms. Regarding race, the association is positive and statistically significant between the percentage of population that is African American and firm birth (the coefficients range between β=0.194 and β=0.744). The unemployment rate is a positive factor influencing firm birth for smaller firms (firms with less than five employees, β=2.656), but negative for larger firms (firms with more than five employees, β=-1.126). The property tax coefficients are negative and statistically significant for firm birth in the large-size category (firms with more than one employee, β=-0.090, and firms with more than five employees, β=-0.294) suggesting that larger firms are more likely to locate in areas with low property taxes. The association between property tax and firm birth is positive for firms with sole proprietor (β= 0.004) but the influence is statistically insignificant.

Another unexpected finding is related to the variable representing the spatial context, that is the distance to nearest CBD (i.e. the CBD of either Washington D.C. or Baltimore City). The distance to CBD coefficients in Table 16 (ranging between β=0.034 and β=0.042) suggest a positive association between the distance to CBD and the number of firm births[1]. That is, the farther a Census block is from the CBD, the higher it is likelihood to experience a firm birth, all else held equal. The highly suburban and polycentric urban form of the study area may well explain the insignificance of proximity to CBD as a factor in the location decision of firms. The section on firm closure provides a more detailed account of the net influence of control variables on firm birth compared to their influence on firm closure.

4.1.2. Firm birth by industry sector: regression results

Firms across different industry sectors may value proximity to rail stations differently. Labor-dependent firms (e.g. manufacturing and retail firms) may benefit more from areas with improved accessibility to the labor force such as that provided by rail systems; whereas knowledge-dependent firms (e.g. professional services and FIRE firms) may value dense urban areas, due to knowledge spillover benefits, more than proximity to rail stations. Therefore, this research examines the probability of firm birth in relation to rail proximity across the six industry sectors described earlier in this chapter. The dependent variable for this set of models is the number of firm births in each of the six industry sectors at the Census block level.

[1] The association between firm birth and distance to CBD remains negative even when the regression models include a dummy control variable that determines whether or not a Census block is within two miles from the nearest CBD.

Table 18 summarizes the results of the standard negative binomial regression model for each of the six industry categories. The analysis by industry sector shows, overall, that areas within a short distance of rail stations positively influence firm birth across all the six specific sectors examined in this research. This is suggested by the negative coefficients, in Table 18, of the continuous distance-to-nearest-station variable across all industry sectors (the coefficients range between -0.037 and -0.060). The negative coefficients mean that the farther a Census block is from the rail stations, the lower is the number of firm births. The regression results of firm birth by sector are mixed, however, for the three distance-to-station buffers and the three station groups (group A, B, and C based on the level of maturity of the stations).

Table 18. Standard negative binomial estimated coefficients: firm birth by selected industry sectors

Dependent variable: count of firm births	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail Station	-0.037***	-0.060***	-0.044***	-0.051***	-0.056***	-0.038***
Group A stations: within <=1/4 mile	0.463***	0.328***	0.578***	-0.191*	0.171	0.083
Group A stations: within 1/4 to 1/2 mi	0.159*	0.236***	0.229**	-0.043	-0.05	0.240*
Group A stations: within 1/2 to 1 mi	-0.236***	-0.284***	-0.434***	-0.335***	-0.179**	-0.091
Group B stations: within <=1/4 mile	0.323***	0.299***	0.477***	0.061	0.495***	0.349*
Group B stations: within 1/4 to 1/2 mi	0.092	0.181**	0.275***	0.178***	0.007	0.098
Group B stations: within 1/2 to 1 mi	0.181***	0.132**	0.131*	0.178***	0.034	0.116
Group C stations: within <=1/4 mile	0.012	1.390***	0.967*	1.464***	0.531	0.759
Group C stations: within 1/4 to 1/2 mi	-0.020	0.445	0.353	-0.110	0.263	-0.795
Group C stations: within 1/2 to 1 mi	-0.303	0.459***	0.246	0.369***	-0.09	0.334
Ratio of Transit to Auto Accessibility	1.084***	0.414**	0.531***	0.421***	0.627***	0.478*
Population density in 1000	-0.006	-0.035***	-0.038***	-0.043***	-0.011*	-0.052***
Employment density in 1000s per sq. mi.	0.006	0.075***	0.069***	0.107***	0.022*	0.055**
Number of firms	0.009***	0.011***	0.010***	0.002***	0.010***	0.014***
Number of firms in the same sector	0.000	-0.009***	0.001	0.196***	0.004***	0.039***
Number of firm closures	0.004***	0.002***	0.003***	-0.009***	0.002***	0.005***
Median HH Income in $1000s	0.000	0.007***	0.008***	0.015***	0.004***	-0.001
Unemployment rate	-0.004	0.016***	0.020***	0.036***	0.011***	-0.003
Percent college educated	0.001	0.012***	0.008***	-0.003***	0.012***	0.002
Percent African American	0.004***	0.007***	0.007***	0.008***	0.008***	0.002**
Median housing rent in $1000	0.035	0.153***	0.163***	0.323***	0.067**	0.036
Distance to nearest highway exit in miles	-0.084***	-0.059***	-0.114***	-0.025***	-0.090***	-0.072***
Distance to CBD in miles	0.035***	0.060***	0.041***	0.041***	0.035***	0.041***
Property tax in 2010	-0.127***	-0.183***	-0.109***	0.025	-0.142***	-0.019
Constant	0.494***	-1.189***	-0.922***	-2.604***	-1.175***	-0.336
ln_r_cons	1.566***	1.403***	1.552***	1.460***	1.144***	2.570***
ln_s_cons	-1.230***	-0.725***	-1.111***	0.646***	-1.186***	-0.812***
N. of cases	116820	116820	116820	116820	116820	116820
Log Likelihood	-46863.579	-57005.155	-41102.511	-75826.398	-37037.397	-16573.333
chi2	3503.402	9999.008	6223.196	24212.128	4650.614	2078.322

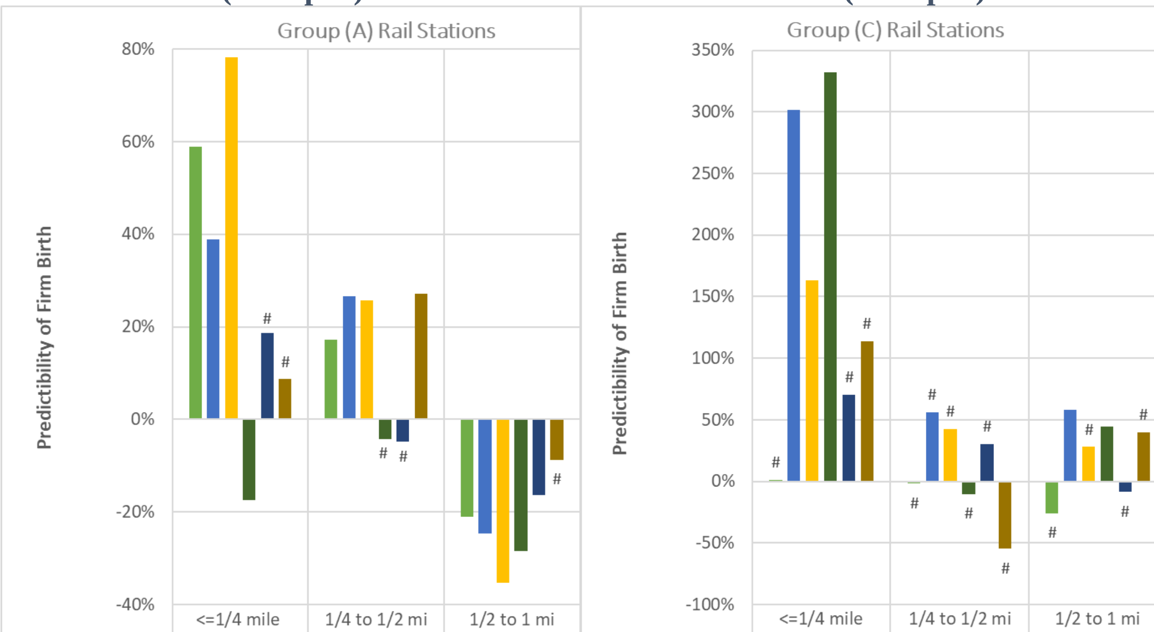
** p<0.05, ** p<0.01, *** p<0.000*

Focusing on the mature rail stations (group A), the probability of firm birth is positive and statistically significant within the *quarter mile* buffer of the stations for three out of the six industry sectors (the three sectors are: Retail trade, β=0.328; Professional services, β=0.463; and FIRE, β=0.578) compared to control areas, all else held equal (see Table 18). Unlike other sectors, however, the likelihood of firm birth in the administrative sector is negative (β=-0.191) for blocks located within the *quarter mile* buffer of group A stations. In the case of more recently opened stations (group C), the probability of firm birth is positive and statistically significant within the *quarter mile* buffer of the stations for all industry sectors; the results, however, are statistically significant only for firms belonging to the professional sector (β= 1.390), FIRE sector (β=0.967), and the administrative sector (β= 1.464). As shown in Table 18, most of the estimated coefficients for group C station buffers are statistically insignificant (likely because there are only six stations in group C).

To clearly illustrate these effects, Figure 7 shows the predicted probability of firm birth within the three station buffers of group A and group C stations, and across the six industry sectors. The predicted probabilities are calculated using coefficients of the three station buffers in Table 18. For example, holding all else equal, blocks located within a *quarter mile* radius from a rail station in group A are 78% more likely to experience a firm birth belonging to FIRE sector compared to control areas located more than a mile from the stations. Positive relationships are found between blocks located within a *quarter to half mile* of group A stations and firm birth belonging to retail sector (17%), professional service sector (27%), FIRE sector (26%), and manufacturing sector (27%).

Similar to the results of the analysis by firm size, blocks within *half to one mile* of the mature stations (group A) have experienced negative probabilities of firm birth compared to control areas (see Figure 7). These negative predicted probabilities can be attributed to the two explanations discussed earlier in the firm birth analysis by firm size.

Figure 7. Predicted probability of firm birth by industry sector for the three distance-to-station buffers, comparing the difference in outcome between areas near mature stations (Group A) to areas near more recent stations (Group C)



Group (A)	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	59%	17%	-21%
Proffesional	39%	27%	-25%
FIRE	78%	26%	-35%
Admininstrative	-17%	-4%	-28%
Health	19%	-5%	-16%
Manufacturing	9%	27%	-9%

Group (C)	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	1%	-2%	-26%
Proffesional	301%	56%	58%
FIRE	163%	42%	28%
Admininstrative	332%	-10%	45%
Health	70%	30%	-9%
Manufacturing	114%	-55%	40%

Note: The y-axis shows the predicted effect of firm birth relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients using (e(β*i*) – 1), where β*i* is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989, and group C stations are those opened after 2000. The symbol (#) refers to the statistically insignificant values.

For group B stations (opened between 1990 and 1998), the influence of areas near stations on firm birth is positive for all the six industry sectors and within all the three station buffers (*quarter mile*, *quarter to half-mile*, and *half to one mile* buffers), as estimated by the standard NB method in Table 18. It is important to note that the regression results of the PS-weighted NB method also indicate a positive association between areas located within proximity to group B stations and the number of firm births across the industry-specific models (see Table 19). The negative and statistically significant coefficients of the continuous distance-to-station variable in Table 19 suggest that there is a negative association between the distance to rail stations and the number of firm births. In other words, the shorter is the distance from a Census block to the nearest group B rail station, the higher is the number of firm births, all else held equal. However, the results are mixed on the probability of firm birth by industry sector for the three buffers of group B stations.

Several coefficients in the PS-weighted NB method are statistically insignificant for the three station buffer variables. In Table 19, positive and statistically significant coefficients of the station buffers are found in four sectors: Retail, professional services, FIRE, and administrative sectors (model 1 through model 4). For instance, Census blocks located within the *quarter mile* and the *half to one mile* buffers of group B stations have positive and statistically-significant influence on probability of firm birth belonging to retail sector (β=0.249 and β=0.203, consecutively). Census blocks located within the *quarter mile* buffer of group B stations also experienced a positive and statistically significant association with firm birth belonging to FIRE sector (β=0.539). These results suggest that even after controlling for the endogeneity of station locations, areas within short walking distance of group B rail stations experienced higher number of firm births of firm belonging to retail, professional services, FIRE, and administrative sectors compared to areas located more than a mile from the stations, all else held equal.

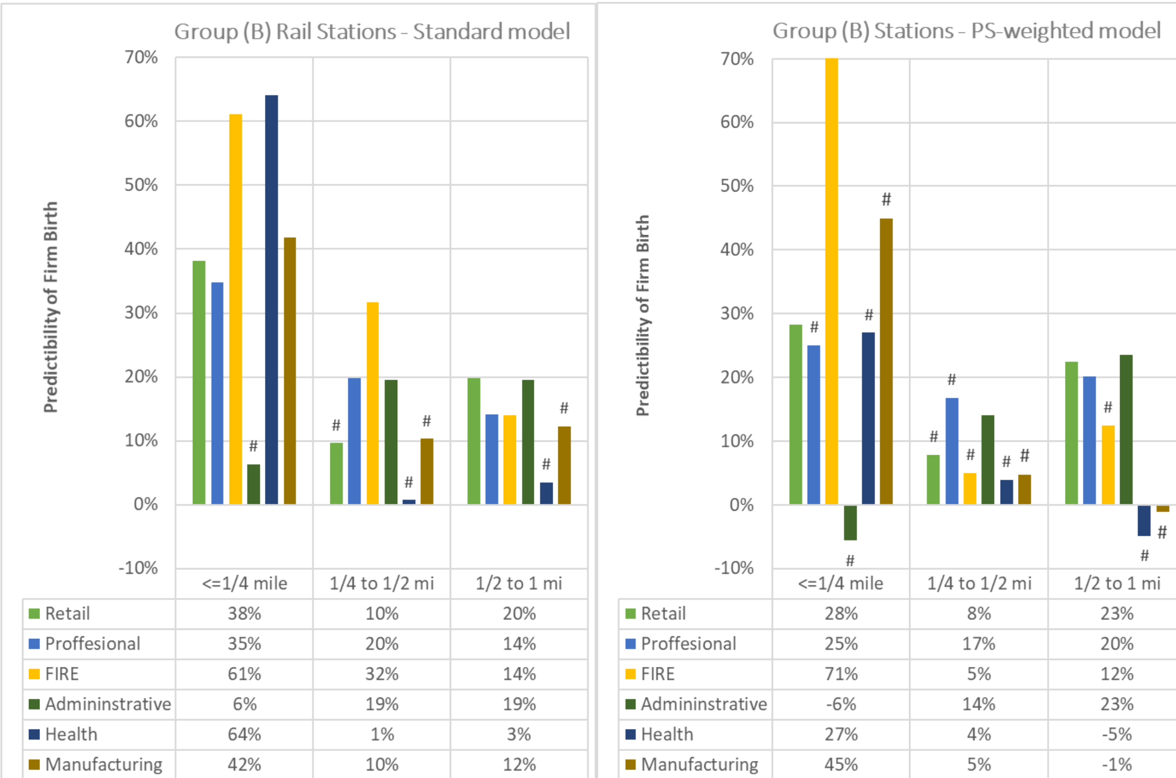
Table 19. PS-weighted negative binomial estimated coefficients: firm birth by selected industry sectors

Dependent variable: number of firm births	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(1) Manufacturing
Distance to Rail station (in mi)	-0.019***	-0.035***	-0.026***	-0.039***	-0.029***	-0.031***
Within <=1/4 mile (Group B)	0.249*	0.223	0.539***	-0.058	0.239	0.371
Within 1/4 to 1/2 mi (Group B)	0.075	0.155	0.049	0.132*	0.038	0.046
Within 1/2 to 1 mi (Group B)	0.203*	0.184**	0.117	0.211***	-0.05	-0.011
Accessibility ratio	0.636***	-0.166	-0.377	-0.04	0.186	0.121
Population per sq. mi. (in 1000s)	-0.011	-0.043***	-0.031***	-0.046***	-0.007	-0.042***
Employee per sq. mi. (in 1000s)	0.017	0.099***	0.066***	0.121***	0.018	0.046*
Number of firms	0.037***	0.043***	0.042***	0.020***	0.037***	0.028***
Firms in the same industry sector	0.085***	0.071***	0.110***	0.309***	0.129***	0.183***
Firm closures	-0.032***	-0.031***	-0.028***	-0.005***	-0.030***	-0.020***
Median HH Income (in $1000s)	0.000	0.006***	0.008***	0.015***	0.003***	0.000
Unemployment rate	-0.749**	1.502***	1.652***	4.123***	0.683*	-0.45
Percent college educated	-0.099	1.001***	0.450***	-0.410***	1.131***	0.057
Percent African-American	0.459***	0.843***	0.710***	0.698***	0.702***	0.265***
Median housing rent (in $1000s)	0.044	0.146***	0.103***	0.250***	0.052	0.086*
Distance to highway (in mi)	-0.076***	-0.043***	-0.081***	-0.018**	-0.083***	-0.066***
Distance to CBD (in mi)	0.022***	0.040***	0.026***	0.033***	0.017***	0.033***
Property tax (in $1000)	-0.100**	-0.177***	-0.007	0.046*	-0.123**	0.006
Constant	-2.259***	-3.167***	-3.500***	-3.334***	-3.272***	-3.777***
chi2	10296.883	19451.951	17363.326	46286.309	14753.12	8880.405
N. of cases	101859	101859	101859	101859	101859	101859

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

To compare the results across the two regression methods, Figure 8 shows the predicted probabilities of effect on firm birth by industry sector, within each of the three station buffers of group B stations (stations opened between 1990 and 1998). The probability effects are calculated from the estimated coefficients in Table 18 and Table 19. For example, holding all else equal, blocks located within a *quarter mile* radius from a rail station in group B are 61% more likely to experience a firm birth belonging to FIRE sector compared to control areas located more than a mile from the stations (see Figure 8). Evidently, in the *half to one mile* buffers, the probability of firm birth is positive in all industry-specific models in the case of group B stations. Positive effects on firm birth have extended up to a mile for stations in group B, yet the effect beyond the half-mile threshold is not as high as the one experienced in blocks located within the *quarter mile* buffer of the rail stations (see Figure 8).

Figure 8. Predicted probability of firm birth by proximity to group B rail stations across selected NAICS industry sectors, comparing results from two regression methods: the standard NB (left side) and the PS-weighted NB (right side)



Standard model	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	38%	10%	20%
Proffesional	35%	20%	14%
FIRE	61%	32%	14%
Admininstrative	6%	19%	19%
Health	64%	1%	3%
Manufacturing	42%	10%	12%

PS-weighted model	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	28%	8%	23%
Proffesional	25%	17%	20%
FIRE	71%	5%	12%
Admininstrative	-6%	14%	23%
Health	27%	4%	-5%
Manufacturing	45%	5%	-1%

Note: The y-axis shows the predicted effect of firm birth relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients using ($e(\beta i) - 1$), where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989, and group B stations are those opened between 1990 and 1999. The symbol (#) refers to the statistically insignificant values.

Transit-to-auto accessibility is another control variable tested in the industry-specific regression in Table 18. Across all industry sectors, there is a positive and statistically-significant relationship between the ratio of transit-to-auto accessibility and firm birth. The estimated coefficients for transit-to-auto accessibility range from β=0.414 in the case of professional services to the particularly high value of β=1.084 for retail trade sector (see Table 18). These positive effects suggest that areas more accessible by transit service (whether rail or bus service) experience higher number of firm births across all industry sectors, especially retail firms.

Finally, the industry-specific regression in Table 18 include a measure of the total number of firms, and another measure of the number of firms in the same industry sector. The number of existing firms within a block (urbanization economies) positively influence firm birth across all industry sectors; the estimated coefficients are small in magnitude, however, ranging between β=0.002 in the case of administrative sector to β=0.014 in the case of manufacturing sector. The number of existing firms in the same industry within a block (localization economies) has a positive and statistically-significant influence on firm birth in the case of three out of the six industry sectors (i.e. administrative sector, β= 0.196; manufacturing sector, β= 0.039, and health sector, β= 0.004). The only negative and statistically-significant relationship between the number of existing firms in the same industry within a block and firm birth is found in the case of professional service industry sector. In the case of professional-service sector, the fact that the number of firms in the same sector negatively predict the number of firm births suggests that the competition effect among professional-service firms overrides any localized agglomeration economies.

4.2. Rail transit impact on firm closure

This section examines the pattern of firm closures in relation to proximity to rail stations. The firm birth analysis in the previous section includes the number of firm closures as a control variable since they may influence the number of firm birth, as discussed earlier. One may argue that the analysis of firm closure as an outcome can be redundant since higher incidents of firm birth lead to higher incidents of firm closure in a process called the creative destruction (discussed in Chapter 2). However, the relationship between firm birth and firm closure is not straightforward. For a verity of reasons, some areas experience higher urban densities (firm agglomeration) over time than others, which means there are different relationships between firm birth and firm closure across urban areas. That is, areas that experience higher number of firm births compared to firm closures over time get denser.

The variation between firm birth and firm closure across urban areas can be more apparent at the micro-level (e.g. Census block). Theoretically, neighborhoods that offer a more sustainable economic environment to firms should experience a lower rate of firm closure to firm birth, compared to neighborhoods that are less economically attractive. The purpose of closure analysis, therefore, is to see whether or not areas within the close proximity to passenger rail stations exhibits lower probability of firm closure relative to the probability of firm birth (predicted in the previous section), compared to areas farther away from rail stations.

In general, areas experiencing high rates of firm birth also experience high rates of firm closure, as previously discussed. Table 20 shows that areas within a *quarter mile* from passenger rail stations experience the highest number of firm closures across all firm-size categories within the study area from 1991 to 2009. The same area, however, experienced the highest number of firm births during the same period (see Table 15 in the firm birth section). The economic trend is not positive over time, however, in areas near the stations and in the study area when looking at both the number of firm closures and the number of firm births.

Table 20. Number of firm closures per square mile by distance from Station: all study area Jurisdictions (1991-2009)

Distance to Station	All Firm Closures	Sole Proprietor	Five or Less Employees	More than 5 Employees
Within 1/4 mile	148.0	31.5	109.2	38.7
1/4 to 1/2 mile	92.4	21.6	68.8	23.6
1/2 to 1 mile	53.2	13.5	39.7	13.5
More than 1 mile	17.0	4.9	13.2	3.8
Average for all blocks	21.1	5.9	16.2	4.9

Source: NETS data, birth densities computed using ArcGIS.

Overall, the study area experienced an economic decline in the period between 1991 and 2009. Within the study area, the number of firm births per square mile was lower than the number of firm closures per square mile in nearly each year between 1991 and 2009 (see Figure 9). On the other hand, areas near the passenger rail stations experienced higher number of firm births compared to firm closures for longer periods than the study area. Figure 10 compares the number of firm closures and firm births per square mile for Census blocks located within a mile of the passenger rail stations; these blocks have experienced lower number of firm closures compared to firm births for several years during the period between 1991 and 2008. Controlled statistical analysis is needed to test whether or not areas within proximity to rail stations indeed experience lower probability of firm closure relative to the probability of firm birth (predicted in the previous section), compared to control areas located more than a mile from the rail stations.

Figure 9. Number of firm closures and firm births per square mile within the study area (1991-2008)

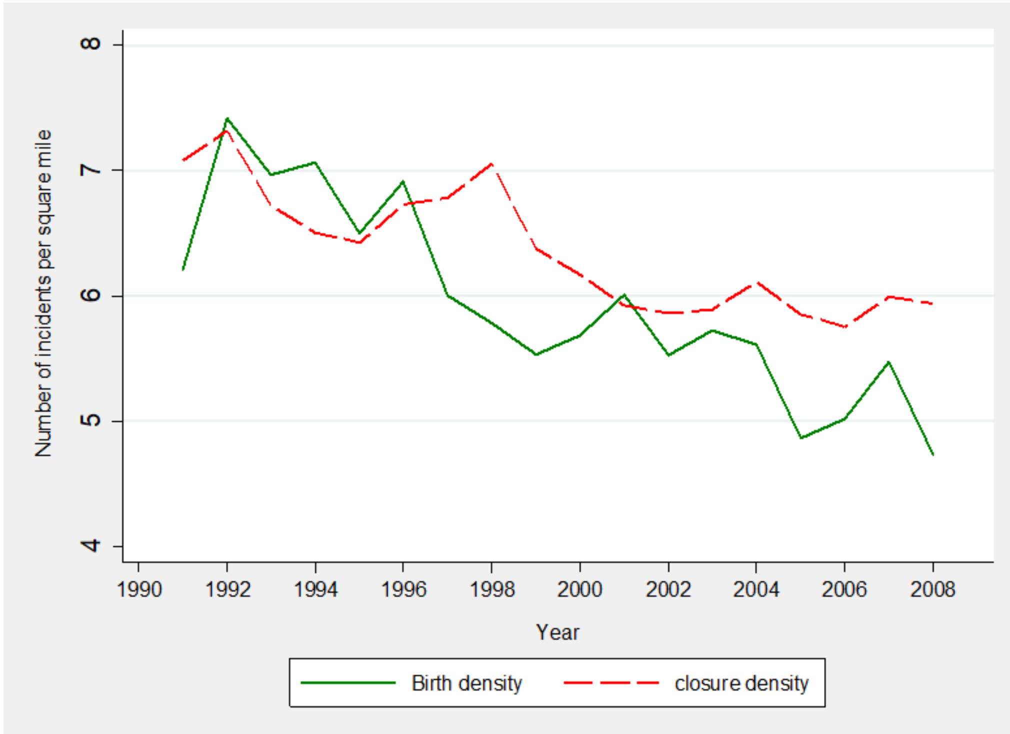
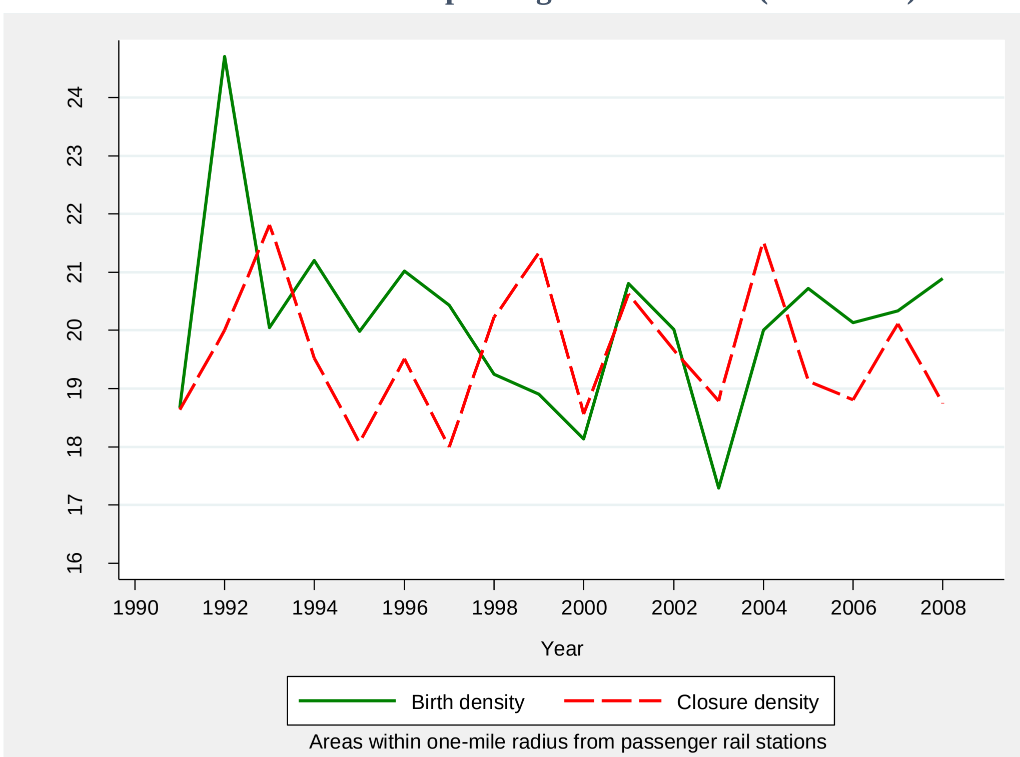


Figure 10. Number of firm closures and firm births per square mile for Census blocks located within one mile of the passenger rail stations (1991-2008)



This section compares the predicted probability of firm birth to the predicted probability of firm closure in three ways:

1) between treatment (the three station buffers) and control areas;
2) across different firm-size categories within the treatment areas; and
3) across selected industry sectors within the treatment areas.

The estimated coefficients from firm closure and firm birth models are compared to determine the net gains for the three comparisons. The following two subsections examine the impact of proximity to passenger rail stations on the probability of firm closure using both regression methods (i.e. the standard NB model and the PS-weighted NB model). Similar to firm birth analysis, the regression models are carried out to examine firm closure across four firm-size categories and across six selected industry sectors.

4.2.1. Firm closure by size category: regression results

This section starts with a discussion on the firm closure regression results of the variables of proximity to rail stations. The section ends with a discussion on the regression results of the control variables on agglomeration, socio-economic characteristics, policy environment, and spatial context. As discussed earlier, the advantage of the firm closure analysis is to determine whether or not areas near the rail stations have higher probability of firm birth relative to the probability firm closure, compared to the control areas (located more than a mile from the stations). For a given area, a positive-sum of the probability of firm birth and the probability of firm closure indicates, on average, a higher probability of firm retainment (i.e. increase in firm agglomeration, which consequently leads to higher economies of scale as discussed in Chapter 2). This section calculates the sum of the probability of firm birth and the probability of firm closure for areas near the passenger rail stations by comparing the estimated coefficients generated from the firm birth and firm closure regressions.

Table 21 and Table 22 show the regression results of firm closure using the standard and the PS-weighted negative binomial regressions, consecutively. Overall, areas within proximity to rail stations experience higher number of firm closures because of high number of firm births as shown in the previous section. This trend is confirmed by the negative estimated coefficients of the continuous distance-to-station variable in Table 21 across all firm-size categories (ranging between $\beta=-0.061$ and $\beta=-0.068$). The negative distance-to-station coefficients mean that there are higher numbers of firm closures in areas within close proximity to the rail stations. The positive estimated coefficients of the three station-buffer variables in Table 21 and Table 22 also confirm the existing positive association between areas in close proximity to the rail stations and the number of firm closures.

However, the estimated coefficients of the three station-buffer variables suggest that there is an exception to the positive association between station proximity and firm closure. As shown in Table 21, Census blocks within a *half to one mile* buffer of the mature rail stations (group A stations) have negative and statistically-significant probability of firm closure (the estimated coefficients ranging between $\beta=-0.387$ for firms with sole proprietor and $\beta=-0.271$ for firms with more than five employees). Without comparing the estimated probabilities of firm closure and firm birth, it is impossible to know whether or not areas near the rail stations experienced a positive economic gain. For instance, areas near rail stations that have positive predicted probability of firm closure compared to control areas may have even higher predicted probability of firm birth, which consequently indicate that these areas have experienced a positive probability of firm retainment. Therefore, this section calculates the total sum of the predicted probabilities of firm closure and firm birth for the three station buffers across the four firm-size categories.

Table 21. Regression coefficients of the standard negative binomial method: firm closure by firm size

Dependent variable: number of firm closures	(1) **All Firms**	(2) **Sole Proprietor**	(3) **Firms > 1 employee**	(4) **Firms <=5 employees**	(5) **Firms >5 employees**
Distance to Rail station (in mi)	-0.062***	-0.068***	-0.066***	-0.063***	-0.061***
Group A stations: within <=1/4 mile	0.348***	0.233**	0.441***	0.284***	0.672***
Group A stations: within 1/4 to 1/2 mi	0.152**	0.062	0.175**	0.106*	0.280***
Group A stations: within 1/2 to 1 mi	-0.323***	-0.387***	-0.353***	-0.378***	-0.271***
Group B stations: within <=1/4 mile	0.658***	0.588***	0.697***	0.662***	0.728***
Group B stations: within 1/4 to 1/2 mi	0.485***	0.369***	0.461***	0.495***	0.295***
Group B stations: within 1/2 to 1 mi	0.270***	0.244***	0.239***	0.284***	0.118
Group C stations: within <=1/4 mile	1.384***	1.731***	1.071**	1.730***	0.490
Group C stations: within 1/4 to 1/2 mi	1.176***	1.287***	1.035***	1.249***	0.959**
Group C stations: within 1/2 to 1 mi	0.868***	0.769***	0.732***	0.860***	0.655***
Accessibility ratio	1.369***	0.854***	1.283***	1.274***	1.210***
Population per sq. mi. (in 1000s)	-0.045***	-0.056***	-0.037***	-0.045***	-0.043***
Employee per sq. mi. (in 1000s)	0.101***	0.138***	0.074***	0.108***	0.048***
Number of firms	0.005***	0.005***	0.005***	0.005***	0.007***
Firm births	0.005***	0.009***	0.004***	0.006***	0.001
Average age of firms	0.002**	0.004***	0.002**	0.001	0.020***
Median HH Income (in $1000s)	0.019***	0.018***	0.015***	0.020***	0.003***
Unemployment rate	4.626***	5.192***	3.759***	5.049***	1.865***
Percent college educated	-0.483***	-0.608***	-0.140*	-0.486***	0.121
Percent African-American	0.503***	0.805***	0.472***	0.598***	0.136**
Median housing rent (in $1000s)	0.534***	0.581***	0.466***	0.564***	0.379***
Distance to highway (in mi)	-0.066***	-0.027***	-0.092***	-0.052***	-0.164***
Distance to CBD (in mi)	0.031***	0.048***	0.034***	0.033***	0.040***
Property tax (in $1000)	0.113***	0.072***	0.058*	0.096***	-0.065
Constant	-2.388***	-3.755***	-1.545***	-2.741***	-0.494***
ln_r Constant	0.424***	0.719***	0.627***	0.430***	1.187***
ln_s Constant	-0.528***	0.119*	-1.002***	-0.455***	-1.758***
N. of cases	116820	116820	116820	116820	116820
Log Likelihood	-110399.3	-66258.33	-86168.574	-103193.656	-34541.043
chi2	24121.688	13418.113	17184.323	22147.153	4789.795

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

Table 22. Regression coefficients of the PS-weighted negative binomial method: firm closure by firm size.

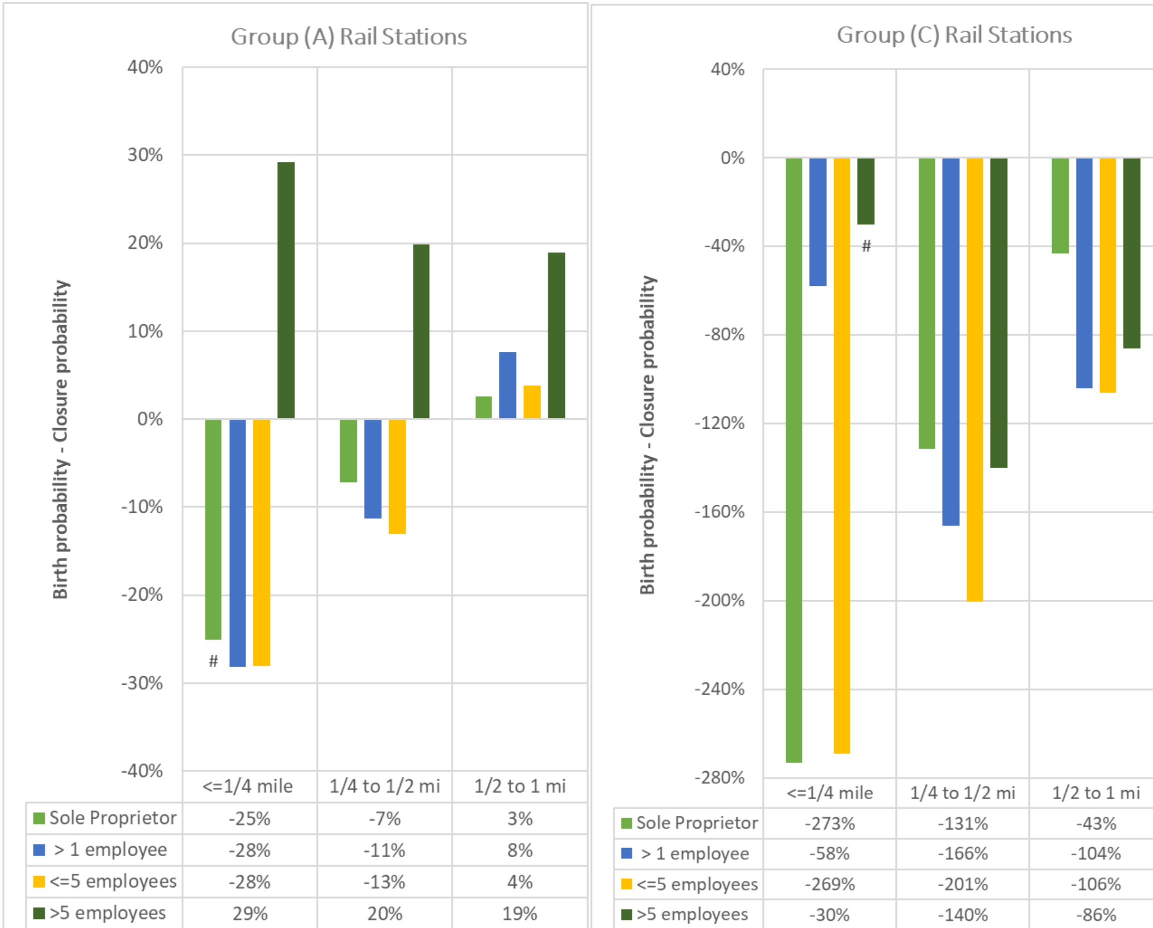
Dependent variable: number of firm closures	(1)	(2)	(3)	(4)	(5)
	All Firms	**Sole Proprietor**	**Firms > 1 employee**	**Firms <=5 employees**	**Firms >5 employees**
Distance to Rail station (in mi)	-0.035***	-0.047***	-0.032***	-0.036***	-0.040***
Within <=1/4 mile (Group B)	0.209**	0.108	0.320***	0.171*	0.557***
Within 1/4 to 1/2 mi (Group B)	0.310***	0.271***	0.294***	0.299***	0.242**
Within 1/2 to 1 mi (Group B)	0.261***	0.286***	0.236***	0.269***	0.151
Accessibility ratio	0.782***	0.745***	0.733***	0.770***	0.570*
Population per sq. mi. (in 1000s)	-0.044***	-0.060***	-0.040***	-0.046***	-0.057***
Employee per sq. mi. (in 1000s)	0.100***	0.145***	0.079***	0.109***	0.062**
Number of firms	0.065***	0.028***	0.064***	0.057***	0.041***
Firm births	0.061***	0.080***	0.032***	0.069***	-0.011***
Average age of firms	0.007***	0.005***	0.009***	0.006***	0.016***
Median HH Income (in $1000s)	0.012***	0.015***	0.009***	0.013***	0.001
Unemployment rate	4.909***	5.841***	3.680***	5.384***	1.614***
Percent college educated	-0.660***	-0.895***	-0.437***	-0.666***	-0.375**
Percent African-American	0.309***	0.519***	0.264***	0.365***	0.088
Median housing rent (in $1000s)	0.284***	0.302***	0.257***	0.295***	0.258***
Distance to highway (in mi)	-0.036***	-0.012	-0.059***	-0.026***	-0.164***
Distance to CBD (in mi)	0.021***	0.038***	0.018***	0.023***	0.029***
Property tax (in $1000)	0.108***	0.055*	0.091***	0.095***	-0.019
Constant	-2.758***	-3.872***	-2.726***	-3.007***	-2.880***
chi2	121434.305	67509.632	66937.121	114516.032	6132.932
N. of cases	101859	101859	101859	101859	101859

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

Figure 11 shows the sum of the predicted probabilities of firm birth and firm closure across the four firm size categories for each of the three station buffers of the mature rail stations (group A stations), and the more recently opened stations (group C stations). The sums of the predicted probabilities in Figure 11 are relative to the control areas that are located more than a mile from the rail stations. The sum of the predicted probabilities is calculated by deducting firm closure predicted-probabilities from firm birth predicted-probabilities, using the estimated coefficients of distance-to-station dummy variables (the *quarter mile*, *quarter to half mile*, and *half to one mile* buffers). The firm birth and firm closure analyses show that the probability of areas near the mature rail stations (group A) to retain larger firms is much higher than smaller ones.

Larger firms with more than five employees have, on average, the highest positive-sum of the predicted probabilities of firm birth and firm closure in areas within short walking distance to a passenger rail stations in group A. Figure 11 shows that blocks located within up to one mile of the mature rail stations (group A) have experienced a considerably higher predicted probability of firm retainment (firm birth - firm closure) of firms with more than five employees compared to the control areas located more than a mile of the rail stations. For instance, the probability of the *quarter mile* buffer of the group A stations to retain firms is 29%, all else held equal (see Figure 11), which is calculated by subtracting the estimated probability of firm closure (96%) from the estimated probability of firm birth (125%).

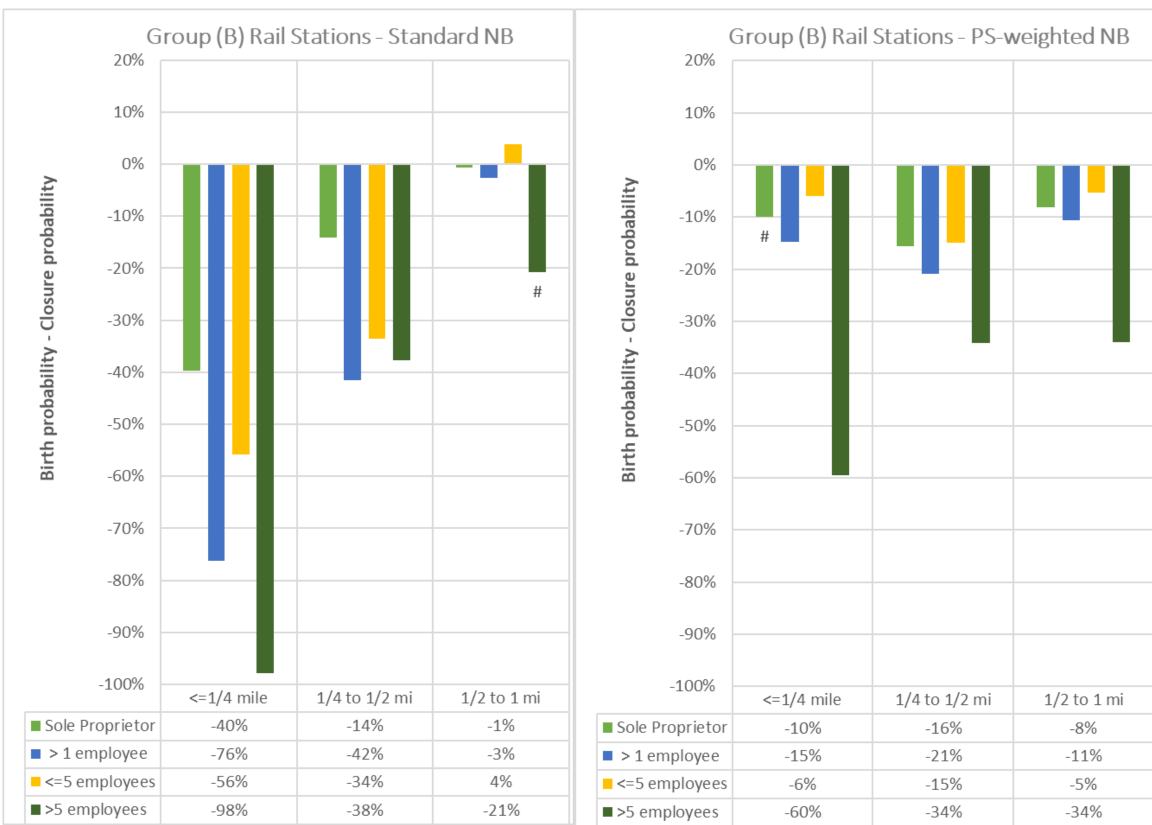
Figure 11. The difference between the predicted probability of firm birth and firm closure by firm size and distance from mature rail stations (group A stations)



Note: The y-axis shows firm birth-to-closure net effects (ratio) relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients of firm birth and firm closure models using [birth(e(βi) – 1) - closure(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989. Group C passenger rail stations are those opened after 2000. The symbol (#) indicates the statistical insignificance of the estimated coefficients of both firm birth and firm closure.

For the less mature stations, however, the areas near rail stations have experienced negative sums of the probabilities of firm birth and closure across nearly all firm sizes. The dominantly negative sums of the probabilities shown in Figure 12 suggest that areas near group B stations exhibit, on average, lower probabilities to retain firms compared to areas located more than a mile of the stations. That is, the probability of firm birth minus the probability of firm closure yield negative firm retainment probabilities in the three group B station buffers. The negative sign of the predicted probability to retain firms is consistent across the finding of the two regression methods but the magnitude is lower in the PS-weighted method (see Figure 12). Areas near the recently opened stations (group C stations) also experienced negative sums of probabilities of firm birth and firm closure (birth – closure) across all firm sizes. The negative sum of the probability of firm birth-closure is not surprising for the six rail stations in group C. All group C stations are located within Prince George's County, which up to recent years had zoning and land use policies unwelcoming to transit oriented development.

Figure 12. The difference between the predicted probability of firm birth and firm closure by firm size and distance from group B stations (comparing the results of the Standard and the PS-weighted NB methods)



Note: The y-axis shows firm birth-to-closure net effects (ratio) relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients of firm birth and firm closure models using [birth(e(βi) – 1) - closure(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B passenger rail stations are those opened between 1990 and 1999. The symbol (#) indicates the statistical insignificance of the estimated coefficients of both firm birth and firm closure.

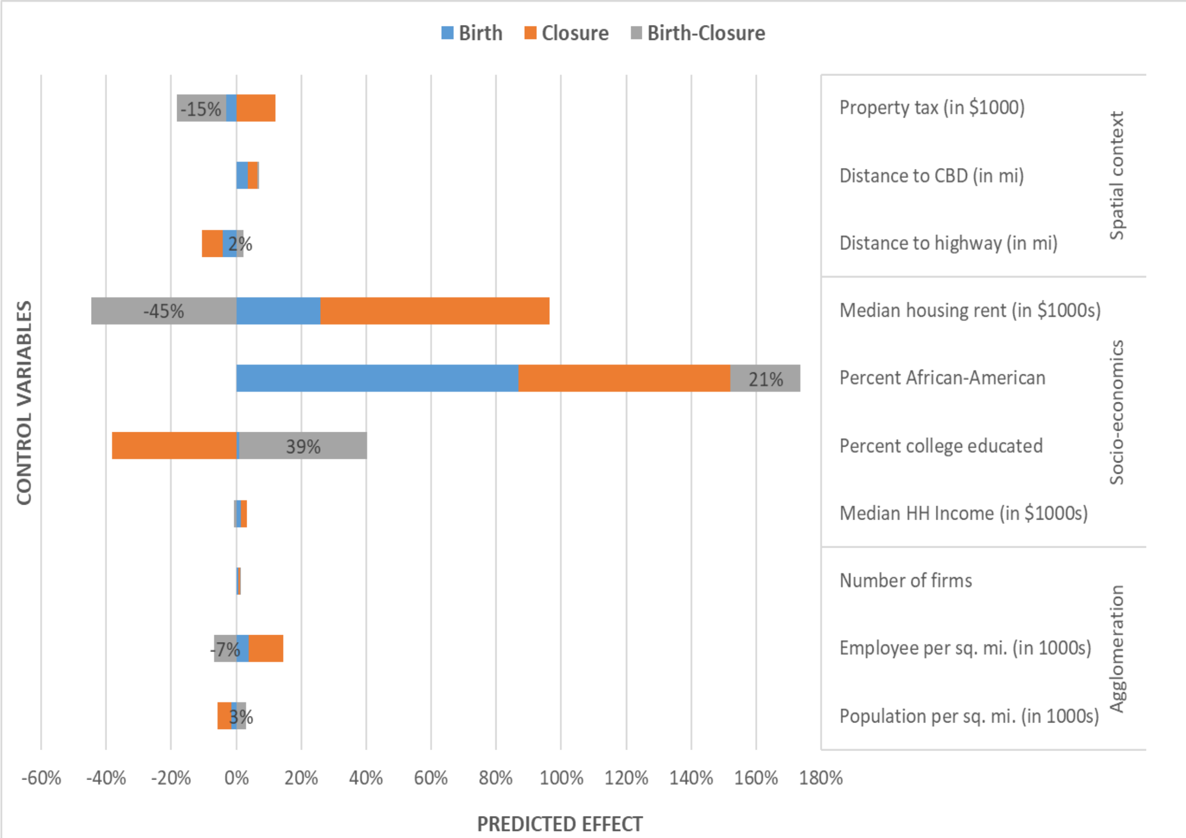
Turning to the other control variables, the association between firm closure and population density is negative across all firm size categories (βs ranging between -0.037 and -0.045), which means the higher the population density in a Census block, the lower the likelihood of that block to experience firm closure, all else held equal. The population density coefficient for all firms (-0.045) in Table 21 suggests that Census blocks are 4% less likely to experience firm closure for every additional unit of population density of a thousand persons per square mile (the unit of population-density variable is in thousands of population). On the other hand, the higher the employment density in a Census block, the higher the likelihood of firm closure (coefficients ranging between 0.048 and 0.101).

In a Census block, the association is positive but small (coefficient ranging between 0.005 and 0.007) between the number of existing firms and the number of firm closures (see Table 21). Similar positive association is found between the number of firm births and the number of firm closures (ranging between 0.005 and 0.009). The average age of firms in a Census block also has a small but positive association with the probability of firm closure, as suggested by the positive sign of the coefficients in Table 21 (ranging between 0.002 and 0.020). These positive estimated coefficients of existing firms and average-firm-age variables suggest that blocks with well-established existing firms have slightly higher likelihood of firm closure compared to blocks with less-established firms.

Most of the coefficients of the socio-economic variables (i.e. unemployment rate, the share of African-American population, household income, and housing rent) show a positive association with incidents of firm closure except for the variable of the level of education (see Table 21). The percentage of population that is college educated in an area is negatively associated with the number of firm closure for smaller firms with five or fewer employees (-0.486). However, this association is positive but statistically insignificant for larger firms with more than five employees (0.121), as shown in Table 21. The sums of the predicted probabilities of firm closure and the predicted probabilities of firm birth for the control variables provide more relevant predictions of net gain or loss in firm density, as discussed earlier.

Figure 13 summarizes the predicted effect of firm birth, firm closure, and the sum of the probabilities of firm birth and firm closure for the control variables related to agglomeration, socio-economic characteristics, and spatial context. The sum of birth and closure predicted probability is a relevant predicted effect to pay attention to. For any given control variable, a positive sum of birth-closure probabilities suggests an overall positive influence on firm retainment. For example, while the predicted effect of population density is negative in the firm birth analysis of all firms (-0.015), the effect is also negative in the firm closure analysis (-0.045). Therefore, as shown in Figure 13, the sum of firm birth and firm closure predicted probabilities (birth-closure) for population density is overall positive suggesting that blocks with higher population densities are more likely to retain firms compared to blocks with lower population densities, all else held equal.

Figure 13. Predicted probabilities of firm birth, firm closure, and the difference between the two (probability of retainment) for selected control variables (all firms)



Note: The x-axis shows the predicted effect on firm birth, firm closure, and the net birth-to-closure relative to control Census blocks, all else held equal. The y-axis shows the predicted effects of each control variable except transit-related variables. The percentages are calculated from the regression coefficients of firm birth and firm closure models. The birth-to-closure net effects are calculated using [birth(e(βi) – 1) - closure(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective control variable.

A similar interpretation of predicted probabilities to retain firms (birth probability – closure probability) is true for other control variables presented in Figure 13. For example, the overall predicted effect of property tax is negative (-15%) because the predicted effect on firm closure is much higher (0.113) than the predicted effect on firm birth (-0.032) for every additional thousand U.S. dollars of property tax.

The firm retainment probability of distance to CDB is near zero because this distance variable has a positive association with the probability of firm birth (0.034) and the probability of firm closure (0.031) that are almost equal in magnitude. The distance to CBD results suggest that areas within the study area that are near CBD have no impact on firm retainment compared to areas farther away from the CBD, holding all else equal. On the contrary, blocks within close proximity to a highway exit have experienced a positive firm retainment, which is suggested by the negative association between the distance to highway and the probability of firm birth (-0.043) that is smaller in magnitude than the negative association with the probability of firm closure (-0.066), as shown in Figure 13.

4.2.2. Firm closure by industry sector: regression results

This section examines firm closure across the six selected industry sectors within the study area in relation to proximity to passenger rail stations. As previously indicated, areas within proximity to rail stations may experience high number of firm closures because of the high number of firm births within the same areas. The objective of the closure analysis by industry sectors is to determine whether industry sectors have different probability of firm retainment (probability of firm birth - probability of firm closure) in areas within short walking distance to the passenger rail stations. In other words, the analysis observes what industry sectors are more likely to benefit more from the improved accessibility provided by the rail stations.

Table 23 and Table 24 show the regression results of the firm closure analysis across the six industry sectors using the standard and the PS-weighted negative binomial models. The dependent variable is the number of firm closures in each industry sector regressed on the control variables including distance to rail station, agglomeration, socio-economic, and spatial context variables. Across the six industry sectors, the number of firm closures in areas near the rail stations are, on average, higher than the number of firm closures in areas located more than a mile of the stations (control areas). This is indicated by the negative sign of the estimated coefficients of the continuous distance-to-station variable (coefficients ranging between -0.061 for the administrative sector and -0.080 for professional services, as shown in Table 23). There are mixed associations, however, between areas within the three rail-station buffers and the number of firm closures across the six industry sectors.

Table 23. Regression coefficients of the standard negative binomial method: Firm Closure by selected industry sectors

Dependent variable: count of firm closures	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail Station	-0.068***	-0.080***	-0.065***	-0.061***	-0.068***	-0.072***
Group A stations: within <=1/4 mile	0.421***	0.660***	0.755***	0.16	0.345**	0.365*
Group A stations: within 1/4 to 1/2 mi	0.136	0.401***	0.188*	0.138*	0.143	0.146
Group A stations: within 1/2 to 1 mi	-0.532***	-0.427***	-0.562***	-0.289***	-0.369***	-0.404***
Group B stations: within <=1/4 mile	0.710***	0.941***	0.873***	0.421***	0.852***	0.477**
Group B stations: within 1/4 to 1/2 mi	0.407***	0.622***	0.442***	0.362***	0.419***	0.298*
Group B stations: within 1/2 to 1 mi	0.289***	0.306***	0.117	0.169***	0.160*	0.238*
Group C stations: within <=1/4 mile	0.816	0.862	1.376*	1.672***	0.216	1.238
Group C stations: within 1/4 to 1/2 mi	0.867**	1.064***	1.181***	0.708**	1.280***	0.735
Group C stations: within 1/2 to 1 mi	0.488**	0.698***	0.544**	0.653***	0.512**	0.445
Ratio of Transit to Auto Accessibility	1.436***	0.719***	0.869***	0.913***	0.708***	0.860**
Population per sq. mi. (in 1000s)	-0.036***	-0.048***	-0.049***	-0.058***	-0.019***	-0.079***
Employee per sq. mi. (in 1000s)	0.077***	0.110***	0.106***	0.131***	0.064***	0.100***
Number of firms	0.008***	0.000	0.001*	-0.009***	-0.001*	0.007***
Number of firms in the same sector	0.005**	0.024***	0.032***	0.292***	0.079***	0.051***
Firm births	0.001	0.008***	0.009***	0.005***	0.014***	0.004**
Average age of firms	0.008***	0.007***	0.009***	0.006***	0.012***	0.017***
Median HH Income in $1000s	0.010***	0.012***	0.011***	0.012***	0.009***	0.006***
Unemployment rate	3.232***	3.989***	4.186***	4.403***	3.043***	3.217***
Percent college educated	-0.327***	0.737***	0.355***	-0.261***	0.461***	-0.322*
Percent African American	0.482***	0.716***	0.635***	0.711***	0.817***	0.260***
Median housing rent in $1000	0.461***	0.420***	0.570***	0.417***	0.397***	0.440***
Distance to nearest highway exit in miles	-0.087***	-0.108***	-0.137***	-0.036***	-0.102***	-0.100***
Distance to CBD in miles	0.043***	0.070***	0.048***	0.052***	0.046***	0.055***
Property tax in 2010	0.031	-0.156***	-0.055	0.038	-0.051	0.073
Constant	-1.490***	-2.736***	-2.226***	-3.515***	-2.631***	-1.976***
ln_r _Cons.	1.129***	1.010***	1.123***	1.631***	1.307***	1.761***
ln_s_Cons.	-1.463***	-1.177***	-1.550***	0.652***	-1.115***	-1.436***
N. of cases	116820	116820	116820	116820	116820	116820
Log Likelihood	-37944.921	-33369.603	-28548.174	-41226.376	-23365.097	-13332.614
chi2	5825.457	7660.483	5783.905	15457.923	5241.246	2140.828

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

Table 24. Regression coefficient of the PS-weighted negative binomial method: firm closure by selected industry sectors

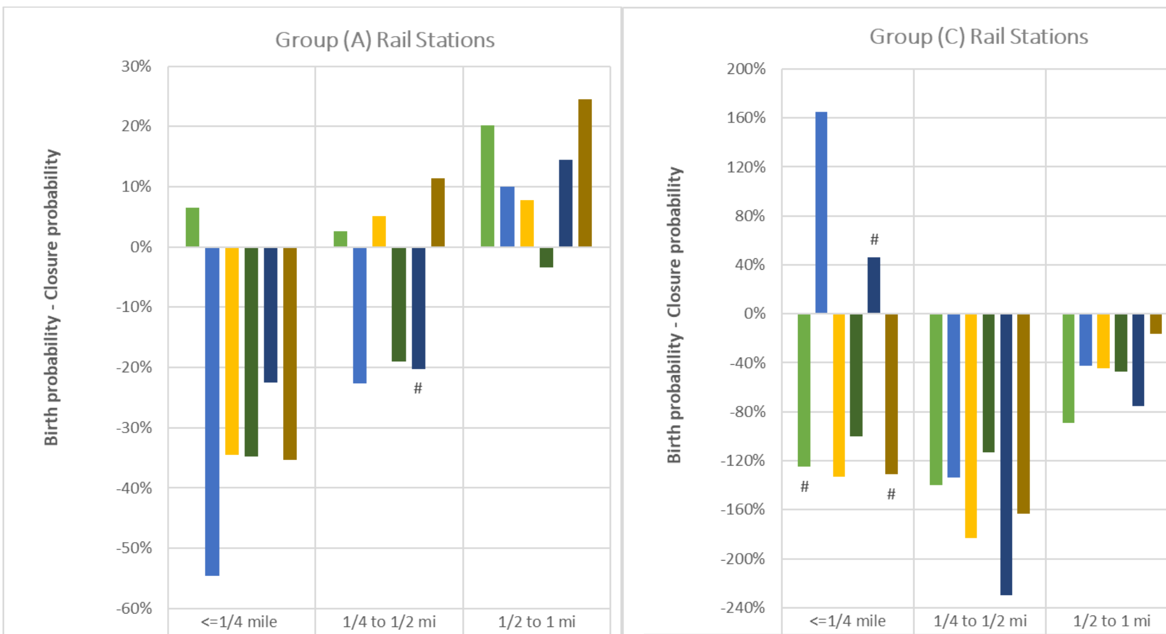
Dependent variable: number of firm closures	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail station (in mi)	-0.036***	-0.055***	-0.033***	-0.047***	-0.047***	-0.043***
Within <=1/4 mile (Group B)	0.242	0.661***	0.492**	0.237	0.758***	0.370*
Within 1/4 to 1/2 mi (Group B)	0.213**	0.481***	0.232*	0.290***	0.499***	0.001
Within 1/2 to 1 mi (Group B)	0.332***	0.374***	0.149*	0.238***	0.118	0.154
Accessibility ratio	0.706***	0.524**	0.181	0.862***	0.657**	0.626
Population per sq. mi. (in 1000s)	-0.040***	-0.058***	-0.048***	-0.049***	-0.021**	-0.070***
Employee per sq. mi. (in 1000s)	0.087***	0.126***	0.107***	0.123***	0.066***	0.092***
Number of firms	0.008***	0.007***	0.003*	-0.006***	0.006***	0.008***
Firms in the same industry sector	0.187***	0.166***	0.291***	0.395***	0.196***	0.379***
Firm births	0.022***	0.027***	0.026***	0.035***	0.025***	0.008**
Average age of firms	0.010***	0.009***	0.010***	0.006***	0.010***	0.014***
Median HH Income (in $1000s)	0.008***	0.009***	0.009***	0.012***	0.008***	0.006***
Unemployment rate	3.081***	3.948***	3.750***	4.916***	3.296***	2.951***
Percent college educated	-0.794***	0.086	-0.307*	-0.535***	0.229	-0.471**
Percent African-American	0.299***	0.617***	0.455***	0.533***	0.753***	0.213**
Median housing rent (in $1000s)	0.263***	0.241***	0.318***	0.323***	0.355***	0.353***
Distance to highway (in mi)	-0.056***	-0.070***	-0.083***	-0.029**	-0.071***	-0.077***
Distance to CBD (in mi)	0.025***	0.051***	0.031***	0.041***	0.027***	0.035***
Property tax (in $1000)	0.098**	-0.135***	0.03	0.02	-0.08	0.140**
Constant	-3.815***	-4.408***	-4.514***	-4.410***	-4.740***	-5.093***
chi2	26573.365	30450.604	25918.228	46315.151	19525.178	7294.18
N. of cases	101859	101859	101859	101859	101859	101859

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

Census blocks within the *half to one mile* buffer of the mature rail stations (group A stations) have negative estimated coefficients of firm closure across all the six industry sectors (ranging between -0.289 for administrative sector and -0.562 for FIRE sector, as shown in Table 23); the negative estimated coefficients of firm closure is somewhat expected since blocks within the same buffer had negative estimated coefficients of firm birth, as shown previously in the firm birth analysis. However, the estimated coefficients of firm closure of all remaining station buffers are positive (ranging between as low as 0.138 to as high as 1.672, as shown in Table 23 and Table 24). Therefore, the sum of the probability of firm birth and the probability of firm closure (probability of birth – probability of closure) within the three station buffers needs to be calculated to determine differences in the probability of firm retainment across the six industry sectors. The probability of firm retainment is a better measure of whether or not areas within close proximity to rail stations provide an overall economic benefit to firms, as explained earlier in this chapter.

Figure 14 shows the probability of firm retainment (birth – closure) within the three station buffers of the mature rail station (group A) as well as the most recently opened stations (group C) across the six industry sectors. Clearly, firms in the retail trade sector are the most likely to benefit from areas located within close proximity to the mature rail stations (group A). Unlike other sectors, the probability of firm retainment of retail firms is positive in all the three station buffers of group A stations (ranging between 3% to 20%) compared to blocks located more than a mile away from the rail stations, all else held equal (see Figure 14). For example, the probability of firm retainment in the *quarter mile* buffer of group A stations is 7%, which is calculated by deducting the probability of firm closure (e(0.421)-1=0.52=52%) from the probability of firm birth (e(0.463)-1=0.59=59%). Blocks within the *quarter to half mile* buffer and within the *half to one mile* buffer of group A stations also exhibit positive probability of firm retainment of manufacturing firms (11% and 25%, consecutively). There results suggest that retail and manufacturing firms benefit the most from better access to the labor force provided by passenger rail stations.

Figure 14. The difference between the predicted probability of firm birth and firm closure by industry sector and distance from rail stations in group A and C



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	7%	3%	20%
Proffesional	-55%	-23%	10%
FIRE	-35%	5%	8%
Admininstrative	-35%	-19%	-3%
Health	-23%	-20%	14%
Manufacturing	-35%	11%	25%

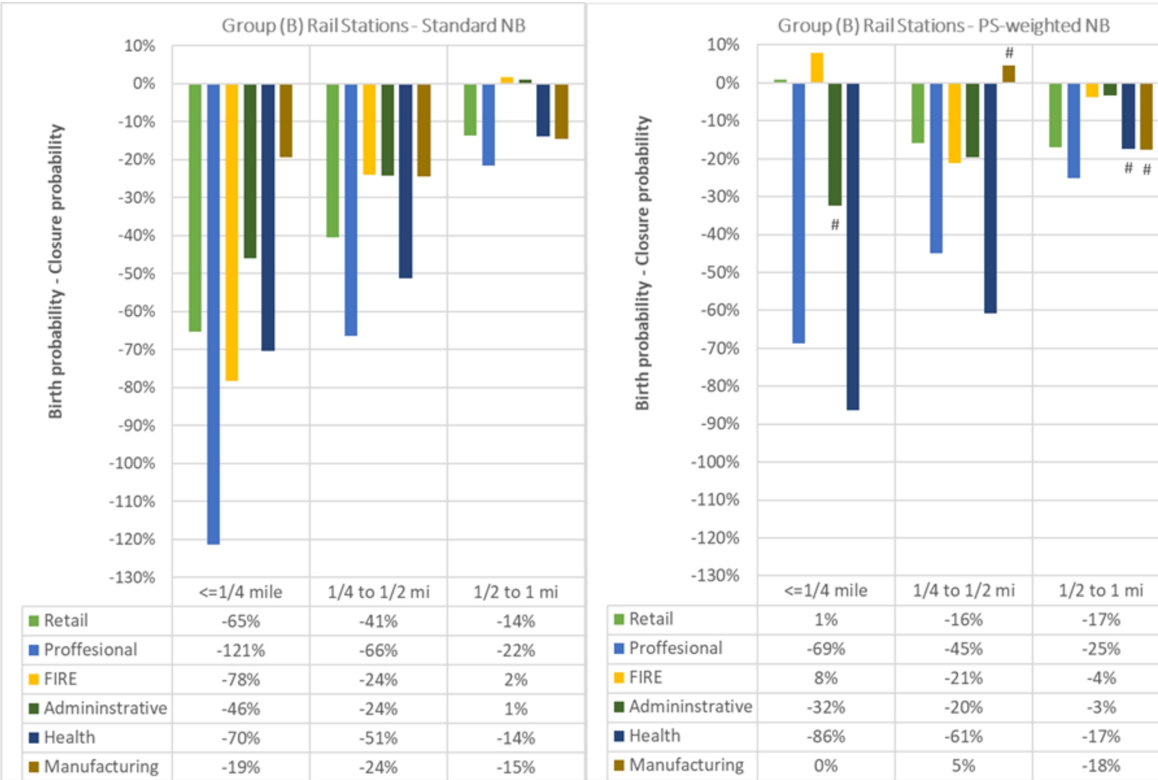
	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-125%	-140%	-89%
Proffesional	165%	-134%	-43%
FIRE	-133%	-183%	-44%
Admininstrative	-100%	-113%	-48%
Health	46%	-230%	-75%
Manufacturing	-131%	-163%	-16%

Note: The y-axis shows firm birth-to-closure net effects relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of firm birth and firm closure models using [birth(e(βi) – 1) - closure(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group (A) passenger rail stations are those opened between 1978 and 1989. Group C passenger rail stations are those opened between 2000 and 2004. The symbol (#) indicates the statistical insignificance of the estimated coefficients of both firm birth and firm closure.

Census blocks within a *half to one mile* distance of group A stations have, on average, experienced positive probability of firm retainment of all industry sectors (ranging between 8% and 25%) except the administrative sector (-3%), compared to control areas (see Figure 14). For the more recently opened stations (group C stations), the probability of firm retainment within the three station buffers are mostly negative, ranging between -16% and -230%, except in two cases. The probability of firm retainment is positive within the *quarter mile* buffer of group C stations for firms in the professional service sector (165%) as well as the health sector (46%), but the probability of firm closure and firm birth are both statistically insignificant for the health sector (see Figure 14).

The three group B station buffers also exhibit a dominantly negative probability of firm retainment. Figure 15 shows the calculated probabilities of firm retainment (probability of birth – probability of closure) within the three group B station buffers, using both regression methods. The probabilities of firm retainment are predominantly negative within group B station buffers (ranging between -121% to 2% using the standard NB method, and between -86% and 8% using the PS-weighted NB method). For example, one exception is the 8% positive probability of blocks located within the *quarter mile* buffer to retain firms belonging to FIRE sector compared to control areas, all else held equal (see Figure 15). The probability of firm retainment analysis by station maturity suggest that areas within close proximity to mature rail stations are more likely to experience gains in firm density essential to transit oriented development.

Figure 15. The difference between the predicted probability of firm birth and firm closure by industry sector and distance from group B stations (comparing the results of the Standard and the PS-weighted NB methods)



Group (B) Rail Stations - Standard NB

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-65%	-41%	-14%
Proffesional	-121%	-66%	-22%
FIRE	-78%	-24%	2%
Admininstrative	-46%	-24%	1%
Health	-70%	-51%	-14%
Manufacturing	-19%	-24%	-15%

Group (B) Rail Stations - PS-weighted NB

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	1%	-16%	-17%
Proffesional	-69%	-45%	-25%
FIRE	8%	-21%	-4%
Admininstrative	-32%	-20%	-3%
Health	-86%	-61%	-17%
Manufacturing	0%	5%	-18%

Note: The y-axis shows firm birth-to-closure net effects (ratio) relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the estimated coefficients of firm birth and firm closure models using [birth(e(βi) – 1) - closure(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B passenger rail stations are those opened between 1990 and 1999. The symbol (#) indicates the statistical insignificance of the estimated coefficients of both firm birth and firm closure.

Turning to the other control variables on agglomeration, socio-economic, and spatial context, the predicted probabilities of firm closure across the six industry sectors are mostly similar in direction to the closure probabilities discussed in the previous section on firm closure by firm size. For example, the firm closure estimated-coefficients of population density range between -0.019 and -0.079 (see Table 23). One exception that stands out is the estimated coefficients of the percentage of population that is college educated on firm closure for the retail sector (-0.327), administrative sector (-0.261), and manufacturing sector (-0.322). For firms in these three sectors, the predicted probability of firm closure in a Census block is lower, the higher is the education level (see Table 23). The negative association between the level of education and firm closure suggests that human capital is a key factor for the longevity (survival) of firms in the retail, administrative, and manufacturing sectors. On the other hand, for the level of education variable, the estimated coefficients of firm closure are positive for the FIRE sector (0.355), health sector (0.46), and professional services sector (0.737), suggesting that the higher the percentage of college-educated population in a Census block, the higher the number of firm closures of firms belonging to these three sectors. This high number of firm closures could merely be the result of high number of firm births, as discussed earlier. For each control variable, both probabilities need to be compared (firm birth and closure probabilities) to understand the overall influence on firm retainment.

Figure 16 and Figure 17 show the probability of firm retainment (probability of firm birth – probability of firm closure) for agglomeration, socio-economic, and spatial-context related variables by industry sector. For instance, for the administrative and health sectors, the association was small but negative between firm closure and the number of existing firms within a block (-0.009 and -0.001, consecutively), whereas the association with firm birth was positive (0.002 and 0.010, consecutively); therefore, the probability of firm retainment is 1% for firms belonging to the administrative and health sectors[1].

[1] For example, the 1% probability of the administrative sector is calculated by deducting the probability of firm closure from the probability of firm birth: [e(-0.009)-1] – [e(0.002)-1]=0.01=1%.

Figure 16. The difference between the predicted probability of firm birth and firm closure for agglomeration and spatial context variables by industry sector

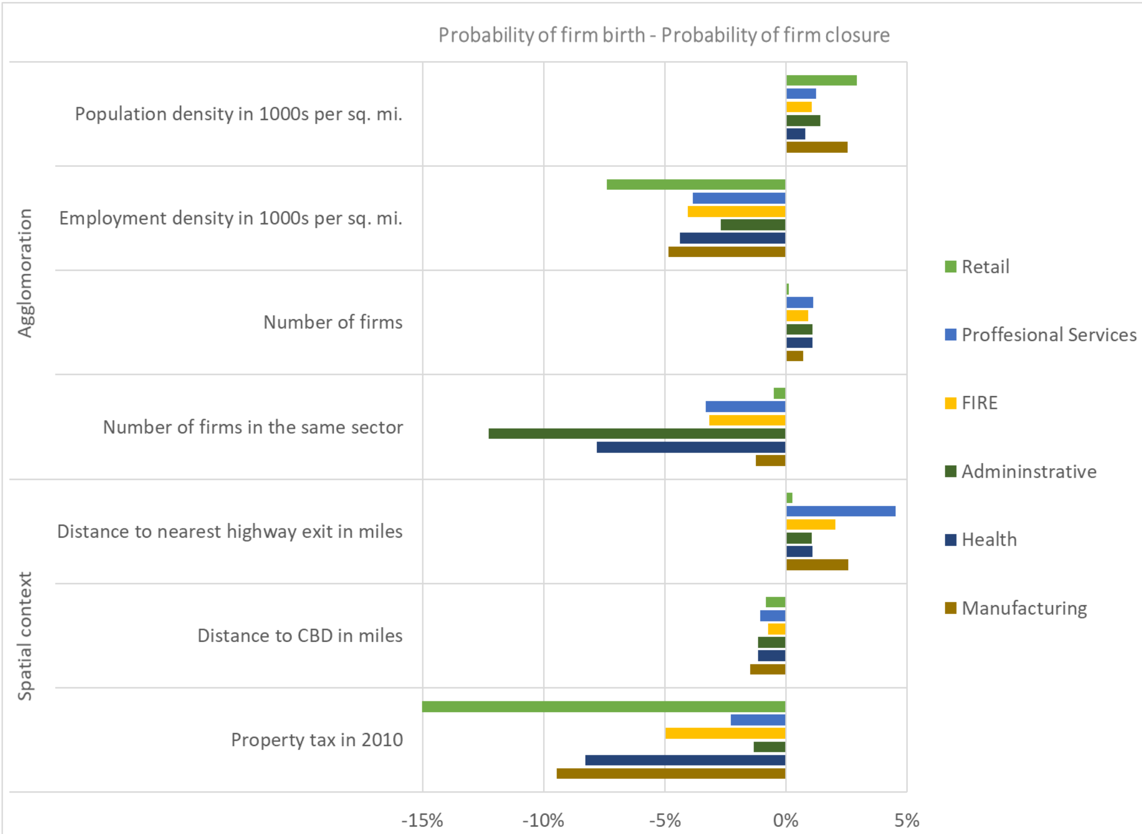
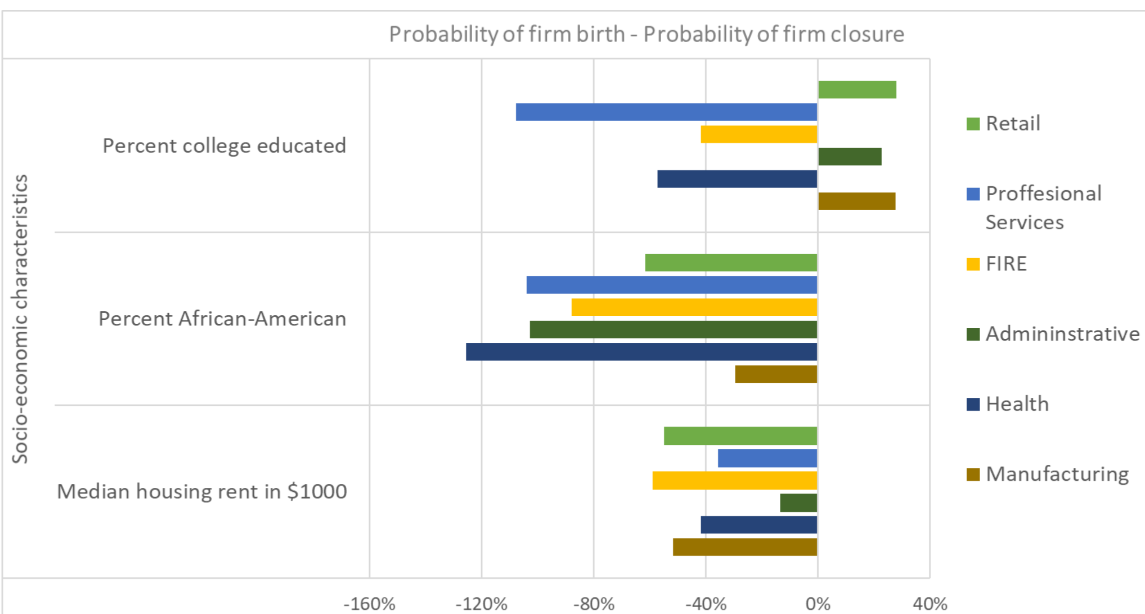


Figure 17. The difference between the predicted probability of firm birth and firm closure for socio-economic variables by industry sector



In addition to proximity to rail station, the firm closure and firm birth regression results suggest that population density, number of existing firms, distance to highway, and the level of education are important factors influencing the probability of firm retainment (see Figure 16 and Figure 17). On the other hand, variables such as the number of firms in the same industry category tends to negatively predict the probability of firm retainment (ranging between -1% to -12%, as shown in Figure 16), which suggests that the competition effect of firms in the same sector overrides localized agglomeration economies.

4.3. Chapter summary

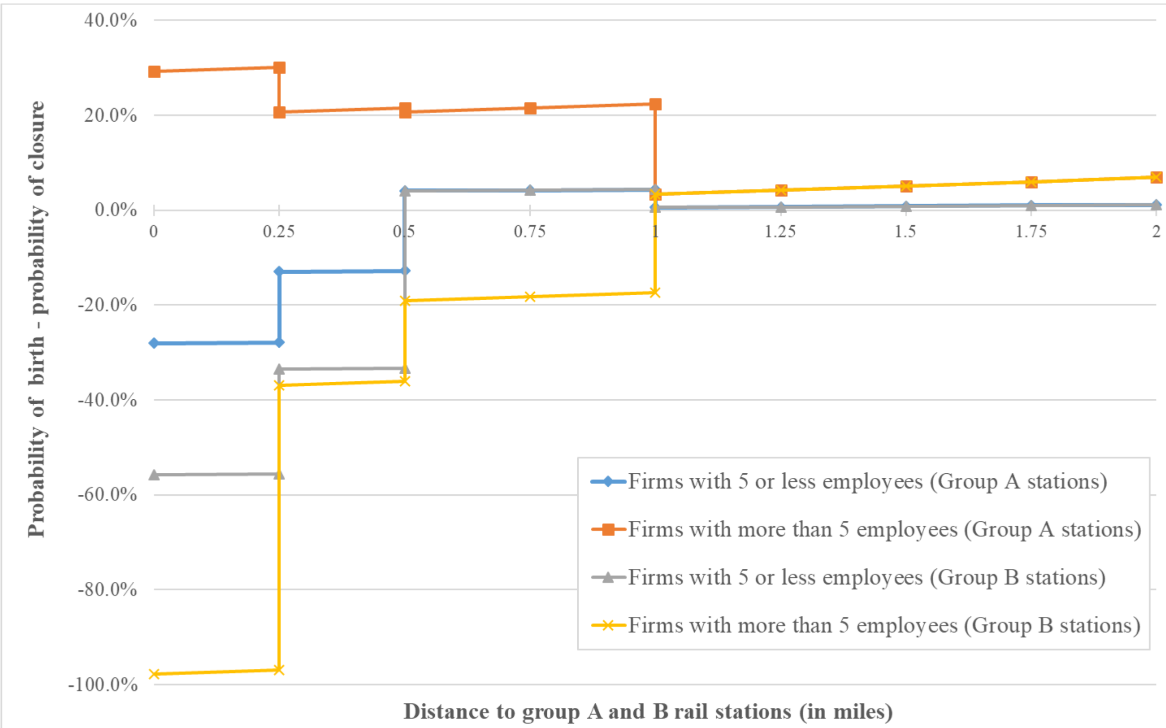
The results in this chapter showed mixed relationships between areas near the passenger rail stations and the probability of firm birth and closure. Most importantly, the results showed that lengthy periods of time elapse before areas near the rail stations exhibit higher probabilities of firm birth than probabilities of firm closure (i.e. positive probability of firm retainment). That is, areas within a mile of the mature rail stations (group A) were more likely to retain firms than areas within a mile of recently opened stations (group B and C). Positive firm retainment in an area indicate an increase in firm density. Figure 18 summarizes the predicted probabilities of firm retainment by distance from group A and group B stations.[1] Evidently, areas within a short walking distance from the mature rail stations exhibits positive probabilities of firm retainment,

[1] See Appendix C for a summary figure of the predicted effects of firm retainment by distance from group C rail stations (stations opened after 2000).

specifically for larger firms with more than five employees, compared to areas further away from the stations.

On the other hand, areas near rail stations opened after 1990 (group B and C stations) exhibit negative probability of firm retainment compared to areas further away from the stations (see Figure 18). In Figure 18, the upward slope of the plotted line for firms with more than five employees indicates that the likelihood of firm retainment increased between 1990 and 2010 in areas that are further in distance from the rail stations.

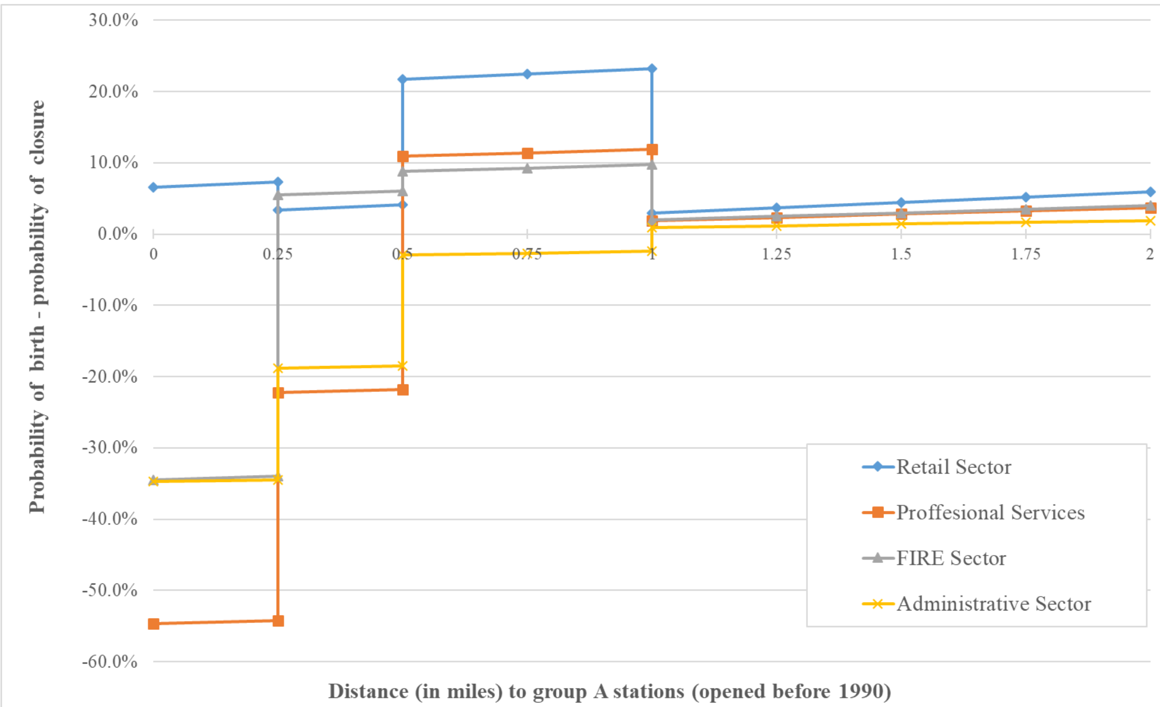
Figure 18. Probability of firm retainment of station distance variables by firm size



The y-axis shows the predicted probabilities of firm retainment as calculated from the coefficient values of firm birth and firm closure. The lines plotted in the graph are calculated as Y = Birth[(e(βi) – 1) + (distance from station * (e(α)-1))] - Closure[(e(βi) – 1) + (distance from station * (e(α)-1))] where βi is the coefficient for the dummy variable of the respective station buffer and α is the coefficient for the continuous distance-to-station variable.

Similar to the analysis by firm size, there are mixed relationships between areas near the rail stations and the probability of firm retainment across industry sectors. Figure 19 summarizes the results of four dominant industry sectors within the study area by distance from the mature rail stations (group A). It shows that firms in the retail trade sector are the most likely to benefit from areas within up to a mile distance of the mature rail stations. Areas within a mile of Group B and C show no signs of positive firm retainment compared to control areas.[1]

Figure 19. Probability of firm retainment of station distance variables by selected industry sectors



The y-axis shows the predicted probabilities of firm retainment as calculated from the coefficient values of firm birth and firm closure. The lines plotted in the graph are calculated as Y = Birth[(e(βi) – 1) + (distance from station * (e(α)-1))] - Closure[(e(βi) – 1) + (distance from station * (e(α)-1))] where βi is the coefficient for the dummy variable of the respective station buffer and α is the coefficient for the continuous distance-to-station variable.

---

[1] See Appendix C for a summary figure of the predicted effects of firm retainment by industry sector and distance from group B and C rail stations.

These results overall show that rail stations have not consistently boosted firm retainment nearby, except in the case of areas near the mature rail stations that were opened before 1990. The inconsistency in firm retainment near rail stations raises the question of what policymakers should do differently to encourage transit-oriented development. Evidently, areas near the stations do not experience an increase in firm density, at least in the short run, without proper urban-growth policies. For more immediate results, policymakers advocating for transit-oriented development should be more proactive in focusing development around transit stations, by adopting policies such as urban growth boundary and maximum parking caps to impede urban sprawl and promote mixed-use and transit-oriented form of urban development.

CHAPTER 5: ANALYSIS RESULTS: FIRM RELOCATION

Firm relocation is one of the least examined patterns of firm dynamics in literature, as discussed in Chapter 2. Studies examining the association between rail transit and firm relocation are particularly rare. The scarcity of data that tracks the exact origins and destinations of relocating firms is the main reason for the lack of empirical research on firm relocation. Using the NETS data, this chapter examines the relocation patterns in relation to proximity to the passenger rail stations within the state of Maryland. This relocation analysis is possible because the NETS dataset provides the coordinates of the origins and destinations of the relocating firms (see Appendix A).

In the period between 1990 and 2009, one in ten firms within the State of Maryland relocated at least once during its lifespan. Table 25 provides summary statistics on the total number of firm relocations between 1990 and 2009 within the State of Maryland and the study area (i.e. the five jurisdictions within Maryland). Firm relocations in Table 25 are categorized by the origin and destination regions. The origin and destination regions in Table 25 are: (a) the study area, (b) the rest of the State of Maryland, and (c) areas outside the State of Maryland. Nearly half of firm relocations have occurred within the study area between 1990 and 2009. That is, nearly half of these relocations have both origins and destinations within the study area. During this same period, 12.3% of all firm relocations in Maryland were relocations to the study area either from the rest of Maryland (3.9%) or from outside of Maryland (8.4%). However, a higher percentage (15.8%) of firms have relocated from the study area to locations outside the study area's five jurisdictions, either in the rest of the State of Maryland (7.2%) or areas outside of Maryland (8.6%).

Table 25. Number of firm relocations between 1990 and 2009 by regions of origin and destination

Origin (O) and destination (D) of firm relocation	Number of relocations	Percentage
O and D within Study Area*	45,919	49.3%
O and D within rest of MD	15,275	16.4%
Study Area to out of MD*	8,031	8.6%
Out of MD to Study Area*	7,796	8.4%
Study Area to rest of MD*	6,666	7.2%
Rest of MD to Study Area*	3,605	3.9%
Rest of MD to out of MD	3,135	3.4%
Out of MD to rest of MD	2,691	2.9%
Total	**93,118**	100%

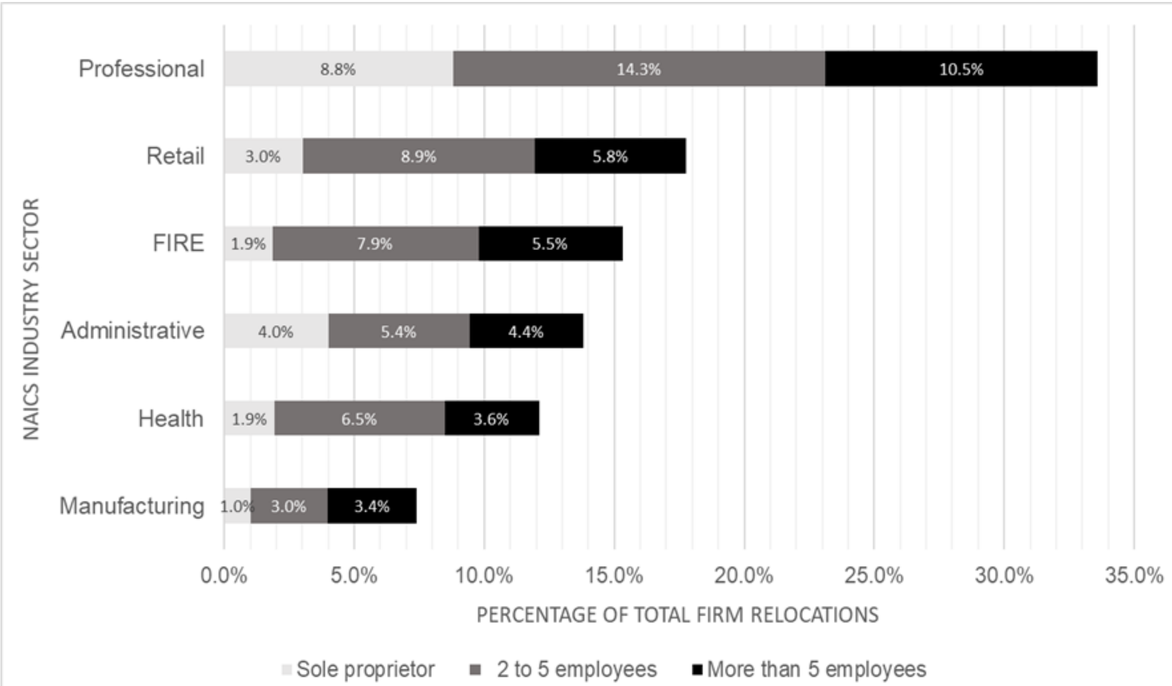
Note: The asterisks (*) indicate firm relocations that have their origion and/or destination within the study area, which are considered for the ananlysis (adding up to 72,017 relocations).

Figure 20 shows the percentage of total firm relocations within the study area across three firm-size categories (sole proprietor, two to five employees, and more than five employees) and six selected industry sectors (Professional service, Retail, FIRE, administrative, health, and manufacturing). The majority of relocating firms within the study area are small in terms of the number of employees. Firms with five or fewer employees accounted for more than 66% of the 72,000 relocating firms that have their origin and/or destination within the study area in the period between 1990 and 2009. In addition, three out of ten relocations were by sole proprietors.

Regarding firm relocation across industry sectors, firms in the professional service sector are about twice as likely to relocate (33.6% relocated), compared to firms in the other five industry sectors (second to professional services is retail firms with

17.7% relocated). Firms belonging to the manufacturing sector on the other hand are the least likely to relocate (only 7.4% relocated).

Figure 20. Percentage of total firm relocations within the study area by size and industry sector (1991-2009).



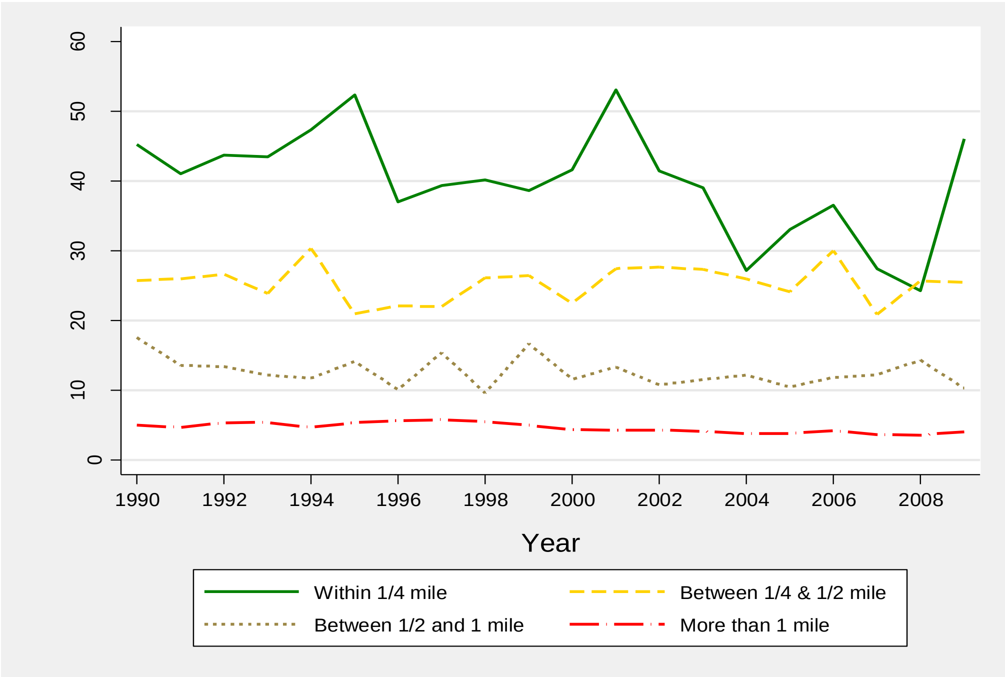
Note: FIRE industry sector refers to firms in the finance, insurance, and real estate industry.

In the study area, the Census blocks that are within a short walking distance of the passenger rail stations have, on average, attracted more relocating firms (i.e. inward firm relocations) compared to control Census blocks located more than one mile away from the stations.[1] In every year between 1990 and 2009, the number of inward firm relocations per square mile (i.e. density of inward relocation) was much higher within the *quarter mile* buffer of rail stations than in the *quarter to half mile* buffer or the *half to one mile* buffer, and higher still than the density of inward firm relocations outside the one

[1] The firm relocations that have their origin and destination within the same Census block are excluded from the analysis.

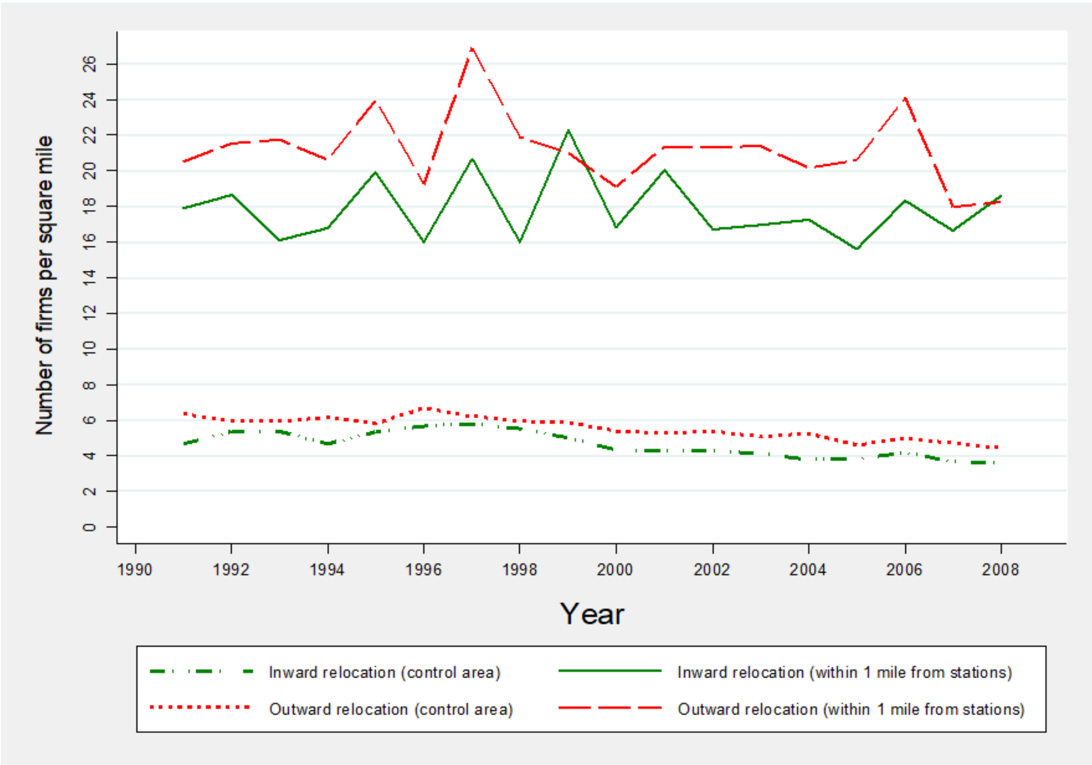
mile threshold from the rail stations (see Figure 21). In any given area, there are two sides to firm relocation, existing firms can be pushed to relocate elsewhere (outward relocation) and others can be attracted from elsewhere to relocated within (inward relocation). Reasons for inward and outward relocation decisions are discussed in Chapter 2.

Figure 21. Number of inward firm relocations per square mile within the study area by distance from rail stations (1990 to 2009).



High number of inward firm relocations in an area can be a push factor for less competitive firms, which may subsequently lead to high number of outward firm relocations. In every year between 1990 and 2009, Figure 22 compares inward and outward relocation densities of areas located within a mile of the passenger rail stations and areas located more than a mile from the stations. Considering both inward to outward firm relocations, the study area has experienced a negative net relocation in the period between 1990 and 2010 (i.e. the number of outward relocations exceeds the number of inward relocations, as shown in Figure 9), so did areas within a mile of the rail stations. Due to variations in push and pull factors (discussed in Chapter 2), inward and outward relocation densities are not homogeneous across Census blocks within the study area. Without a statistically controlled analysis, one cannot determine whether or not areas near rail stations have positive influence on net firm relocation.

Figure 22. Number of inward and outward firm relocations per square mile within the study area by distance from rail stations (1990 to 2009).



This chapter carries out a series of regression analysis to examine inward and outward firm relocations in areas within a short distance to the passenger rail stations, compared to control areas located more than a mile from the stations. At Census block level, the negative binomial method is appropriate for the analysis of firm relocation since a considerable number of blocks have zero number of firm relocations (see Chapter

3 for a detailed discussion on NB regression). This chapter is divided into four main sections. Each section caries out two regression methods, the standard NB and the PS-weighted NB regression. Section one and two examine inward firm relocation by firm size and industry sector, respectively. Section three and four examine outward firm relocation by size and sector, respectively, and compare results to inward firm relocation to determine the net predicted effects by calculating the difference between inward and outward relocation effects. The summary statistics of the outcome and control variables were presented earlier in Chapter 4 (see Table 13). This chapter ends with a section summarizing the results of inward and outward firm relocation relative to the distance from the passenger rail stations.

5.1. Inward firm relocation by size: regression results

The inward relocation analysis indicates that, overall, access to passenger rail stations is a pull factor for relocating firms. The analysis by firm size suggests that larger relocating firms are more likely to relocate within short proximity of mature rail stations (group A), whereas smaller relocating firms are more likely to locate within a close proximity to more recent stations (group B and C). Table 26 and Table 27 show the regression results of the inward relocation analysis using the standard NB and the PS-weighted NB models, respectively. The inward relocation models use control variables similar to the ones used in the firm birth and closure models (see Table 13). As discussed in Chapter 3, data for the control variables are obtained either at the Census block level or at the Census block group level for three periods, 1990, 2000, and 2010. At the Census block level, variables include distance to the nearest rail station, distance to the nearest highway exit, distance to the nearest central business district (either in Washington, DC or Baltimore City), transit-to-auto accessibility ratio, the number of firms, the number of firm births and closures, and the number of outward firm relocations. At the Census block group level, the control variables include the socio-economic characteristics of the local population, such as population and employment densities, income, and education. The negative coefficients of *distance to rail station* variable in Table 26 suggest that the likelihood of inward firm relocation decreases the greater the distance from stations (the coefficients range between -0.050 and -0.052 across the four size categories).

The inward relocation results substantially vary across the three station buffers (i.e. within a *quarter mile* buffer, within a *quarter to half mile* buffer, and within a *half to one mile* buffer). Starting with the mature rail stations opened before 1990 (group A), the regression coefficients suggest a positive association between areas within a *quarter mile* of the rail stations and the probability of inward relocation (e.g. the estimated coefficient is 0.217 for firms with five or less employees and 0.617 for firms larger than five employees). In other words, relocating firms are more likely to choose areas within a *quarter mile* distance of group A rail stations as their new firm location than areas located more than a mile of the stations, all else held equal. On the contrary, areas located within a *half to one mile* distance of group A rail stations experienced negative probabilities of inward firm relocations compared to control areas (e.g. the estimated coefficient is -0.310 for firms with five or less employees and -0.147 but just below the 95% statistical significance for firms larger than five employees, as shown in Table 26). The areas within the *half to one mile* buffer of group A stations have positive firm retainment (as shown in Chapter 4), which may explain the negative probabilities of inward relocation.

Table 26. The number of inward firm relocation by firm-size as a function of proximity to rail stations, agglomeration, and socio-economic characteristics. Using the standard negative binomial method.

Dependent Variable: Inward Relocations	(1) All Firms	(2) Sole Proprietor	(3) Firms > 1 employee	(4) Firms <=5 employees	(5) Firms >5 employees
Distance to Rail station (in mi)	-0.055***	-0.050***	-0.052***	-0.051***	-0.052***
Group A stations: within 1/4 mile	0.282**	0.182	0.346**	0.217*	0.617***
Group A stations: within 1/4 to 1/2 mi	0.021	0.079	0.026	0.002	0.229*
Group A stations: within 1/2 to 1 mi	-0.250***	-0.148*	-0.274***	-0.310***	-0.147
Group B stations: within 1/4 mile	0.439***	0.597***	0.453***	0.499***	0.531***
Group B stations: within 1/4 to 1/2 mi	0.239**	0.370***	0.186*	0.201*	0.309**
Group B stations: within 1/2 to 1 mi	0.065	-0.066	0.086	-0.010	0.162*
Group C stations: within 1/4 mile	1.241**	1.808***	0.994*	1.503**	0.715
Group C stations: within 1/4 to 1/2 mi	0.095	0.759	-0.231	0.088	-0.09
Group C stations: within 1/2 to 1 mi	0.096	0.043	0.044	0.070	0.156
Accessibility ratio	0.738***	0.376	0.827***	0.486**	1.029***
Population per sq. mi. (in 1000s)	-0.047***	-0.039***	-0.048***	-0.041***	-0.076***
Employee per sq. mi. (in 1000s)	0.059***	0.095***	0.043***	0.072***	0.040*
Number of firms	0.014***	0.016***	0.015***	0.016***	0.018***
Firm births	0.003***	0.014***	0.002*	0.005***	0.001
Firm closures	-0.001*	0.008***	-0.001**	0.001**	-0.001
Firm outward relocations	-0.035***	-0.090***	-0.032***	-0.054***	-0.037***
Median HH Income (in $1000s)	0.004***	0.007***	0.002***	0.005***	-0.003***
Unemployment rate	1.244***	2.782***	0.599*	1.570***	-0.011
Percent college educated	0.533***	0.824***	0.527***	0.717***	0.343**
Percent African-American	0.204***	0.637***	0.064	0.342***	-0.141*
Median housing rent (in $1000s)	0.145***	0.286***	0.075**	0.186***	-0.033
Distance to highway (in mi)	-0.072***	-0.021	-0.089***	-0.048***	-0.141***
Distance to CBD (in mi)	0.047***	0.062***	0.042***	0.048***	0.042***
Property tax (in $1000)	-0.030	-0.043	-0.045	-0.018	-0.106**
Constant	-0.865***	-3.024***	-0.445***	-1.499***	-0.298*
ln_r					
Constant	1.318***	2.373***	1.367***	1.664***	1.377***
ln_s					
Constant	-0.957***	0.041	-1.182***	-0.609***	-1.405***
N. of cases	116820	116820	116820	116820	116820
Log Likelihood	-44143.887	-17552.599	-36228.111	-35721.896	-19449.478
chi2	6410.161	4732.732	4939.558	6224.829	2831.312

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

Table 27. The number of inward firm relocation by firm-size as a function of proximity to rail stations, agglomeration, and socio-economic characteristics. Using the PS-weighted negative binomial method.

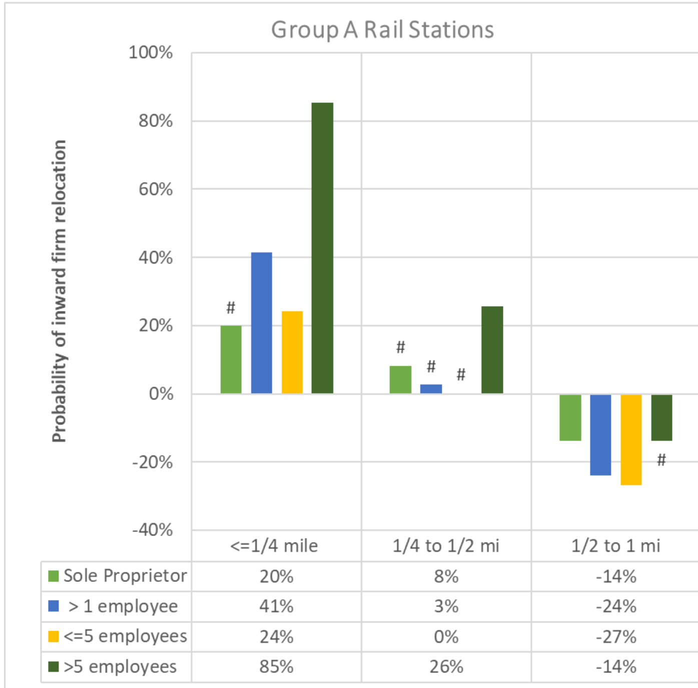
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Inward Relocations	**All Firms**	**Sole Proprietor**	**Firms > 1 employee**	**Firms <=5 employees**	**Firms >5 employees**
Distance to Rail station (in mi)	-0.043***	-0.042***	-0.044***	-0.037***	-0.057***
Group B stations: within 1/4 mile	0.269*	0.506**	0.242	0.287*	0.287
Group B stations: within 1/4 to 1/2 mi	0.238*	0.319*	0.215	0.193	0.289
Group B stations: within 1/2 to 1 mi	0.062	0.101	0.048	0.036	0.106
Accessibility ratio	0.267	0.205	0.257	0.281	0.257
Population per sq. mi. (in 1000s)	-0.044***	-0.030**	-0.047***	-0.038***	-0.054*
Employee per sq. mi. (in 1000s)	0.046***	0.076***	0.029	0.063***	-0.028
Number of firms	0.047***	0.020***	0.049***	0.040***	0.046***
Firm births	0.032***	0.037***	0.022***	0.039***	0.011**
Firm closures	-0.030***	0.002	-0.032***	-0.020***	-0.030***
Firm outward relocations	-0.115***	-0.103***	-0.101***	-0.121***	-0.080***
Median HH Income (in $1000s)	0.004***	0.008***	0.002***	0.006***	-0.002*
Unemployment rate	0.832**	2.873***	-0.035	1.367***	-1.04
Percent college educated	0.069	0.378**	0.042	0.275**	-0.219
Percent African-American	0.078	0.579***	-0.033	0.205***	-0.137
Median housing rent (in $1000s)	0.117***	0.224***	0.076**	0.139***	0.043
Distance to highway (in mi)	-0.066***	-0.01	-0.093***	-0.031**	-0.183***
Distance to CBD (in mi)	0.036***	0.051***	0.034***	0.036***	0.043***
Property tax (in $1000)	0.086**	-0.118*	0.122***	0.055	0.097
Constant	-2.941***	-5.080***	-2.860***	-3.666***	-2.917***
chi2	9400.37	11010.097	7059.409	16432.487	5180.34
N. of cases	101859	101859	101859	101859	101859

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

The results of the station buffers in Table 26 suggest that areas within the *quarter mile* buffer of the stations exhibit positive likelihood of inward firm relocation across all rail stations regardless of their level of maturity. Similar to group A stations, areas within a *quarter mile* of rail stations in group B (opened between 1990 and 1999) and group C (opened after 2000) show positive influence on inward firm relocation compared to control areas, all else held equal. For instance, the estimated coefficients for the *quarter mile* buffer of group B stations is 0.499 for firms with five or less employees and 0.531 for firms larger than five employees. The estimated coefficients of the station buffers are better understood when converted to predicted probability of effect, using the equation $[e(\beta) - 1]$.

Figure 23 presents the predicted probabilities of inward firm relocation within the three station buffers of the mature rail stations (group A), using the estimated coefficients in Table 26. In areas within a half mile distance of group A stations, larger firms with more than five employees have higher probability of inward relocation than smaller firms with less than five employees, compared to control areas (e.g. the probability is 85% for firms larger than five employees but only 24% for firms with five or less employees within the *quarter mile* buffer). These results suggest that labor intensive firms (larger firms) that decided to relocate rank improved access to the labor force provided by rail transit much higher in their relocation decisions than less labor intensive firms (smaller firms).

Figure 23. The predicted effects of inward firm relocation by firm size by distance from the mature rail stations (group A stations).



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	20%	8%	-14%
> 1 employee	41%	3%	-24%
<=5 employees	24%	0%	-27%
>5 employees	85%	26%	-14%

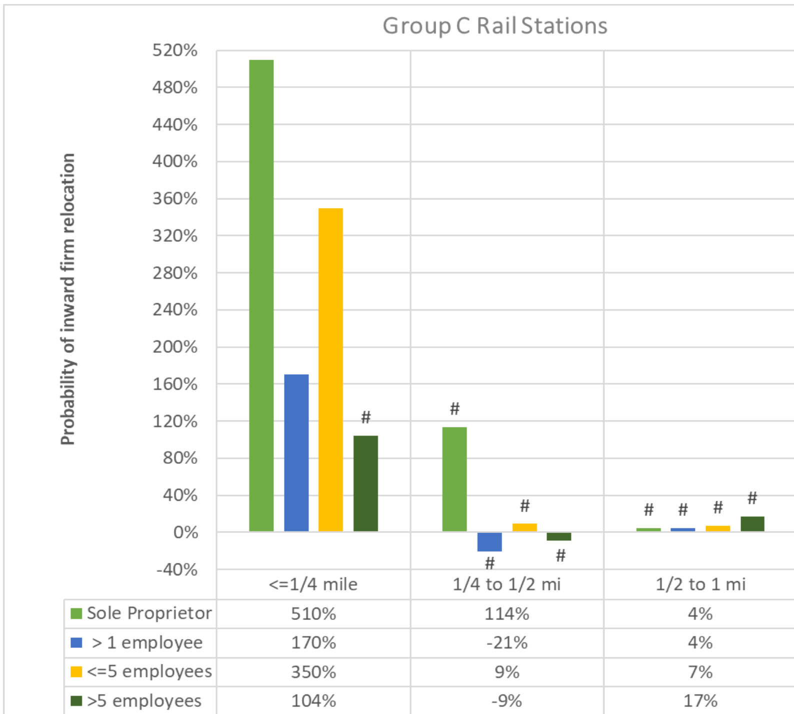
Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using ($e^{(\beta i)} - 1$), where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened before 1990. The symbol (#) refers to statistically insignificant values.

On the other hand, in areas within the *quarter mile* buffer of the more recently opened stations (group C), smaller firms with five or less employees have higher probability of inward relocation (350%) than larger firms with more than five employees (104% but statistically insignificant), as shown in Figure 24. One must not jump to conclusions, however, regarding these results because group C stations consist of six rail stations only, and any results associated with these stations are from one period only (2010), as they were opened after 2000.[1] In addition, most of the estimated coefficients of group C station buffers are statistically insignificant, the only exception being smaller

[1] Census blocks within up to a mile of group B stations are considered treated in 2010 only because in 1990 and 2000 those blocks were not served by any rail stations.

firms with five or less employees within the *quarter mile* buffer, as explained earlier (see Figure 24).

Figure 24. The predicted effects of inward firm relocation by distance from the more recently opened stations (group C stations).

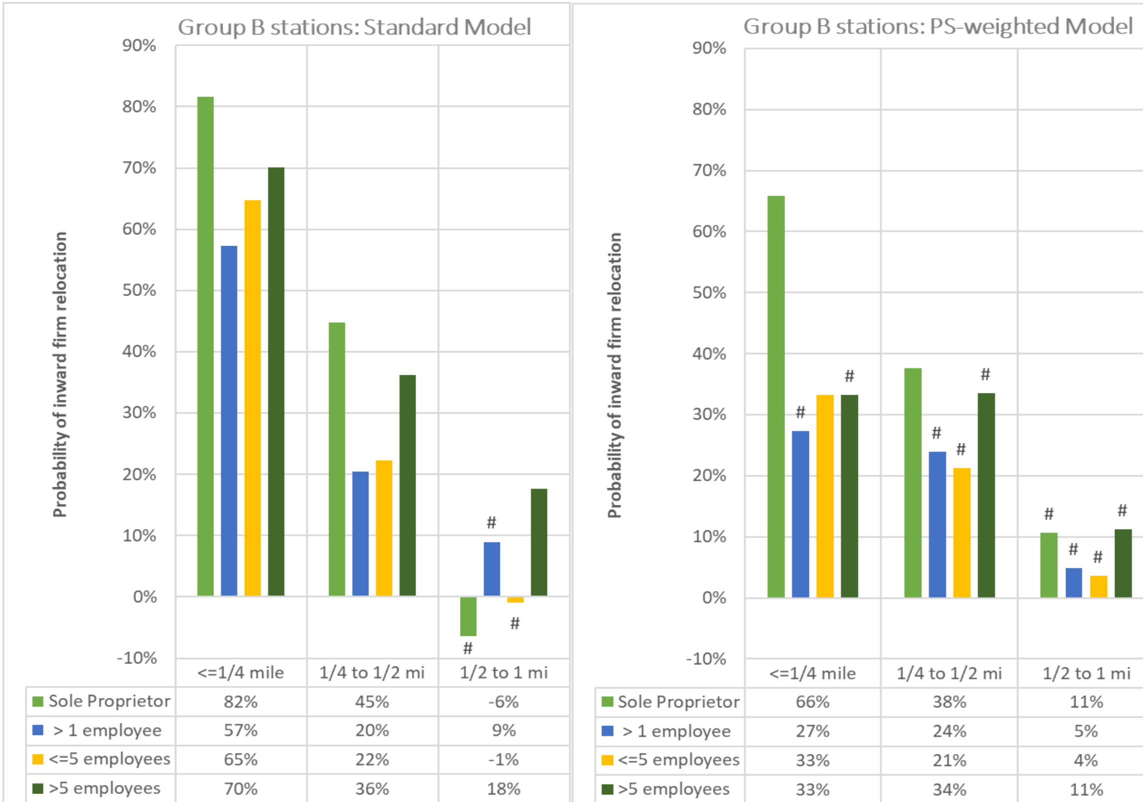


Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using ($e(\beta i) - 1$), where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group C stations are those opened between 2000 and 2004. The symbol (#) refers to the statistically insignificant values.

Turning to group B rail stations, the probability of inward firm relocation of the three station buffers are estimated using the standard NB binomial and the PS-weighted NB methods (Table 26 and Table 27). The PS-weighted NB method restricts the analysis of firm dynamics to group B rail stations because these stations were opened after 1990. It controls for possible endogeneity of rail station placement and firm location decisions, as discussed in Chapter 3. The estimated coefficients across both NB methods are mostly consistent in sign (direction of influence) but inconsistent in statistical power (i.e. level of significance). For instance, in the standard NB method, the estimated coefficient of firms large than five employees is 0.531 in the *quarter mile* buffer of group B stations (Table 26), whereas in the PS-weighted NB method, the same estimated coefficient is 0.287 but below the 95% statistical significance (Table 27).

Figure 25 converts the estimated coefficients obtained from both NB methods to the probability of inward relocation for each of the three station buffers of group B stations, compared to control areas. Both methods indicate a positive and statistically significant probability of smaller relocating firms to locate in areas within a *quarter mile* and a *quarter to half mile* of group B stations, all else held equal. For instance, based on the standard NB method, sole proprietors are 82% more likely to relocate within the *quarter mile* buffer, whereas the probability is 66% based on the PS-weighted method. For firms smaller than five employees, the probability of inward relocation is 65% and 35% across the two methods, both statistically significant (see Figure 25). For larger relocating firms with more than five employees, the probability of inward relocation is only significant for the *quarter mile* buffer (70%) and the *quarter to half mile* buffer (36%) of group B stations, using the standard NB method. The fact that smaller relocating firms are more likely to locate near group B stations than larger relocating firms suggests that access to rail stations is valuable to firms not only in terms of better access to labor force but in terms of access to a wider customer base as well.

Figure 25. The predicted effect of inward firm relocation by distance from rail station in group B, comparing results of two regression methods: the standard NB (left side) and the PS-weighted NB (right side)



Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using ($e(\beta i) - 1$), where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B stations are those opened between 1990 and 1999. The symbol (#) refers to the statistically insignificant values.

Focusing on other transportation-related variables, the coefficients of transit-to-auto accessibility-ratio in Table 26 suggest that better transit access is a key factor influencing the locations selected by relocating firms, especially for larger firms that benefit from better access to a labor force which would be provided by transit service. For instance, the estimated coefficient of the accessibility-ratio is 1.029 for firms larger than five employees, whereas the coefficient is 0.486 for firms with five or less employees. The inward relocation analysis also suggests that larger relocating firms consider proximity to a highway exit a more valuable factor in location decisions than smaller firms. For instance, the estimated coefficient of distance to highway exist in Table 26 is -0.141 for firms larger than five employees, whereas the coefficient is -0.048 for firms with five or less employees. Evidently, the negative coefficients of the continuous variable of the distance to rail station suggests that the closer a block is to a rail station the higher its likelihood of inward relocation (e.g. the coefficient is -0.052 for firms larger than five employees and -0.051 for firms with five or less employees).

There is only one agglomeration-related variable that has an unexpected direction of influence on firm relocation (Chapter 2 provides a discussion on the direction of impact expected for each control variable on firm relocation). The results in Table 26 indicate that population density has a negative and statistically significant effect on inward firm relocation (e.g. the coefficient is -0.076 for relocating firms larger than five employees and -0.041 for relocating firms with five or less employees). The negative effect of population density on inward firm location decisions was also found in the firm birth analysis presented earlier in Chapter 4; this effect can be explained by the highly suburban urban form of the study area. There are numerous suburban Census block groups within the study area that are predominantly residential and densely populated. Moreover, the outward firm relocation analysis, presented later in this chapter, provides more insight on the impact of population density, along with other control variables, on net firm relocation by comparing the regression results of outward relocation to inward relocation.

Employment density, on the other hand, has positive influence on inward firm relocation. The analysis by firm size shows that the coefficients of employment density variables are higher for smaller firms, suggesting that employment density is a more important factor for the location choices of small firms that made the decision to relocate, all else held equal. For instance, the estimated coefficient of employment density for relocating firms with more than five employees is 0.072, while the coefficient is 0.040 for relocating firms with five or less employees (see Table 26). The employment density results indicate that smaller relocating firms benefit more from local agglomeration (through externalities such as information spillovers) than larger relocating firms.

Similar to the findings of employment density, the regression results in Table 26 suggest that the total number of firms within a Census block is also a significant factor influencing inward firm relocation. For instance, the estimated coefficient for firms larger than five employees is 0.018, suggesting that the probability of a larger relocating firm to locate within a Census block is 1.8% higher for each additional existing firm, everything else held equal (the probability is 1.6% for smaller relocating). These results suggest that smaller and larger relocating firms value local agglomeration of firms when moving to a new location for the economic benefits that agglomeration provides, such as local information sharing (See Chapter 2 for detailed discussion of agglomeration externalities).

The number of firm births within a Census block has a positive but small association with the probability of inward firm relocation. For instance, the estimated coefficient is 0.005 for firms with five or less employees; whereas the coefficient for firms with more than five employees is 0.001 but also statistically insignificant (see Table 26). The positive association between firm birth and the likelihood of inward relocation is not surprising because higher number of firm births in an area implies that it fosters business development. The negative coefficients of firm closure (-0.001) and outward relocation (-0.032) for larger firms indicate that the higher the number of firm closures and/or outward relocations within a block, the lower the rate of inward relocation. The results suggest that areas experiencing high number of firm closures and/or outward relocations are not an attractive destination for firms that decide to move to a new location. An exception to this trend is the positive association between firm closure and inward relocation of smaller firms with sole proprietor (0.008) and with five or less employees (0.001), as shown in Table 26. This positive association suggests that higher numbers of closures in an area may imply high competitiveness, which can be an attractive factor for certain relocating firms. These associations can be better understood in the inward relocation analysis by industry sector, which is the subject of the following section.

Focusing on the socio-economic variables, the association is positive but small (0.005) between median household income and inward firm relocation for smaller firms with five or less employees. This association is negative but small (-0.003) for larger firms with more than five employees, however (see Table 26). These opposite directions of association imply that larger relocating firms are more likely to be drawn to areas with lower costs of labor (inferred by low household income), whereas smaller relocating firms are more likely to be drawn to more affluent neighborhoods with higher demand for goods and services. Median housing rent has similar direction of association with the probability of inward relocation to that found for household income. For instance, the estimated coefficient for median housing rent is 0.186 for smaller relocating firms with five or less employees, whereas the coefficient is -0.033 (but below the 95% statistical significance) for larger relocating firms with more than five employees (see Table 26). The positive association between smaller relocating firms and housing rent confirms the earlier conclusion that smaller relocating firms are more likely to locate in more affluent neighborhoods.

Human capital is also a key contributing factor to the location decisions of relocating firms. This is suggested by the positive association between the percentage of population that is college educated and inward firm relocation. The estimated coefficient for relocating firms with five or less employees is 0.717, whereas the coefficient for larger relocating firms with more than five employees is 0.343 (see Table 26). Regarding race, the association is positive (0.342) between the percentage of African American population and inward firm relocation of smaller firms with five or less employees. This association is negative (-0.141), however, for larger relocating firms with more than five employees. Large relocating firms with more than five employees tend to avoid areas with high percentages of African Americans, which suggest a possibility that labor-intensive relocating firms have racial preferences.

Unemployment rate is a positive factor influencing inward firm relocation for smaller firms with less than five employees (β=1.570), but statistically insignificant factor for larger firms with more than five employees (β=-0.011). High unemployment rate in an area can be an indicator of dampening wages which may attract some relocating firms. For larger relocating firms, however, unemployment rate is not a contributing factor to their relocation decisions. A high unemployment rate may also suggest that an area is suffering from poverty and low quality of life, which can be repulsive to certain relocating firms that seek more affluent neighborhoods with skilled labor.

The association between local property tax and inward relocation for larger firms with more than five employees is negative (-0.106), suggesting that large moving firms are more likely to relocate to areas that impose lower property taxes. This negative association is expected since larger moving firms are normally deterred from areas with higher tax burden. For smaller moving firms with five or less employees, local property tax is not a significant factor in their decision to relocate (suggested by the statistically insignificant coefficient in Table 26, -0.018).

Similar to the earlier findings from the firm birth analysis, the continuous variable of the distance to the nearest central business district (i.e. the CBD of either Washington, D.C. or Baltimore City) is positively associated with the probability of inward firm relocation. For relocating firms with more than five employees, the estimated coefficient of the distance to CBD variable is 0.048, whereas the coefficient is 0.042 for moving firms with five or less employees (see Table 26). A positive association means that the farther the distance from the CBD, the higher the likelihood of inward firm relocation. These positive associations can be due to the highly suburban form of the study area and/or the normally high business competitiveness in CBD areas. The study area jurisdictions had not enforced any urban growth policies to concentrate urban development in urban clusters, which consequently drove employment and population to more affordable suburban areas.

5.2. Inward firm relocation by industry sector: regression results

This section examines inward firm relocation by selected industry sector to provide insights on how moving firms from various sectors react to proximity to rail stations and other factors such as agglomeration and socio-economic characteristics. The results overall suggest that moving firms in the professional services, FIRE, and administrative sectors are the most likely to relocate to areas within a short proximity to rail stations. Table 28 and Table 29 report the estimated coefficients of the standard NB and the PS-weighted NB methods, respectively. First, the negative estimated coefficients of the continues distance to station variable suggest that the number of inward firm relocations decrease the farther away the distance from stations, all else held equal (the coefficients range between -0.039 for moving firms in manufacturing sector to -0.076 for moving firms in health sector, as shown in Table 28).

For mature rail stations (group A), the results show evidence that moving firms in the FIRE, professional services, and administrative sectors value areas within short walking distance to the rail stations for the accessibility benefit these areas provide, such as access to a wider customer-base and labor force. This is specifically true for areas within a *quarter mile* of group A stations, where the estimated coefficients of inward firm relocation are positive for moving firms in the FIRE industry sector (0.654), the professional services sector (0.550) and the administrative sector (0.381). For other sectors, the coefficients are also positive in the *quarter mile* buffer of group A stations but lack statistical significance (e.g. the coefficient is 0.008 for retail sector and 0.102 for health sector).

Table 28. The number of inward firm relocations by industry sector as a function of proximity to rail stations, agglomeration, and socio-economic characteristics. Using the standard negative binomial method.

Dependent Variable: Inward Relocations	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail station (in mi)	-0.048***	-0.058***	-0.048***	-0.052***	-0.076***	-0.039***
Group A stations: within 1/4 mile	0.008	0.550***	0.654***	0.381*	0.102	0.279
Group A stations: within 1/4 to 1/2 mi	0.097	0.191	-0.021	0.233	-0.019	0.173
Group A stations: within 1/2 to 1 mi	-0.197	-0.331***	-0.338**	-0.334**	-0.187	-0.287**
Group B stations: within 1/4 mile	0.437*	0.851***	0.682***	0.580**	0.495*	0.446**
Group B stations: within 1/4 to 1/2 mi	-0.193	0.311*	0.175	0.101	0.334*	0.311*
Group B stations: within 1/2 to 1 mi	-0.019	0.139	0.074	-0.180	-0.120	0.045
Group C stations: within 1/4 mile	0.843	1.335*	1.240	1.356	-0.304	1.330
Group C stations: within 1/4 to 1/2 mi	-0.874	-0.183	-21.089	0.059	-1.498	0.528
Group C stations: within 1/2 to 1 mi	-0.441	-0.185	-0.202	0.157	0.294	0.461
Accessibility ratio	0.627*	0.737**	0.602*	0.628*	0.511	1.031***
Population per sq. mi. (in 1000s)	-0.031**	-0.059***	-0.066***	-0.058***	-0.014	-0.074***
Employee per sq. mi. (in 1000s)	0.026	0.084***	0.091***	0.101***	0.012	0.091***
Number of firms	0.020***	0.013***	0.013***	0.013***	0.005**	0.012***
Firms in the same industry sector	0.012**	0.030***	0.043***	0.169***	0.068***	0.154***
Firm births	0.003	0.011***	0.013***	0.005*	0.025***	-0.007***
Firm closures	0.002	-0.002*	0.002	-0.008***	-0.001	0.022***
Firm outward relocations	-0.052***	-0.052***	-0.063***	-0.051***	-0.036***	-0.085***
Median HH Income (in $1000s)	0.001	0.002*	-0.002	0.005***	0.003**	0.004***
Unemployment rate	-0.275	1.410**	-0.145	2.048***	1.070	2.832***
Percent college educated	0.121	1.725***	1.057***	0.255	1.290***	-0.070
Percent African-American	0.015	0.308***	0.216*	0.414***	0.245*	0.139*
Median housing rent (in $1000s)	-0.001	0.206***	0.091	0.126**	0.041	0.285***
Distance to highway (in mi)	-0.049*	-0.094***	-0.111***	-0.077***	-0.064*	-0.094***
Distance to CBD (in mi)	0.031***	0.068***	0.042***	0.055***	0.036***	0.038***
Property tax (in $1000)	-0.041	-0.069	-0.121*	-0.094	0.016	0.084
Constant	-1.157***	-2.697***	-2.460***	-2.283***	-2.543***	-1.912***
ln_r						
Constant	2.582***	1.913***	2.089***	2.903***	2.447***	2.081***
ln_s						
Constant	-0.645***	-0.564***	-0.246*	-0.18	-0.554***	-0.786***
N. of cases	116820	116820	116820	116820	116820	116820
Log Likelihood	-9649.764	-14904.579	-8306.392	-8828.466	-7099.46	-12911.172
chi2	2007.84	3854.607	2791.085	3260.668	2043.189	2873.293

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

Table 29. The number of inward firm relocations by industry sector as a function of proximity to rail stations, agglomeration, and socio-economic characteristics. Using the PS-weighted negative binomial method.

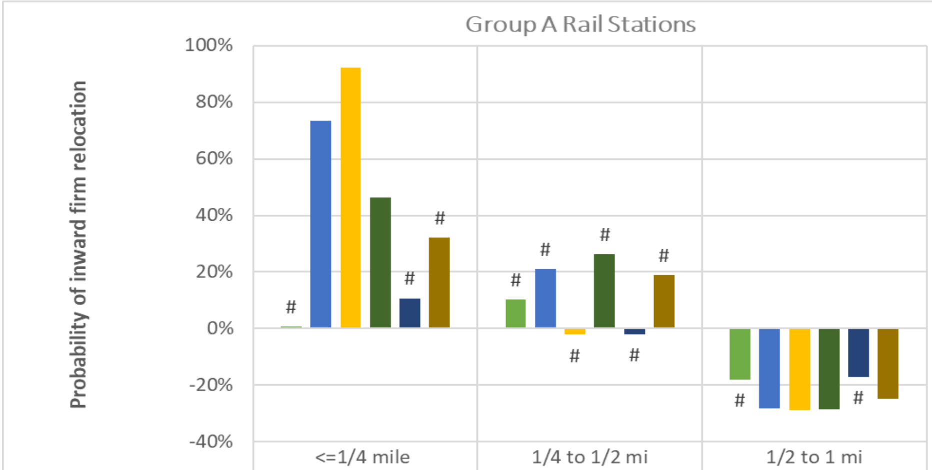
Dependent variable: Inward Relocations	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail station (in mi)	-0.040***	-0.053***	-0.038***	-0.059***	-0.078***	-0.047***
Group B stations: within 1/4 mile	0.149	0.523**	0.360	0.317	0.537	0.439
Group B stations: within 1/4 to 1/2 mi	-0.193	0.266	0.35	0.176	0.315	0.092
Group B stations: within 1/2 to 1 mi	-0.041	0.109	0.177	0.046	-0.058	-0.133
Accessibility ratio	0.203	0.485	0.288	-0.08	0.331	-0.154
Population per sq. mi. (in 1000s)	-0.037*	-0.032*	-0.01	-0.048**	-0.022	-0.108***
Employee per sq. mi. (in 1000s)	0.037	0.038	0.013	0.095***	0.035	0.073
Number of firms	0.022***	0.016***	0.017***	0.016***	0.010***	0.003
Firms in the same industry sector	0.035***	0.105***	0.119***	0.202***	0.088***	0.356***
Firm births	0.023***	0.026***	0.024***	0.009***	0.027***	0.017***
Firm closures	-0.015***	-0.023***	-0.016***	-0.019***	-0.007***	-0.002
Firm outward relocations	-0.063***	-0.065***	-0.076***	-0.048***	-0.047***	-0.053***
Median HH Income (in $1000s)	0.003**	0.003***	0.001	0.006***	0.004**	0.000
Unemployment rate	-0.083	0.703	-2.434**	2.124***	0.977	2.156***
Percent college educated	-0.281	1.124***	0.557*	-0.218	1.004***	-0.476
Percent African-American	-0.08	0.254**	0.308**	0.354***	0.297*	-0.259
Median housing rent (in $1000s)	0.007	0.211***	0.160*	0.118*	0.093	0.08
Distance to highway (in mi)	-0.047*	-0.092***	-0.123***	-0.069***	-0.060*	-0.157***
Distance to CBD (in mi)	0.021**	0.058***	0.045***	0.054***	0.031***	0.038***
Property tax (in $1000)	-0.024	-0.005	0.066	-0.033	0.033	0.223*
Constant	-4.288***	-5.000***	-5.169***	-5.301***	-5.504***	-5.032***
chi2	6549.037	10449.478	8088.87	8301.753	6053.484	5375.969
N. of cases	101859	101859	101859	101859	101859	101859

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

In areas within a *half to one mile* of group A stations, the association is negative with inward firm relocation, which is similar in direction to the association found previously in the firm birth analysis (the coefficients range between -0.187 and -0.338, as shown in Table 28). This negative association suggests that blocks within a *half to a mile* of the mature rail stations are either (1) less attractive to moving firms in their selections of new firm locations, (2) at a saturation point of urban density and therefore allowing for limited possibility for moving firms to relocate within, or (3) have land use and zoning regulations hindering moving firms from relocating within the block area (the analysis of outward firm relocation in the following sections provides more insight on the interaction between inward and outward relocation patterns within areas near rail stations).

The station buffer coefficients are converted to predicted probabilities to better understand their magnitude and direction of influence. Figure 26 shows the predicted probability of inward firm relocation within the three station buffers of group A stations, by industry sector. For instance, areas within a *quarter mile* of the mature rail stations have experienced 92% higher probability of inward relocation of moving firms belonging to FIRE sector, compared to control areas. Areas within a *quarter to half mile* distance of group A stations mostly indicate positive probabilities of inward relocation across the industry sectors but lack the statistical significance (the probabilities are -2% for FIRE and health sectors but statistically insignificant). One cannot infer any conclusions from the estimated coefficients of the *quarter to half mile* buffer of group A stations due to lack of statistical significance.

Figure 26. The predicted effects of inward firm relocation for selected industry sectors by distance from the mature rail stations (group A stations).



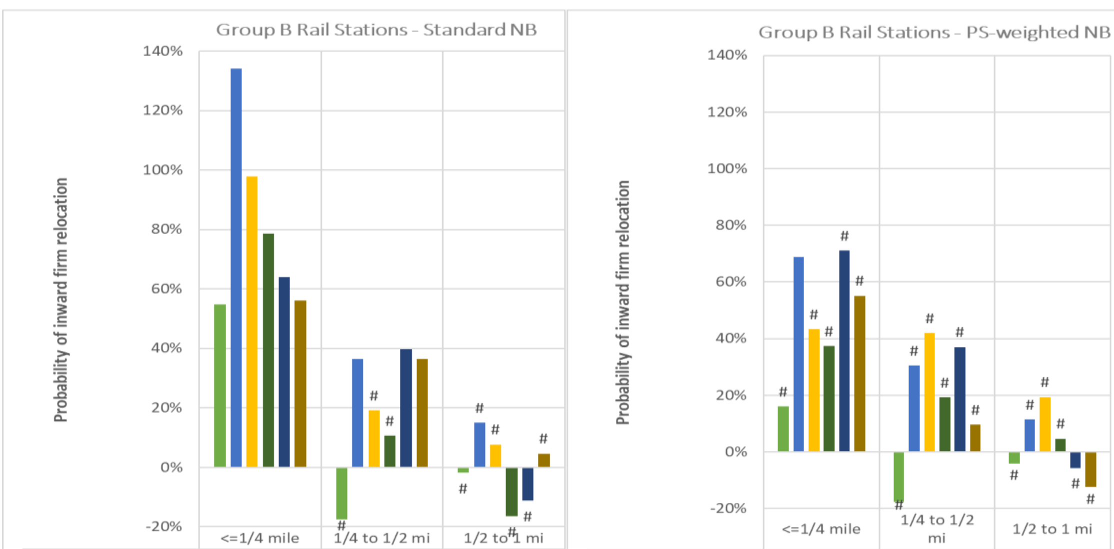
	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	1%	10%	-18%
Proffesional Services	73%	21%	-28%
FIRE	92%	-2%	-29%
Admininstrative	46%	26%	-28%
Health	11%	-2%	-17%
Manufacturing	32%	19%	-25%

Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using $(e^{(\beta i)} - 1)$, where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A stations are those opened between 1978 and 1989. The symbol (#) refers to the statistically insignificant values.

The results for Group B stations also provide evidence that close proximity to stations is a key factor in the location decisions of moving firms. Areas within a *quarter mile* of group B stations (stations opened between 1990 and 1999) show a positive association with inward firm relocation across all industry sectors. For instance, the highest estimated coefficient for the *quarter mile* buffer of group B stations is 0.851 for professional services, whereas the lowest coefficient is 0.437 for moving firms in retail sector. In the PS-weighted regression, the direction of influence remains positive in the *quarter mile* buffer of group B stations but the coefficients of five out of the six sectors lack the statistical significance (see Table 29). The only exception is moving firms belonging to professional services sector, where the coefficient is positive (0.523) and statistically significant.

The PS-weighted regression results provide a compelling evidence, that moving firms belonging to professional services sector strongly value areas located within a short walking distance to rail stations (PS-weighted regression controls for the endogeneity of the placement of group B stations). Figure 27 shows side by side the probabilities of inward relocation generated by the two regression methods for the three station buffers of group B stations by the six industry sectors. For instance, the predicted probability of inward relocation in the standard NB method is the highest for professional services sector (134%) within the *quarter mile* buffers of group B stations, whereas the same predicted probability is (69%) in the PS-weighted model. The standard NB estimates suggest that, as shown in Figure 27, blocks within a quarter mile of group B stations are more likely to experience inward firm relocations compared to blocks located more than a mile of the stations, all else held equal.

Figure 27. The predicted effect of inward firm relocation by distance from group B rail stations and selected industry sectors, comparing results of two regression methods: the standard NB (left side) and the PS-weighted NB (right side)



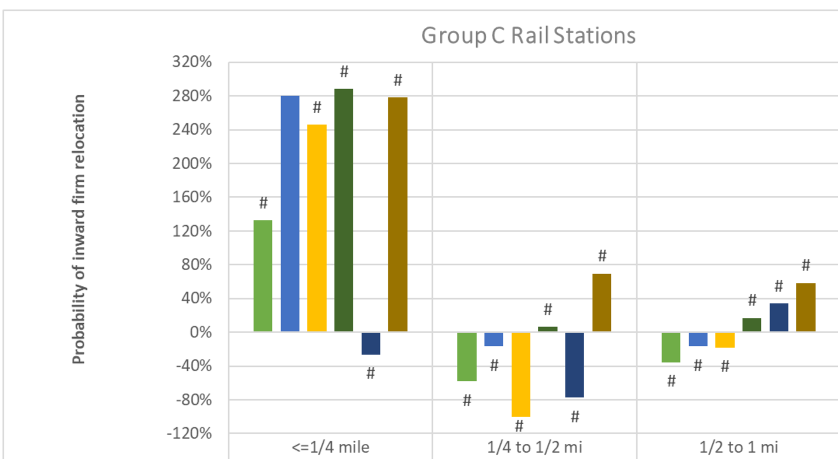
	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	55%	-18%	-2%
Proffesional Services	134%	36%	15%
FIRE	98%	19%	8%
Admininstrative	79%	11%	-16%
Health	64%	40%	-11%
Manufacturing	56%	36%	5%

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	16%	-18%	-4%
Proffesional Services	69%	30%	12%
FIRE	43%	42%	19%
Admininstrative	37%	19%	5%
Health	71%	37%	-6%
Manufacturing	55%	10%	-12%

Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using (e(β*i*) – 1), where β*i* is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B stations are those opened between 1990 and 1999. The symbol (#) refers to the statistically insignificant values.

For group C stations, professional services sector, yet again, is the only one with positive and statistically significant estimated coefficient of the *quarter mile* buffer (1.335). For the other five industry sectors, the estimated coefficients are statistically insignificant within all the three station buffers of group C stations (see Table 28 and Figure 28). Group C stations have several limitations, however, such as small sample size (only six stations) that are analyzed at one period (2010 only), as discussed in the preceding section. Therefore, it should not be surprising to see most estimates being statistically insignificant within the three station buffers of group C.

Figure 28. The predicted effects of inward firm relocation for selected industry sectors by distance from the more recently opened stations (group C stations).



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	132%	-58%	-36%
Proffesional Services	280%	-17%	-17%
FIRE	246%	-100%	-18%
Admininstrative	288%	6%	17%
Health	-26%	-78%	34%
Manufacturing	278%	70%	59%

Note: The y-axis shows the predicted effect of inward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients using (e(β*i*) – 1), where β*i* is the coefficient for the dummy variable of the respective distance-to-station buffer.

Group C stations are those opened after 2000. The symbol (#) refers to the statistically insignificant values.

Overall, the inward relocation analysis by industry sector provide compelling evidence that the areas within a *quarter mile* of the passenger rail stations experienced higher levels of inward relocations of relocating firms in the professional services, FIRE, and administrative industries, compared to control areas. In other words, close proximity to rail stations is a more important factor for the location decisions of relocating firms in the professional services, FIRE, and administrative industries, compared to other industries. Some evidence of positive inward relocation was also found for moving firms belonging to retail, health, and manufacturing sector within the *quarter mile* of group B stations.

Turning to the control variables on agglomeration, the sign of the coefficients for population density is negative in all the inward relocation models by industry sector (see Table 28). The negative association between population density and inward relocation is similar to the association found and discussed in previous section (inward relocation analysis by firm size). Employment density, on the other hand, is positively associated with inward firm relocation across all the six industry sectors, but the magnitude and the significance of the influence varies. For instance, employment density is not a statistically significant factor for moving firms in the retail and health sectors (coefficient is 0.026 and 0.012, respectively), whereas employment density clearly matters to the location decisions of relocating firms in the other four sectors (coefficients range between 0.084 and 0.101 and are statistically significant). To better understand the influence of agglomeration on inward relocation, one should examine how the presence of firms in an area influences the location decisions of relocating firms. Two variables are of particular importance: (1) the total number of firms, which represents urbanization economies, and (2) total number of firms in the same industry category, which represents localization (specialization) economies.

The estimated coefficients of these two variables suggest that localization economies matter more than urbanization economies in the location decisions of moving firms belonging to five out of the six industry sectors. The exception is moving firms belonging to the retail sector, where urbanization economies is valued above localization economies. The importance of urbanization economies (total number of firms) for firms in the retail sector is also above that of any other sector (i.e. for retail moving firms the estimated coefficient for the total number of firms variable is 0.020, which is higher than any other sector, as shown in Table 28). The number of firms in the same industry category also positively predicts the number of inward firm relocations across all of the six sectors (coefficients ranging between 0.012 for retail moving firms and 0.169 for administrative moving firms, as shown in Table 28), suggesting that relocating firms are more attracted to areas with a high presence of firms in their own-industry sectors (i.e. evidence of localization economies).

Focusing on socio-economic variables, the median household income within a Census block is not a strong predictor of inward firm relocation. For instance, the association is positive but small for moving firms in the administrative sector (0.005) and in the professional services sector (0.002), as shown in Table 28. This positive association with inward relocation suggests that moving firms in these sectors are slightly more likely to be drawn to high-income neighborhoods that normally exemplify higher demand for goods and services than low-income neighborhoods. Median housing rent also has positive associations with inward relocation. For instance, the estimated coefficient for median housing rent is 0.206 for moving firms in the professional services sector and 0.126 for moving firms in the administrative sector (see Table 28).

In a Census block, the percentage of population that is college educated is a strong predictor of the location selections of relocating firms belonging to the professional services (1.725), administrative (1.290), and FIRE sector (1.057), as shown in Table 28. It is not surprising that human capital greatly matters to moving firms belonging to these sectors because they are knowledge-oriented. Regarding race, the association is positive (0.342) between the percentage of African American population and inward firm relocation of smaller firms with five or less employees. This association is negative (-0.141), however, for larger relocating firms with more than five employees. Large relocating firms with more than five employees tend to avoid areas with high percentages of African Americans, which suggest a possibility that labor-intensive moving firms have racial preferences.

Unemployment rate is also positively associated with inward firm relocation of firms in the professional services (1.410), administrative (2.048), and manufacturing (2.832) sectors. This positive association suggests that the availability of labor due to high rates of unemployment attracts firms belonging to these three sectors. High rates of unemployment in an area may also have diminishing influences on wages, which can be a desirable location for certain firms to move to. Firms in the manufacturing sector, for instance, often relay on inexpensive labor. The unemployment rate coefficients are negative and statistically insignificant for the retail (-0.275) and FIRE (-0.145) sectors. High unemployment rate can also be a sign of low business competitiveness and lower quality of life in an area, which can be unattractive local attributes for certain relocating firms, such as retailers.

Related to the spatial context, the distance to CBD is positively associated with inward firm relocation across all industry sectors (coefficients ranging between 0.031 for retail sector and 0.068 for professional services). The positive association suggest that proximity to CBD is not a key factor in the location decisions of relocating firms. Distance to the nearest highway exit, on the other hand, is a key factor in the location decisions of moving firms. The coefficients of the distance to highway variable are negative across all industry sectors, ranging between -0.049 for the retail sector and -0.111 for the FIRE sector. These negative coefficients suggest that the greater the distance from the nearest highway exit, the lower the number of inward firm relocations will be within any given area (see Table 28). The property tax variable has a statistically significant and negative coefficient only in the model of FIRE industry sector (-0.121), suggesting that higher tax rates decrease the likelihood of inward firm relocation of firms in the FIRE industry.

5.3. Outward firm relocation by size: regression results

The analysis in this section confirms the trend found earlier in the firm retainment analysis by firm size in Chapter 4. When controlling for other factors, areas within a short walking distance to the mature rail stations (group A) experienced higher rates of inward firm relocation than outward firm relocation (i.e. positive net firm relocation) of larger firms with more than five employees, compared to control areas located more than one mile from stations. In comparison, areas near stations opened after 1990 (group B and C) experienced lower rates of inward relocation than outward relocation (i.e. negative net firm relocation), compared to control areas. This section discusses the results of the analysis on outward firm relocation considering variables related to proximity to rail stations and other control variables on agglomeration, socio-economics, and spatial context.

Table 30 and Table 31 show the regression results of outward firm relocation using the standard NB and the PS-weighted NB regression methods. Outward firm relocation patterns are analyzed to determine whether or not areas near the rail stations have experienced positive net probability of firm relocation (i.e. the difference between the probability of inward relocation and the probability of outward relocations), compared to control areas located more than a mile from the stations. A positive net relocation in a given area signals its appeal to relocating firms. This section calculates the net probability of firm relocation within the three distance-to-station buffers by comparing the regression coefficients obtained from each set of inward relocation and outward relocation models.

Note that the outward relocation regressions include a control variable capturing the average age of firms within Census blocks since the age of firm may have an influence on the decision of firms on whether or not to relocate elsewhere, as discussed in Chapter 2. The estimated coefficients of *average age of firms* are positive and statistically significant across all of the models by firm size (ranging between 0.011 and 0.029), suggesting that higher average age of firms within a census block has a positive association with the number of outward firm relocations (see Table 30). In other words, the results suggest that a high average age of firm survival in a Census block is a push factor. This finding is surprising because older firms are normally more embedded in their spatial environment. One plausible explanation, however, is that the presence of a high number of more embedded firms in a block raises competitiveness which consequently force less competitive firms to relocate out.

Table 30. Regression coefficients of the standard NB method: outward firm relocation by firm size.

Dependent Variable: Outward Relocations	(1) All Firms	(2) Sole Proprietor	(3) Firms > 1 employee	(4) Firms <=5 employees	(5) Firms >5 employees
Distance to Rail station (in mi)	-0.063***	-0.075***	-0.058***	-0.064***	-0.056***
Group A stations: within 1/4 mile	0.442***	0.256	0.502***	0.377***	0.556***
Group A stations: within 1/4 to 1/2 mi	0.079	0.037	0.089	0.027	0.263**
Group A stations: within 1/2 to 1 mi	-0.283***	-0.316***	-0.309***	-0.364***	-0.223**
Group B stations: within 1/4 mile	0.795***	0.869***	0.785***	0.788***	0.822***
Group B stations: within 1/4 to 1/2 mi	0.337***	0.283*	0.355***	0.247**	0.514***
Group B stations: within 1/2 to 1 mi	0.117*	0.274***	0.063	0.104	0.273***
Group C stations: within <=1/4 mile	1.349*	1.557*	1.212	1.721**	0.548
Group C stations: within 1/4 to 1/2 mi	0.716*	0.869*	0.494	0.822**	-0.022
Group C stations: within 1/2 to 1 mi	0.274	0.475*	0.113	0.317	-0.004
Accessibility ratio	1.256***	0.916***	1.261***	1.161***	1.199***
Population per sq. mi. (in 1000s)	-0.062***	-0.066***	-0.064***	-0.066***	-0.068***
Employee per sq. mi. (in 1000s)	0.109***	0.158***	0.093***	0.132***	0.061***
Number of firms	0.010***	0.008***	0.012***	0.010***	0.017***
Firm births	0.003***	0.011***	0.001*	0.004***	-0.002
Firm closures	0.002***	0.007***	0.001***	0.002***	0.003***
Firm inward relocations	-0.020***	-0.038***	-0.021***	-0.026***	-0.026***
Average age of firms	0.020***	0.011***	0.023***	0.016***	0.029***
Median HH Income (in $1000s)	0.009***	0.011***	0.005***	0.010***	-0.002*
Unemployment rate	2.942***	3.925***	2.421***	3.474***	1.251***
Percent college educated	0.125	0.151	0.270**	0.129	0.391***
Percent African-American	0.396***	0.795***	0.270***	0.543***	0.044
Median housing rent (in $1000s)	0.276***	0.370***	0.239***	0.316***	0.152***
Distance to highway (in mi)	-0.104***	-0.038**	-0.118***	-0.083***	-0.166***
Distance to CBD (in mi)	0.052***	0.076***	0.048***	0.057***	0.049***
Property tax (in $1000)	-0.094***	-0.092*	-0.108***	-0.110***	-0.101**
Constant	-1.820***	-3.780***	-1.439***	-2.277***	-1.224***
ln_r					
Constant	1.144***	1.806***	1.217***	1.335***	1.510***
ln_s					
Constant	-1.008***	-0.266**	-1.237***	-0.865***	-1.393***
N. of cases	116820	116820	116820	116820	116820
Log Likelihood	-40556.32	-16739.196	-33036.482	-33011.808	-17805.714
chi2	8171.234	4809.592	6271.058	7365.198	3656.917

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

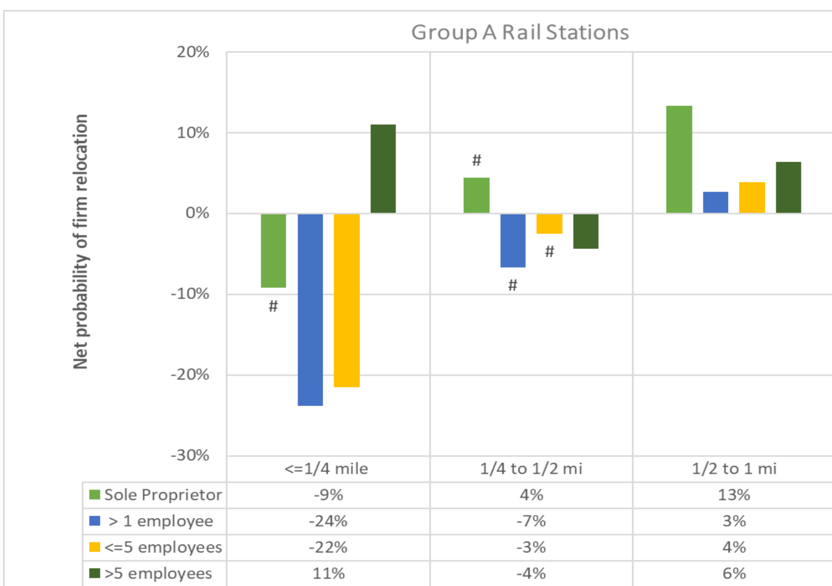
Table 31. Regression coefficients of the PS-weighted NB method: outward firm relocation by firm size.

Dependent Variable: Inward Relocations	(1) All Firms	(2) Sole Proprietor	(3) Firms > 1 employee	(4) Firms <=5 employees	(5) Firms >5 employees
Distance to Rail station (in mi)	-0.043***	-0.059***	-0.042***	-0.043***	-0.048***
Within <=1/4 mile (Group B)	0.427***	0.696***	0.337**	0.493***	0.347*
Within 1/4 to 1/2 mi (Group B)	0.197*	0.161	0.192	0.137	0.288*
Within 1/2 to 1 mi (Group B)	0.096	0.320**	0.000	0.125	0.042
Accessibility ratio	0.894***	0.974***	0.789***	0.834***	0.911***
Population per sq. mi. (in 1000s)	-0.056***	-0.057***	-0.049***	-0.054***	-0.046*
Employee per sq. mi. (in 1000s)	0.092***	0.137***	0.061***	0.109***	0.007
Number of firms	0.040***	0.010***	0.045***	0.031***	0.047***
Firm births	0.027***	0.041***	0.014***	0.034***	-0.004
Firm closures	-0.014***	0.015***	-0.020***	-0.001	-0.022***
Firm outward relocations	-0.022**	-0.091***	-0.021**	-0.077***	-0.001
Average firm age	0.016***	0.010***	0.018***	0.014***	0.020***
Median HH Income (in $1000s)	0.008***	0.012***	0.006***	0.010***	0.000
Unemployment rate	2.342***	3.943***	1.456***	2.977***	-0.293
Percent college educated	-0.273*	-0.093	-0.235*	-0.224*	-0.213
Percent African-American	0.268***	0.716***	0.187***	0.388***	0.056
Median housing rent (in $1000s)	0.173***	0.249***	0.156***	0.191***	0.147**
Distance to highway (in mi)	-0.087***	-0.028	-0.113***	-0.061***	-0.181***
Distance to CBD (in mi)	0.037***	0.063***	0.033***	0.042***	0.036***
Property tax (in $1000)	0.011	-0.107*	0.033	-0.019	0.018
Constant	-3.647***	-5.656***	-3.587***	-4.227***	-3.763***
chi2	20504.199	13508.474	13612.192	20703.096	12705.284
N. of cases	101859	101859	101859	101859	101859

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

Starting with the mature rail stations (group A), Figure 29 shows the difference between the predicted probability of inward firm relocation and the predicted probability of outward firm relocation (net firm relocation) by firm size within three station buffers. Areas within a quarter mile of the mature rail stations have, on average, experienced higher net probability of firm relocation (11%) of larger firms with more than five employees, compared to areas more than a mile from the stations. The net probability of firm relocation is also positive within the *half to one mile* buffer of group A stations. For instance, the net probability of firm relocation is 6% for firms with more than five employees and 4% for firms with five or less employees, compared to control areas (see Figure 29). The probability of firm retainment was also positive and higher for larger firms within the *half to one mile* buffer of group A stations, as shown in Chapter 4. These results suggest that larger firms benefit from access to the workforce provided by rail transit more than smaller firms.

Figure 29. The difference between the predicted probability of inward relocation and the predicted probability of outward relocation by firm size and distance from the mature rail stations (group A stations).

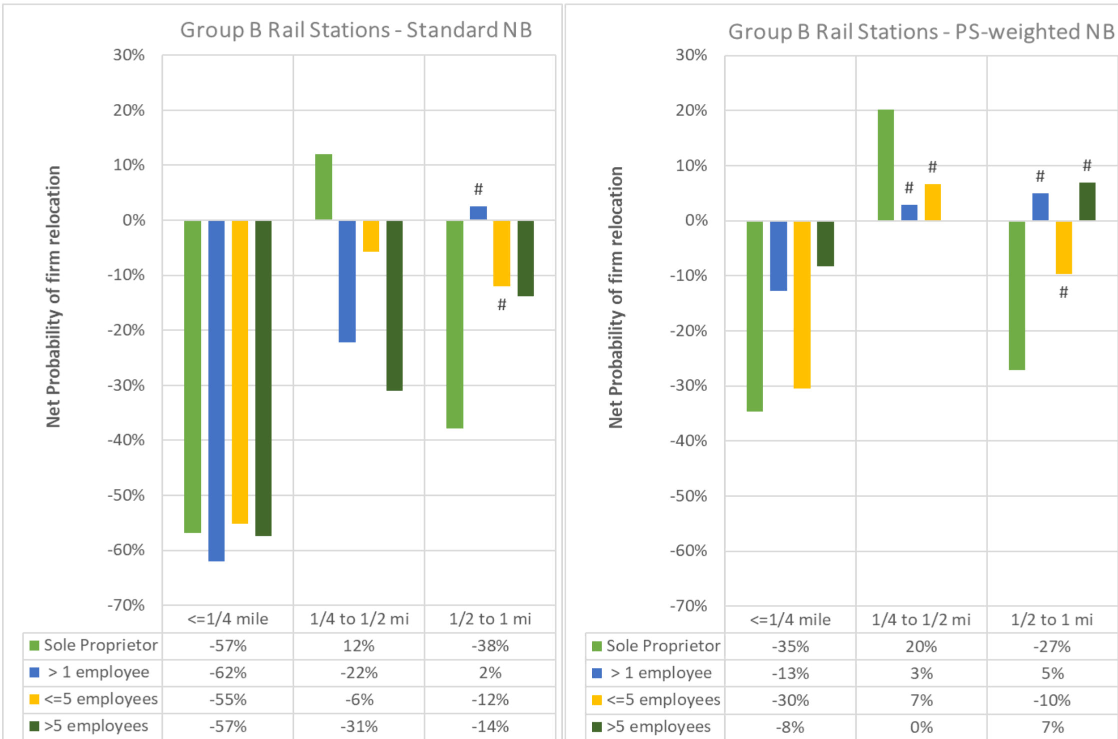


Note: The y-axis shows the difference between inward and outward firm relocation relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward firm relocation and outward firm relocation using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989. The symbol (#) shows the statistical insignificance of the estimated coefficients of both inward and outward relocation.

The net predicted probability of firm relocation is mostly negative in areas within close proximity to group B and C stations (stations opened after 1990). For group B stations (opened between 1990 and 1999), the difference between inward and outward relocation is negative across the three station buffers (ranging between -14 and -57% for relocating firms with more than five employees and between -6% and -55% for relocating firms with five or less employees), compared to control areas (see Figure 30). These negative net probabilities of firm relocation suggest that blocks near group B stations experienced a net loss in the number of relocating firms. The only exception is the positive net probability of relocation for sole proprietors in the *quarter to half mile* buffer (12%).

The estimated coefficients remain mostly negative for group B station-buffers even after controlling for the endogeneity of the stations. In the PS-weighted method, for instance, the net probability of firm relocation within the *quarter mile* buffer of group B stations is -8% for relocating firms with more than five employees and -30% for relocating firms with five or less employees, compared to control areas (see Figure 30). Clearly, the accessibility benefits provided by passenger rail in Maryland are still to be realized in areas near the stations opened after 1990. Areas near many of the stations opened after 1990 remained predominantly residential because the State of Maryland lacked any deliberate plans to promote economic development around these stations.

Figure 30. The difference between the predicted effects of inward relocation and the predicted effect of outward relocation within group B station buffers by firm size (comparing the Standard and PS-weighted NB methods).



Group B Rail Stations - Standard NB

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	-57%	12%	-38%
> 1 employee	-62%	-22%	2%
<=5 employees	-55%	-6%	-12%
>5 employees	-57%	-31%	-14%

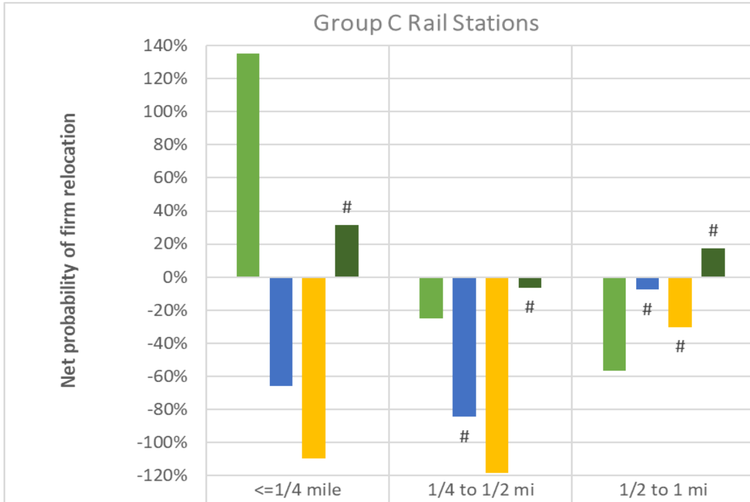
Group B Rail Stations - PS-weighted NB

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	-35%	20%	-27%
> 1 employee	-13%	3%	5%
<=5 employees	-30%	7%	-10%
>5 employees	-8%	0%	7%

Note: The y-axis shows inward-to-outward relocation net effects (ratio) relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward and outward firm relocation models using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B passenger rail stations are those opened between 1990 and 1999. The symbol (#) shows the statistical insignificance of the estimated coefficients of both inward and outward relocation.

Areas near more recently opened stations (group C, opened after 2000) also experienced negative net probability of firm relocation in most cases (see Figure 31). For instance, the net probability of firm relocation is -109% within the *quarter mile* buffer and -118% within the *quarter to half mile* buffer of group C stations, compared to control areas. The only exception is areas within a *quarter mile* of the stations which experienced a positive net relocation for sole proprietor (135%). The net relocation of larger firms with more than five employees is also positive within the *quarter mile* buffer but statistically insignificant (31%).

Figure 31. The difference between the predicted effects of inward relocation and the predicted effect of outward relocation within group C station buffers by firm size.



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Sole Proprietor	135%	-25%	-56%
> 1 employee	-66%	-85%	-7%
<=5 employees	-109%	-118%	-30%
>5 employees	31%	-6%	17%

Note: The y-axis shows the net predicted effect of inward-to-outward relocations relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward firm relocation and outward firm relocation models using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group C passenger rail stations are opened after 2000. The symbol (#) shows the statistical insignificance of the estimated coefficients of both inward and outward relocation.

Turning to other control variables, Figure 32 summarizes the predicted probability of inward firm relocation, outward firm relocation, and most importantly, the net probability of firm relocation for each of the control variables that represent agglomeration, socio-economic characteristics, and spatial context. The inward and outward relocation probabilities are obtained from the estimated coefficients in Table 26 and Table 30, respectively. The net probability of firm relocation is the difference between the probability of inward relocation and the probability of outward relocation, as explained earlier.

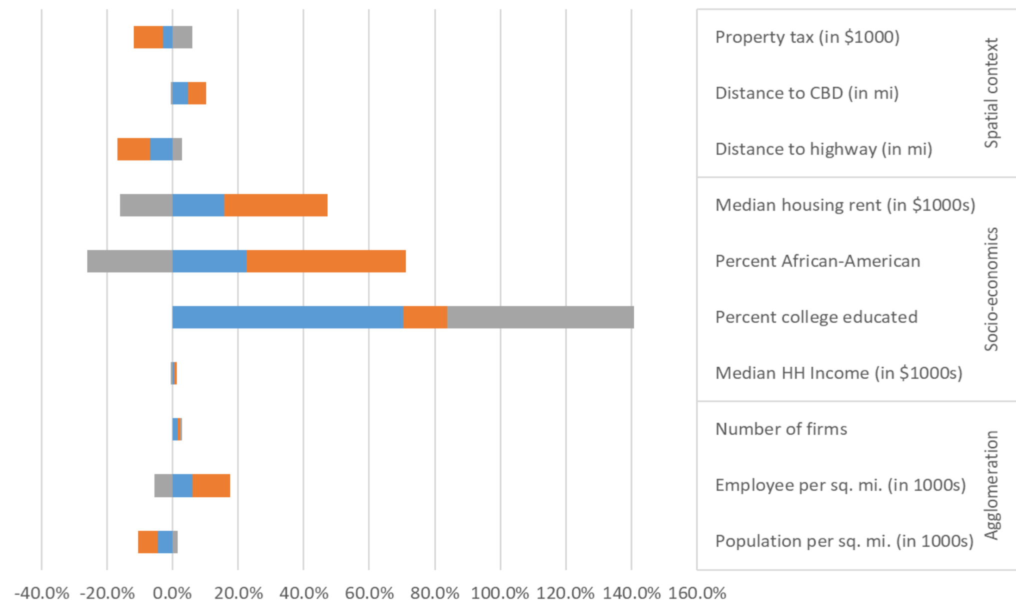
As this study hypothesized, population density has a positive effect on the net probability of firm relocation. The net predicted effect of population density on firm relocation is positive (1.4%) since the magnitude of the negative effect of population density on outward firm relocation ($e^{-0.062} - 1$= -0.060) is larger than the magnitude of the negative effect of population density on inward firm relocation ($e^{-0.047} - 1$= -0.046) (see Figure 32). In other words, the higher the population density at the micro-level, the lower the likelihood of existing firms to relocate out, holding all else equal. This measure may capture urbanization economies which appear to matter to firms even at the very local level.[1]

The predicted effect of employment density on net firm relocation is opposite to the one found for population density. As discussed earlier, employment density has, on average, a positive influence on inward firm relocation (6.1%). On the other hand, employment density has a positive association with outward firm relocation (11.5%) that is stronger in magnitude compared to its positive association with inward relocation leading to a negative effect (-5.4%) of employment density on net firm relocation (see Figure 32). In other words, employment density is both a push and a pull factor for relocating firms, but the push factor is stronger. Higher employment densities at the micro-level may suggest higher employment competitiveness which can consequently push less competitive firms to relocate elsewhere. Interestingly, the coefficient of outward relocation is larger for smaller firms with five or less employees (0.132) than larger firms with more than five employees (0.06), suggesting that higher employment

[1] Not that each Census block has its population density data from the Census block group level, as discussed in Chapter 3.

competitiveness is more likely to push smaller firms than larger firms to relocate out (see Table 30).

Figure 32. Predicted probabilities of inward firm relocation, outward firm relocation, and net relocation (inward - outward) for selected control variables (all firm sizes).



	Agglomeration			Socio-economics				Spatial context		
	Population per sq. mi. (in 1000s)	Employee per sq. mi. (in 1000s)	Number of firms	Median HH Income (in $1000s)	Percent college educated	Percent African-American	Median housing rent (in $1000s)	Distance to highway (in mi)	Distance to CBD (in mi)	Property tax (in $1000)
Inward relocation	-4.6%	6.1%	1.4%	0.4%	70.4%	22.6%	15.6%	-6.9%	4.8%	-3.0%
Outward relocation	-6.0%	11.5%	1.0%	0.9%	13.3%	48.6%	31.8%	-9.9%	5.3%	-9.0%
Net relocation	1.4%	-5.4%	0.4%	-0.5%	57.1%	-26.0%	-16.2%	2.9%	-0.5%	6.0%

Predicted Probability

Note: The x-axis shows the predicted effects of inward relocation, outward relocation, and the net inward-to-outward relocation relative to control Census blocks, all else held equal. The y-axis shows the predicted effects of each control variable except transit-related variables. The percentages are calculated from the regression coefficients of inward relocation and outward relocation models of all firms. The net effect of inward-to-outward relocation is calculated using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective control variable.

Similar to employment density, the total number of firms in a block is positively associated with inward firm relocation (1.4%) and also with outward firm relocation (1.0%), as shown in Figure 32. Unlike employment density, however, the net probability of relocation for the total number of firms is positive (0.4%), which provides additional evidence that urbanization economies can still occur at the micro-level, in smaller magnitudes, however.

In relation to the socio-economic factors, the percentage of college educated population has the highest positive influence on the net probability of firm relocation across the control variables (57.1%), as shown in Figure 32. The association is positive between the percentage of college educated population and the probability of both inward relocation (70.4%) and outward relocation (13.3%), which means that higher education levels in a block can be both a push and a pull factor for firms. Clearly, the inward firm relocation effects of the percentage of college educated population is considerably stronger than the outward firm relocation effects, resulting in the net positive effect on firm relocation. Higher percentages of college educated population in a block suggest high quality of its workforce in terms of ability to learn new skills and quality of life, which can be a key factor in the relocation decisions of firms that seek high quality labor.

Regarding race, the net probability of firm relocation is negative for the percentage of African American population (-26%), suggesting that the higher the percentage of African American population within a block, the lower the net probability of firm relocation. The regression results of the two variables on income suggest that, on average, the higher the income level within a block, the lower the net probability of relocation. Median household income has a positive effect on outward relocation (0.9%) that is stronger than its effect on inward relocation (0.4%), resulting on a net negative effect on firm relocation (-0.5%), as shown in Figure 32. The median housing rent also has a net negative effect on firm relocation (-16.2%).

The regression results of the control variable on unemployment rate may appear surprising at first glance because of its large in magnitude and statistically significant association with outward firm relocation suggested by the large regression coefficients in Table 30. The regression results on unemployment suggest, however, that what drives firms to relocate out of a certain location is a loss in the number employees. Alternatively, firms that have already made the decision to relocate out from a certain location may have chosen to lay off employees hence explaining the positive association between unemployment rate and outward firm relocation.

Focusing on the last group of control variables, on spatial context, distance to highway is negatively associated with incidents of outward firm relocation suggested by the negative coefficients across all the models by firm size (see Table 30). The negative coefficients suggest that areas within proximity to a highway exit have lower likelihood of outward firm relocation compared to areas farther away from a highway exit. As shown in Figure 32, the predicted effect on net firm relocation is overall positive for proximity to highway since its impact on averting outward firm relocation is stronger than its negative impact on attracting inward relocation.

5.4. Outward firm relocation by industry sector: regression results

This section examines outward firm relocation across selected industry sectors within the study area in relation to proximity to the passenger rail stations. The main objective of this section is to determine the net probability of firm location of the three station buffers across the six selected industry sectors (i.e. the difference between the probability of inward relocation and the probability of outward relocation). The results overall show that areas within close proximity to the mature rail stations experienced positive probabilities of net firm relocation of firms belonging to FIRE and administrative sectors, while areas within close proximity to group B and C stations generally experienced negative net probabilities of firm relocation.

Table 32 and Table 33 provide the regression results of the outward firm relocation across selected industry sectors using the standard NB and the PS-weighted NB methods. The dependent variable is the count of outward firm relocations in each industry sector (six models) regressed on control variables related to proximity to rail stations, agglomeration, socio-economic, and spatial context. The results suggest that there are considerable differences in the predicted effect of areas near the rail stations on outward firm relocation across different industry sectors. The predicted effects of the control variables on agglomeration, socio-economics, and spatial context also differ across the models by industry sectors in Table 32.

The continuous *distance to rail station* variable has negative coefficients across all industry sectors analyses (ranging between -0.040 and -0.074), suggesting that, on average, areas closer to the rail stations experience higher rates of outward firm relocation, all else held equal. However, the inward relocation analysis showed that the association between the distance to rail station and inward firm relocation is also negative, which mean that areas closer to the rail stations also experience higher rates of inward relocation. Therefore, this section examines in detail the predicted effects of areas near the rail stations (the three distance-to-station buffers) on net firm relocation by comparing the predicted effects of inward and outward firm relocation. Starting with the mature stations (group A stations), Figure 33 shows that, all else held equal, blocks within a quarter mile of the mature rail stations had on average positive impact on net firm relocation for four out of the six analyzed sectors (i.e. professional services, FIRE, administrative, and manufacturing sectors).

Table 32. Estimated coefficients of the standard NB method: outward firm relocation by selected industry sectors.

Dependent Variable: Outward Relocations	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail station (in mi)	-0.040***	-0.076***	-0.074***	-0.060***	-0.053***	-0.044**
Group A stations: within 1/4 mile	0.487**	0.886***	0.190	0.147	-0.041	0.261
Group A stations: within 1/4 to 1/2 mi	0.233	0.235*	0.170	-0.153	0.046	-0.232
Group A stations: within 1/2 to 1 mi	-0.177	-0.303***	-0.612***	-0.469***	-0.234	-0.107
Group B stations: within 1/4 mile	0.698***	1.215***	0.708***	0.441*	0.547**	0.636**
Group B stations: within 1/4 to 1/2 mi	0.326*	0.397**	0.272	0.326*	0.176	0.464*
Group B stations: within 1/2 to 1 mi	0.031	0.110	0.010	0.200	-0.225	0.324*
Group C stations: within <=1/4 mile	0.964	1.894**	1.530	1.915**	1.160	-18.923
Group C stations: within 1/4 to 1/2 mi	-0.471	1.439***	0.643	-2.423*	-1.527	-18.861
Group C stations: within 1/2 to 1 mi	0.193	0.428	0.205	0.230	0.095	-0.101
Accessibility ratio	0.933**	0.918***	1.229***	1.026***	1.383***	1.519***
Population per sq. mi. (in 1000s)	-0.044***	-0.074***	-0.093***	-0.073***	-0.028*	-0.107***
Employee per sq. mi. (in 1000s)	0.063**	0.145***	0.140***	0.138***	0.075***	0.114**
Number of firms	0.019***	0.010***	0.014***	0.006***	0.001	0.014***
Firms in the same industry sectors	0.006	0.008**	0.047***	0.180***	0.097***	0.173***
Firm births	-0.002	0.007***	0.008***	0.003	0.012***	-0.002
Firm closures	0.006***	0.003***	0.001	-0.004***	0.009***	0.006***
Firm inward relocations	-0.036***	-0.030***	-0.057***	-0.035***	-0.019***	-0.046***
Average age of firms	0.019***	0.017***	0.019***	0.013***	0.017***	0.019***
Median HH Income (in $1000s)	-0.001	0.007***	0.001	0.008***	0.005***	0.002
Unemployment rate	1.536**	3.158***	2.767***	2.241***	2.073***	1.082
Percent college educated	-0.069	1.121***	0.817***	-0.062	0.845***	-0.609*
Percent African-American	0.076	0.577***	0.189	0.516***	0.472***	0.066
Median housing rent (in $1000s)	0.232***	0.304***	0.043	0.194***	0.284***	0.043
Distance to highway (in mi)	-0.107***	-0.117***	-0.129***	-0.039*	-0.088**	-0.176***
Distance to CBD (in mi)	0.035***	0.084***	0.054***	0.055***	0.035***	0.049***
Property tax (in $1000)	-0.085	-0.158***	-0.149**	-0.065	-0.093	0.051
Constant	-1.156***	-3.509***	-2.354***	-3.193***	-3.311***	-2.181***
ln_r						
Constant	2.672***	1.560***	2.191***	2.459***	2.253***	2.757***
ln_s						
Constant	-1.074***	-0.845***	-0.803***	-0.122	-0.769***	-0.679***
N. of cases	116820	116820	116820	116820	116820	116820
Log Likelihood	-8684.918	-14100.135	-7846.35	-9001.28	-6459.769	-4592.683
chi2	1991.137	3929.085	2357.543	3715.128	2216.41	1536.544

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

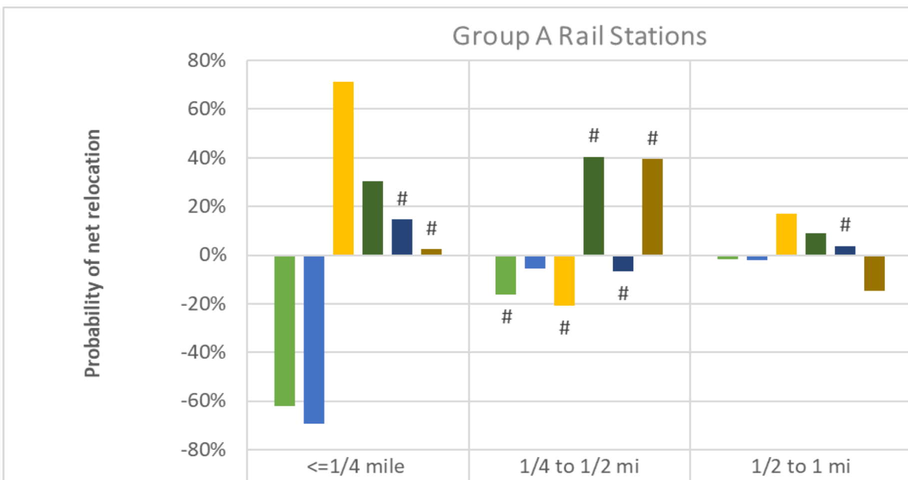
Table 33. Estimated coefficients of the PS-weighted NB method: outward firm relocation by selected industry sectors.

Dependent Variable: Outward Relocations	(1) Retail	(2) Professional Services	(3) FIRE	(4) Administrative	(5) Health	(6) Manufacturing
Distance to Rail station (in mi)	-0.033**	-0.066***	-0.058***	-0.055***	-0.042***	-0.049***
Within <=1/4 mile (Group B)	0.437	0.711***	0.323	0.215	0.898***	0.254
Within 1/4 to 1/2 mi (Group B)	0.139	0.253	0.195	0.403*	0.139	0.128
Within 1/2 to 1 mi (Group B)	-0.157	0.207	-0.135	0.233	-0.194	0.107
Accessibility ratio	0.243	0.999**	1.395***	0.658	1.888***	1.304*
Population per sq. mi. (in 1000s)	-0.049**	-0.048***	-0.037	-0.070***	-0.016	-0.070*
Employee per sq. mi. (in 1000s)	0.082**	0.106***	0.063	0.133***	0.051	0.042
Number of firms	0.026***	0.011***	0.013***	0.003	0.010***	0.008***
Firms in the same industry sectors	0.042***	0.098***	0.137***	0.194***	0.086***	0.296***
Firm births	0.008**	0.020***	0.015***	0.015***	0.015***	0.015***
Firm closures	-0.010***	-0.004*	-0.007***	0	0.005**	-0.003*
Firm inward relocations	-0.085***	-0.049***	-0.073***	-0.023**	-0.058***	-0.055***
Average age of firms	0.016***	0.014***	0.016***	0.012***	0.015***	0.018***
Median HH Income (in $1000s)	0.003*	0.009***	0.005***	0.008***	0.006***	0.005**
Unemployment rate	1.799***	2.144***	1.869**	2.274***	1.550*	0.982
Percent college educated	-0.551*	0.166	0.334	-0.280	0.654**	-1.118***
Percent African-American	0.065	0.577***	0.249*	0.604***	0.549***	0.100
Median housing rent (in $1000s)	0.148*	0.253***	0.073	0.170**	0.321***	0.034
Distance to highway (in mi)	-0.093**	-0.113***	-0.117***	-0.039	-0.081**	-0.181***
Distance to CBD (in mi)	0.021*	0.073***	0.044***	0.049***	0.025**	0.051***
Property tax (in $1000)	0.021	-0.116	-0.135*	-0.073	-0.092	0.197*
Constant	-4.885***	-5.667***	-5.488***	-5.681***	-6.306***	-5.747***
chi2	6422.559	11840.065	6733.981	8978.717	6730.168	5891.078
N. of cases	101859	101859	101859	101859	101859	101859

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

The net predicted effects on firm relocation are mostly positive but small in magnitude in the *quarter to half mile* buffer and the *half to one mile buffer* of group A stations. Surprisingly, blocks within a half mile distance of the mature rail stations have experienced a negative net effect on firm relocation. In other words, for retail firms, the predicted effect of outward relocation is much higher than the predicted effect of inward relocation in areas within a half mile distance of group A stations (see Figure 33). The same blocks, however, had high net predicted effect of firm birth to firm closure of retail firms, as discussed earlier.

Figure 33. The difference between the predicted effects of inward relocation and the predicted effect of outward relocation within group A station buffers by selected industry sectors.

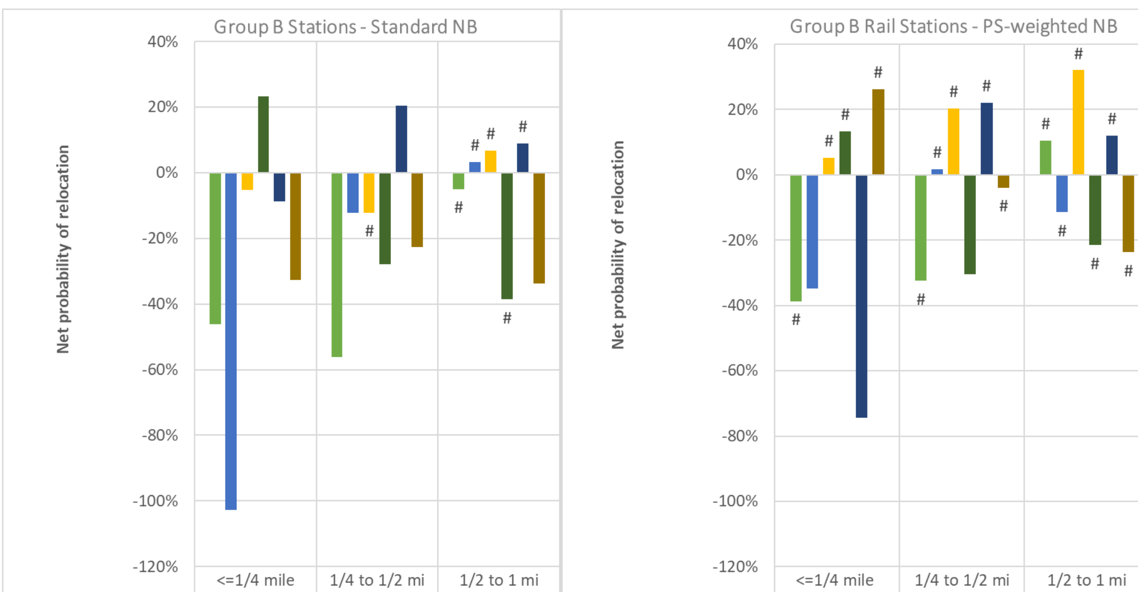


	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-62%	-16%	-2%
Proffesional Services	-69%	-5%	-2%
FIRE	71%	-21%	17%
Admininstrative	31%	40%	9%
Health	15%	-7%	4%
Manufacturing	2%	40%	-15%

Note: The y-axis shows the inward-to-outward net firm relocation effects relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward firm relocation and outward firm relocation models using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group A passenger rail stations are those opened between 1978 and 1989. The symbol (#) shows the statistical insignificance of the estimated coefficients of both inward and outward relocation.

The probability of net firm relocation in areas near group B rail stations also varies across different industry sectors. Some probabilities of net relocation vary in direction (whether positive or negative) across the two regression methods, which means that controlling for the endogeneity of areas near the rail stations can lead to different predicted effects of the measured outcome. For instance, the PS-weighted NB method predicts a positive net probability of firm relocation for firms belonging to FIRE sector in areas located within a *quarter to half mile* of group B stations (20%), whereas the net probability of relocation is negative in the standard NB method (-12), as shown in Figure 34.

Figure 34. The difference between the predicted effects of inward relocation and the predicted effect of outward relocation within group B station buffers by selected industry sectors (comparing the Standard and PS-weighted NB methods).



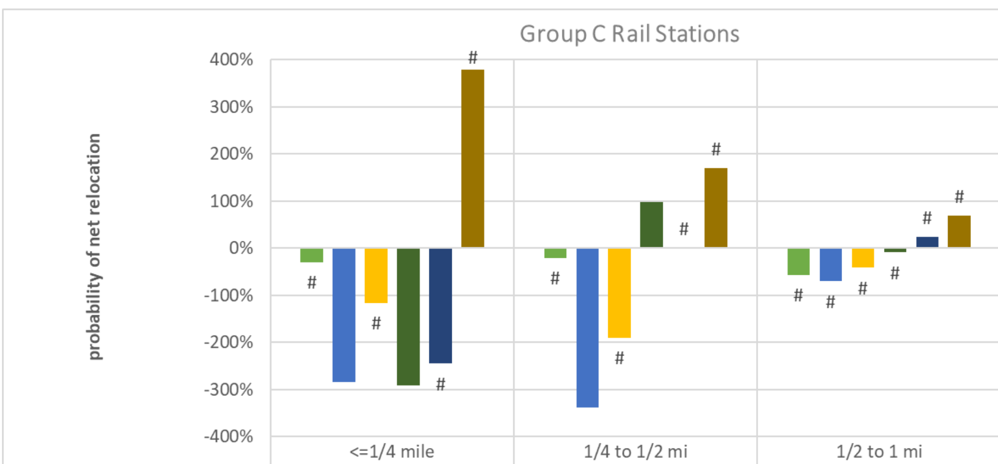
	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-46%	-56%	-5%
Proffesional Services	-103%	-12%	3%
FIRE	-5%	-12%	7%
Admininstrative	23%	-28%	-39%
Health	-9%	20%	9%
Manufacturing	-33%	-23%	-34%

	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-39%	-32%	11%
Proffesional Services	-35%	2%	-11%
FIRE	5%	20%	32%
Admininstrative	13%	-30%	-22%
Health	-74%	22%	12%
Manufacturing	26%	-4%	-24%

Note: The y-axis shows inward-to-outward relocation net effects relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward and outward firm relocation models using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group B passenger rail stations are those opened between 1990 and 1999. The symbol (#) shows the statistical insignificance of the estimated coefficients of both inward and outward relocation.

The probabilities of net relocation are mostly negative and statistically insignificant in areas located within close proximity to the more recently opened stations (group C stations opened after 2000). For instance, areas within a *quarter mile* of group C stations have negative probabilities of net relocation (ranging between -30% and -291%) across five of the six sectors (the only positive probability of net relocation is for manufacturing sector, but the estimated coefficients were statistically insignificant), as shown in Figure 35.

Figure 35. The difference between the predicted effects of inward relocation and the predicted effect of outward relocation within group C station buffers by selected industry sectors.



	<=1/4 mile	1/4 to 1/2 mi	1/2 to 1 mi
Retail	-30%	-21%	-57%
Proffesional Services	-285%	-338%	-70%
FIRE	-116%	-190%	-41%
Admininstrative	-291%	97%	-9%
Health	-245%	1%	24%
Manufacturing	378%	170%	68%

Note: The y-axis shows the inward-to-outward net firm relocation effects relative to control Census blocks, all else held equal. The x-axis shows the three distance-to-station buffers. The percentages are calculated from the regression coefficients of inward firm relocation and outward firm relocation models using [inward(e(βi) – 1) - outward(e(βi) – 1)], where βi is the coefficient for the dummy variable of the respective distance-to-station buffer. Group C passenger rail stations are those opened between 2000 and 2004.

For other control variables, the rest of this section discusses only the probabilities of net firm relocation that are different in direction across the selected industry sectors and opposite to the probabilities discussed in the previous section (section 5.4). The variable on the level of education shows some clear differences in the direction of influence on outward relocation across the six examined industry sectors. The percent of population with a college or higher degree has a negative association with outward firm relocation for firms in the retail, administrative, and manufacturing sectors, whereas the association is positive for the remaining sectors (see Table 32). The negative association is statistically significant (at the 95% level of confidence) for firms in the manufacturing sector only, suggesting that blocks with on average highly educated population are more likely to uphold firms in the manufacturing industry from outward relocation compared to firms belonging to other industry sectors.

The presence of high number of firms of the same sector within a block is positively associated with the number of outward relocations. For example, the higher the number of retail firms within a block, the higher the probability of outward relocation of retail firms within that block, which is obvious. What is not obvious is the effect of the presence of high number of retail firms within a block on net retail firm relocation. As discussed earlier, comparing the predicted effects of inward relocation and outward relocation provide an overall measure of what impact a control variable has on net firm relocation. For instance, turning back to the level of education variable, the average predicted effects of this variable on net firm relocation (inward predicted effect – outward predicted effect) reveal that higher average level of education within a block is a more important factor for firms in the professional service, Health, and FIRE industry sectors.

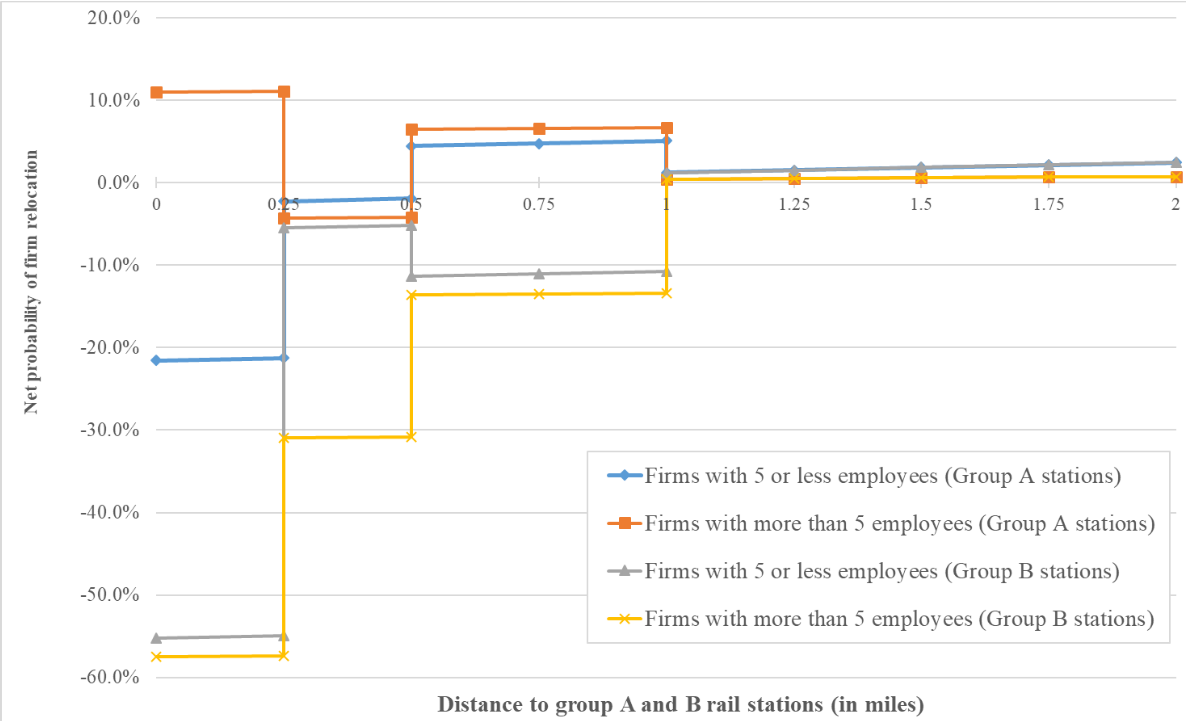
That is, blocks that have populations with high average education levels pull firms in the professional service, health, and FIRE sectors to relocate within, more than pushing them to relocate out. Similarly, across all industry sectors, the higher the number of firms of own-industry within a block, the higher the net probability of firm relocation.

5.5. Chapter summary

This chapter tested several factors that influence firm relocation. The main focus was on factors related to proximity to the passenger rail stations in The State of Maryland. Inward relocation and outward relocation are examined separately to determine the net probability of firm relocation within three station buffers (the *quarter mile, quarter to half mile,* and *half to one mile* buffer). The relocation analysis overall suggests that areas within close proximity to the mature rail stations (group A stations, opened before 1990) have experienced a net gain in the number of larger relocating firms with more than five employees, compared to control areas located more than a mile of the stations. On the contrary, areas within a mile of group B stations (opened between 1990 and 1999) have experienced a net loss in the number of relocating firms, compared to control areas (see Figure 36).

Areas within a mile of group C stations (opened after 2000) also experienced negative net probability of firm relocation across the industry sectors, with one exception. Firms in the administrate sector had positive but statistically insignificant net probability of firm relocation (see Appendix D). It is important to note that most of the estimated coefficients for group C station-buffers were statistically insignificant, which can lead to incorrect conclusions.

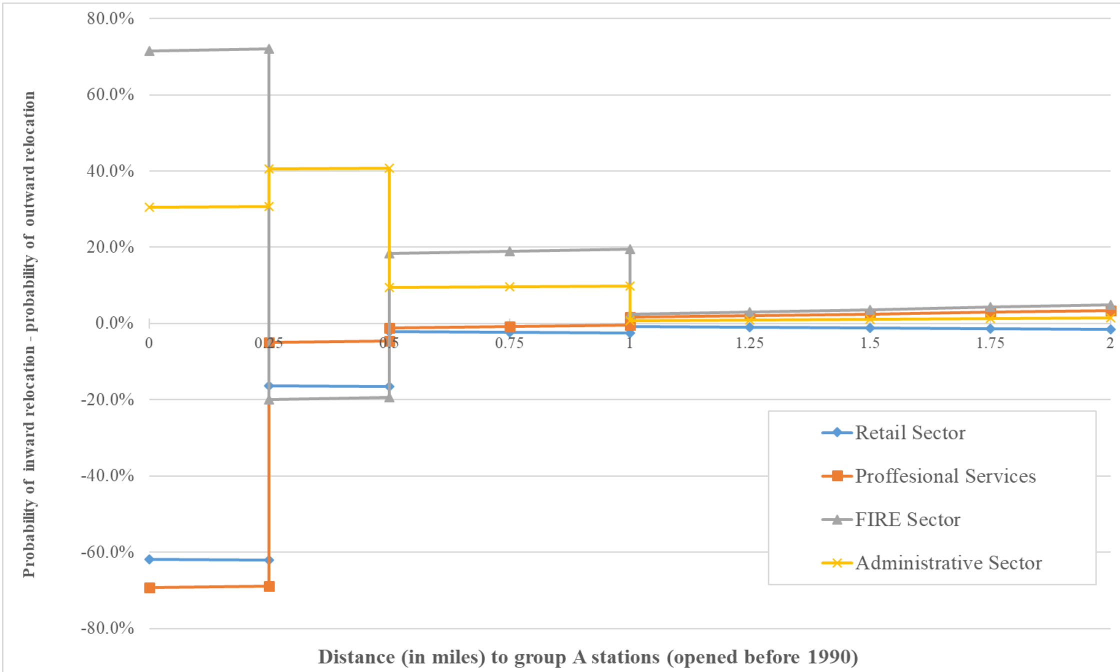
Figure 36. The net probability of firm relocation of station distance variables by firm size (group A and B stations).



The y-axis shows the predicted probabilities of net firm relocation as calculated from the coefficient values of inward relocation and outward relocation. The lines plotted in the graph are calculated as Y = inward[(e(βi) – 1) + (distance from station * (e(α)-1))] - outward[(e(βi) – 1) + (distance from station * (e(α)-1))] where βi is the coefficient for the dummy variable of the respective station buffer and α is the coefficient for the continuous distance-to-station variable.

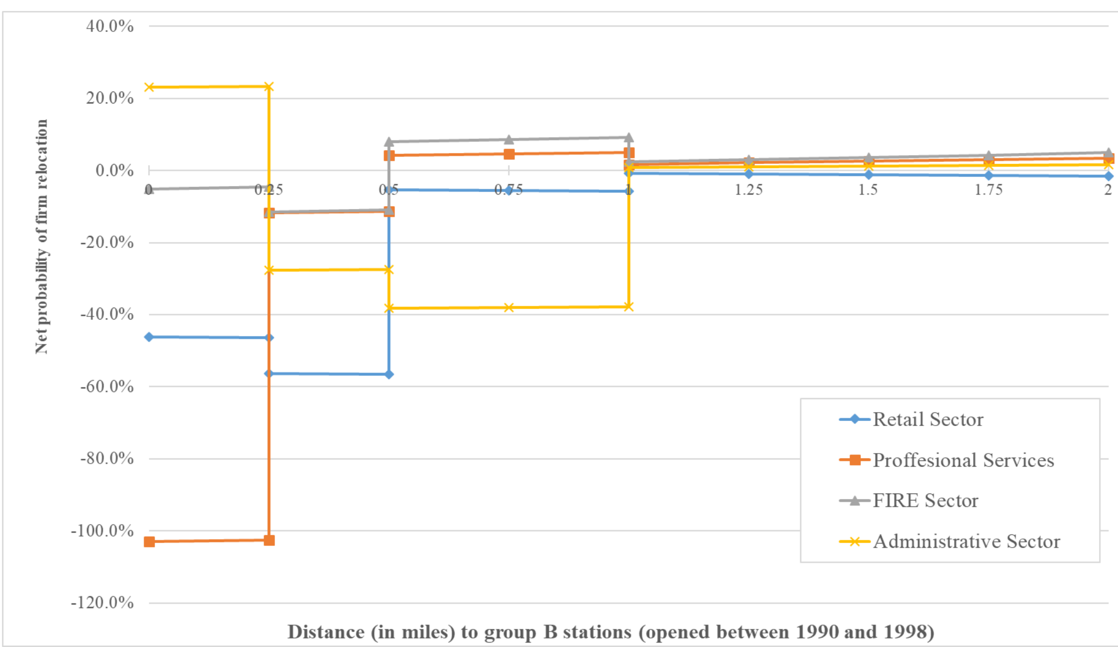
Similar to the firm relocation analysis by firm size, there are mixed relationships between proximity to rail stations and the net probability of firm relocation across the six industry sectors. Figure 37 summarizes the results of four industry sectors that have dominant presence in the study area by distance from the mature rail stations (group A). Firms belonging to the FIRE and administrative sectors are the most likely to benefit from areas within a mile of the mature rail stations. In comparison, areas within a quarter mile of group B stations also had a positive net probability of firm relocation for the administrative sector, but the net probability was negative in between a quarter to one mile of the stations (see Figure 38).

Figure 37. Probability of net firm relocation of station distance variables by selected industry sectors (group A stations).



The y-axis shows the predicted probabilities of net firm relocation as calculated from the coefficient values of inward relocation and outward relocation. The lines plotted in the graph are calculated as Y = inward[(e(βi) – 1) + (distance from station * (e(α)-1))] - outward[(e(βi) – 1) + (distance from station * (e(α)-1))] where βi is the coefficient for the dummy variable of the respective station buffer and α is the coefficient for the continuous distance-to-station variable.

Figure 38. Probability of net firm relocation of station distance variables by selected industry sectors (group B stations).



The y-axis shows the predicted probabilities of net firm relocation as calculated from the coefficient values of inward relocation and outward relocation. The lines plotted in the graph are calculated as Y = inward[(e(βi) – 1) + (distance from station * (e(α)-1))] - outward[(e(βi) – 1) + (distance from station * (e(α)-1))] where βi is the coefficient for the dummy variable of the respective station buffer and α is the coefficient for the continuous distance-to-station variable.

The firm relocation analysis, overall, shows that areas within close proximity to rail stations have not consistently attracted relocating firms, except in the case of areas near the mature rail stations that were opened before 1990. Similar trend was found earlier for the firm retainment analysis in Chapter 4. The inconsistency in net firm relocation near rail stations raises the question of what policymakers should do differently to attract relocating firms to select areas near the rail stations. Areas near the rail stations do not appear to experience positive net relocation, at least in the short run, without proper zoning and land use regulations that make station areas a more desirable place for firms to relocate within. For more immediate results, policymakers advocating for transit-oriented development should be more proactive in focusing development around transit stations, as discussed in Chapter 4.

CHAPTER 6: SUMMARY, CONCLUSIONS, AND POLICY IMPLICATIONS

Rail transit network plays a critical role in the U.S. transportation system and economy. Investments in rail transit systems contribute to: (1) improved overall efficiency of transportation systems, (2) environmental sustainability, (3) reducing automobile dependence and congestion, and (4) promoting economic development. While the first three influences of rail investments are well-documented, the contribution of these investments to economic development is less understood and has recently attracted close attentions by transportation scholars and economists, as well as local officials and planners.

The objective of this dissertation research is to examine the magnitude of impact that the close proximity to the passenger rail station has on firm dynamics, controlling for other influential factors. The central question this dissertation answers is—how transit investments influence the spatial dynamics of economic activities? In other words, the study examines the spatial variation of three patterns of firm dynamics—firm birth, firm closure, and firm relocation patterns—in relation to areas within a short walking distance to the passenger rail stations in the State of Maryland.

This dissertation uses standard negative binomial and propensity-score-weighted negative binomial regression methods to analyze the dependent variables of firm dynamics constructed from the National Establishment Time Series (NETS) panel data for the period from 1990 to 2010. In particular, this study considers six important research design aspects when examining the association between the patterns of firm dynamics and proximity to passenger rail stations:

1) maturity of passenger rail stations (i.e. the analysis considers the opening date of the passenger rail stations);
2) distance-to-station threshold (i.e. whether the examined areas are within a *quarter mile, quarter to half mile*, or *half to one mile* distance from rail stations);
3) firm size category in terms of number of employees (i.e. whether a firm has a sole proprietor, more than one employee, five or less employees, or more than five employees);
4) disaggregation by industry sector (i.e. whether a firm is categorized as retail, professional service, FIRE, health, administrative, or manufacturing firm based on NAICS classification); and
5) the method used for the statistically controlled analysis, specifically, whether or not the analysis controls for the endogeneity of the treatment (i.e. the placement of passenger rail stations).
6) the choice of absolute numbers of firms as dependent variables, as compared to the proportion of firms relative to existing firms or relative to the size of labor force in the past studies on firm dynamics. Negative binomial regressions are applied to the panel data set as an appropriate method for count dependent variables.

In addition to taking these six research design aspects into account, the analyses control for relevant factors related to transportation, agglomeration, and socio-economic characteristics that influence the patterns of firm dynamics, identified through a review of relevant empirical and theoretical literature (see Chapter 2). Several inferences can be drawn from the collection of results presented in Chapter 4 and Chapter 5.

When other factors are controlled for, higher rates of firm births (or startups) and relocating firms have located within a short walking distance to the passenger rail stations regardless of the differences in the level of maturity of the stations. The level of maturity of the rail stations impacted the magnitude but not the direction of influence. That is, areas within a short walking distance to *more recently opened stations* attracted higher numbers of smaller startups and relocating firms (with five or fewer employees) than areas within a short walking distance to the mature stations that were *opened before 1990*. The mature rail stations were more likely to attract larger firms (with more than five employees) than stations that were opened after 1990. Because of the improved level of market accessibility, more recent rail stations tend to influence the spatial setting for smaller firms more substantially compared to mature rail stations that have a higher influence on the spatial setting for larger firms. Evidently, there is a strong effect when new transit stations are introduced to a less developed site in terms of jobs, and thus, locational decisions of new and relocating firms tend to be unconstrained by the existing employment densities.

Although most of the past studies on firm dynamics examined firm birth and inward relocation (i.e. positive impacts), two other types of firm dynamics should be examined to get a more comprehensive understanding of firm dynamics in relation to proximity to passenger rail stations; this dissertation provides deeper analysis to examine firm closure and outward relocation as well (i.e. negative impacts). By comparing the positive effects of firm dynamics (firm birth and inward relocation) to the negative predicted impacts (firm closure and outward relocation), it was possible to estimate the combined probability of firm retainment and net relocation for areas near the passenger rail stations.

In the period between 1990 and 2010, there has been inconsistent growth in urban density near the passenger rail stations in the State of Maryland. The results in this dissertation suggest that areas near the passenger rail stations have belated positive economic impacts, shown by positive probabilities of firm retainment and net relocation around the mature rail stations that were opened before 1990 (see Figure 18 in Chapter 4 and Figure 36 in Chapter 5).[1] In comparison, areas near the less mature stations that were opened after 1990 had predominantly negative probabilities of firm retainment and net relocation, compared to the rest of the study area (see Appendix C). Clearly, the State of Maryland lacked deliberate planning to encourage urban densification near rail stations, and some regulations may have actively discouraged densification near the stations in favor of continuous suburbanization.

The industry-specific analysis of this research shows evidence that areas within a short walking distance to the mature passenger rail stations experienced a positive probability of firm retainment in the case of firms belonging to the retail, FIRE, and professional services sectors (see Figure 19 in Chapter 4). For relocating firms, areas near

[1] This dissertation calculates the probability of firm retainment by subtracting the probability of firm closure from the probability of firm birth. Similarly, the probability of net firm relocation equals the difference between the probability of inward relocation and the probability of outward relocation.

the mature rail stations experienced a positive net firm relocation in the case of firms belonging to the FIRE and administrative sectors (see Figure 37 in Chapter 5). This research finds no positive probabilities of firm retainment and net firm relocation in areas around the rail stations opened after 1990 (see Appendix C and D). Therefore, policymakers should introduce proactive policies that encourage businesses belonging to these sectors to locate near the passenger rail stations, lest these areas face extended delays to realizing development potentials.

The firm birth and inward relocation results by industry sector suggest that industries are more likely to locate in specialized economic environments to share a common pool of specialized workers. These results may lead one to wrongly conclude that a more specialized economic environment at the local level is a more suitable strategy to guide development. The reality is completely the opposite when considering all the four spatial aspects of firm dynamics (i.e. firm birth, closure, inward relocation, and outward relocation). This research provides empirical evidence that urbanization economies lead to higher probabilities of firm retainment than localization economies at a micro-level. In a Census block, the number of firms in own-industry (localization economies) negatively influences the probability of firm retainment, while the total number of firms (urbanization economies) has a positive influence (see Figure 16 in Chapter 4). Therefore, a more diverse economic environment around rail stations can lead to higher probability of firm retainment compared to a more specialized economic environment.

Both a more diverse economic environment and higher population densities can lead to higher probability of firm retainment. The results of this research on population density appeared contradictory at first. The results show that at a micro-level, population density may not necessarily have a positive association with the positive effects of firm dynamics—firm birth and inward relocation. The micro analysis carried out recently by Chatman and Noland (2016) on firm birth also presented some evidence of a negative association between population density and the rate of firm birth in Portland, Oregon. However, the analysis conducted in this dissertation shows that population density has an overall positive influence on the firms' economic activities at a very local level, taking into account all patterns of firm dynamics. Population density positively influences the probability of firm retainment and net relocation at the Census block level (see Figure 13 in Chapter 4 and Figure 34 in Chapter 5). Not surprisingly, the impact of population density at the local level on the probability of firm retainment is higher for retailors compared to the other analyzed sectors (see Figure 16 in Chapter 4).

The inconsistency in firm retainment near rail stations raises the question of what policymakers should do differently to encourage transit-oriented development. After all, the densification of station areas (implied by positive firm retainment and net firm relocation) is what advocates of transit-oriented-development promote to bring about, including improved accessibility, reduced traffic congestion and air pollution due to modal shift, and increased walkability which accommodates more healthy and active lifestyles. Areas near the stations do not experience an increase in firm density, at least in the short term, without proper policies. If policymakers want to encourage transit-oriented development, they should realize that without proactive interventions decades may pass before any urban agglomeration occurs near the rail stations. For more immediate results, policymakers advocating for transit-oriented development should be more proactive in directing development near rail stations, by adopting policies such as removing minimum parking requirements and even setting maximum parking caps, and by promoting higher residential density (Cervero, 2008), more mixed land use (Cervero and Duncan, 2006), better street connectivity and landscape (Cervero, 2007), and other development that provides locations for people's social and economic activities.

The tension between the two types of agglomeration economies (i.e. localization and urbanization economies) at a micro-level is another key factor that policymakers should consider when directing policies towards transit-oriented development. Policymakers should also pay close attention to the linkage between firms, industrial sectors, and rail station areas. Certain industry sectors may gain more from the accessibility benefits provided by rail station, such as access to a large pool of workforce and customers, compared to other sectors. For any transit-oriented development to emerge and thrive, policymakers must proactively encourage both diverse economic activities and residential development around passenger rail stations. This is certainly facilitated not only by mixed land use development in each station area, but could be also facilitated among a few stations close to each other along a transit line. In other words, while one station has a more focus on residential development, the nearby stations have more employment. As a group of stations in proximity, these stations can work as a transit-oriented development.

While the analysis in this dissertation would have ideally included data related to changes in land use/land cover, such data is not always easily available. Spatial-temporal data on land use better depicts changes in local policies over time. Land use related covariates would also allow for more robust propensity score matching analysis, given that planners locate transit stations depending on the existing land use patterns.

Future studies can take a step further when analyzing the connection between rail transit and firm dynamics to include factors related to the level of service and the physical characteristics of railway lines and stations. Differences in the speed and the frequency of passenger rail services may have varying influences on the patterns of firm dynamics. Certain physical characteristics of transit stations can also be more appealing in the location decisions of certain firms compared with others. Moreover, qualitative research on factors influencing the decisions of startups and relocating firms to locate near the passenger rail stations would provide a deeper understanding of the influence passenger rail systems have on firm dynamics.

Appendices

Appendix A: The geographic accuracy of the latitudes and longitudes of the NETS dataset

The level at which the NETS dataset provided the Longitude and latitude of firm location

Field Name	Data Type	Size	Description
Latitude	Decimal (8,4)	8	Establishment Latitude
Longitude	Decimal (8,4)	8	Establishment Longitude
LevelCode	Text	1	Level at which latitude/longitude provided (D = Block Face, B = Block Group, T = Census Tract Centroid, Z = ZIP Code Centroid, N = Not Coded, S = Street Level)
OriginLatitude	Decimal (8,4)	8	Latitude of Move Origin
OriginLongitude	Decimal (8,4)	8	Longitude of Move Origin
OriginLevelCode	Text	1	Level at which origin latitude/longitude provided (D = Block Face, B = Block Group, T = Census Tract Centroid, Z = ZIP Code Centroid, N = Not Coded, S = Street Level)
DestLatitude	Decimal (8,4)	8	Latitude of Move Destination
DestLongitude	Decimal (8,4)	8	Longitude of Move Destination
DestLevelCode	Text	1	Level at which destination latitude/longitude provided (D = Block Face, B = Block Group, T = Census Tract Centroid, Z = ZIP Code Centroid, N = Not Coded, S = Street Level)

Source: Walls & Associates, NETS Database: 2012 Database Description.

The variable "LevelCode" highlighted in the table above shows the geographic level that NETS data recorded the longitude and latitude of the location of firms. For relocating firms, NETS data provided the longitude and latitude of the location of origin and destination also at various level as highlighted in the table. The table below shows the level code of the longitude and latitude of the location of firms within the study area by year. Clearly, NETS data provided most of the latitude and longitude of firm locations at the Census block level (D).

	levelcode					
year	B	D	S	T	Z	Total
1990	171	78,777	123	150	32,276	111,497
1991	181	83,065	128	164	32,431	115,969
1992	189	86,376	138	172	30,301	117,176
1993	201	98,757	161	206	35,146	134,471
1994	216	104,434	163	209	35,337	140,359
1995	241	114,034	188	226	36,730	151,419
1996	253	118,653	197	237	34,239	153,579
1997	270	127,287	218	263	33,521	161,559
1998	289	134,770	235	274	33,359	168,927
1999	298	136,249	246	282	29,382	166,457
2000	307	138,463	260	290	24,399	163,719
2001	315	148,026	279	300	22,930	171,850
2002	457	176,556	326	348	23,368	201,055
2003	601	196,504	341	420	18,749	216,615
2004	629	202,556	361	453	12,867	216,866
2005	565	209,785	383	451	6,849	218,033
2006	337	223,653	397	373	6,431	231,191
2007	264	235,858	407	322	4,968	241,819
2008	265	257,221	444	269	5,616	263,815
2009	267	286,686	483	247	5,797	293,480
2010	219	265,226	538	156	4,283	270,422
Total	6,535	3,422,936	6,016	5,812	468,979	3,910,278

Appendix B: The steps to calculate the Inverse Probability of Treatment (IPT) weights and including them in the negative binomial regression.

Following are the steps undertaken to calculate the Inverse Probability of Treatment (IPT) weights that are used to adjust the NB regression:

1. **Running a logistic regression**

The study uses logistic regression to calculate the propensity scores. The Stata commands used to calculate the propensity scores are as follow:

logistic t x1 x2 …. xn

predict propensity

where x1 to xn are the covariates that determine the value of the propensity scores; and t is the treatment dummy variable (treatment=1 if Census blocks are within one-mile distance from rail stations; treatment=0 if the blocks are more than one-mile away from the stations). The treatment is restricted to passenger rail stations that were opened between 1990 and 1998 (all the 43 group-B stations). Census blocks within one-mile from the other stations (stations opened before 1990 or after 2000) are omitted from the PS-weighted analysis. Below are the results from the logistic regression used to calculate the propensity scores, followed by goodness of fit test and the PS distribution.

The results of logistic regression to calculate propensity scores are presented in the regression table below. The initial goodness-of-fit test showed that the logistic regression does not fit the data until an interaction term between two covariates was added to the logistic regression (the two covariates are housing rent and unemployment rate). The interactions between all of the covariates were tested to determine which interaction to be added to the logistic regression (Lunt, 2014). The covariates used in the PS logistic regression are: population and employment densities, household income, unemployment rate, percent of population that are college graduate, percent of the population that are African American, housing rent, and distance to the nearest highway exit. The only interaction term added to the logistic regression (as shown in the table below) is the interaction between unemployment rate and housing rent covariates.

```
Logistic regression                               Number of obs     =     33,953
                                                  LR chi2(9)        =     769.34
                                                  Prob > chi2       =     0.0000
Log likelihood =  -1838.044                       Pseudo R2         =     0.1731
```

b_one_mi_B	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
population_density	1.097182	.0341401	2.98	0.003	1.032269	1.166178
employment_density	.892292	.0516094	-1.97	0.049	.7966623	.9994009
HH_income	.9320958	.0049613	-13.21	0.000	.9224225	.9418706
unemployment_rate	1.21e-18	1.03e-17	-4.87	0.000	7.47e-26	1.97e-11
percent_college_edu	94.32743	33.65915	12.74	0.000	46.87067	189.8344
percent_black	.4377389	.1219774	-2.96	0.003	.2535275	.7557971
housing_rent	154.1104	75.09762	10.34	0.000	59.29848	400.5165
distance_to_highway	.6333226	.0355322	-8.14	0.000	.5673732	.7069378
rent_x_unemp	7.81e+23	7.88e+24	5.45	0.000	2.02e+15	3.02e+32
_cons	.0032808	.0011423	-16.43	0.000	.0016582	.0064914

2. Testing goodness of fit of the PS logistic model

The Hosmer-Lemeshow test is used to determine the goodness of fit of the propensity score logistic model. Data is first regrouped by ordering the predicted probabilities and forming the number of groups. The formula of HL test is in the following form:

$$G^2_{HL} = \sum_{j=1}^{10} \frac{(O_j - E_j)^2}{E_j(1 - E_j/n_j)} \sim \chi^2_8$$

Where: χ^2 = chi squared, nj = number of observations in the jth group, Oj = number of observed cases in the jth group, and Oj = number of expected cases in the jth group. Small p-values of the LH test mean that the model is a poor fit.

The HL goodness-of-fit test below determines whether the predicted probabilities deviate from the observed probabilities. The goodness of fit test below (of the improved propensity model) shows that the p-value for the goodness-of-fit test is lower than the significance level (95%), which means that the predicted probabilities do not deviate from the observed probabilities hence a best fit of the PS model is achieved.

Logistic model for b_one_mi_B, goodness-of-fit test

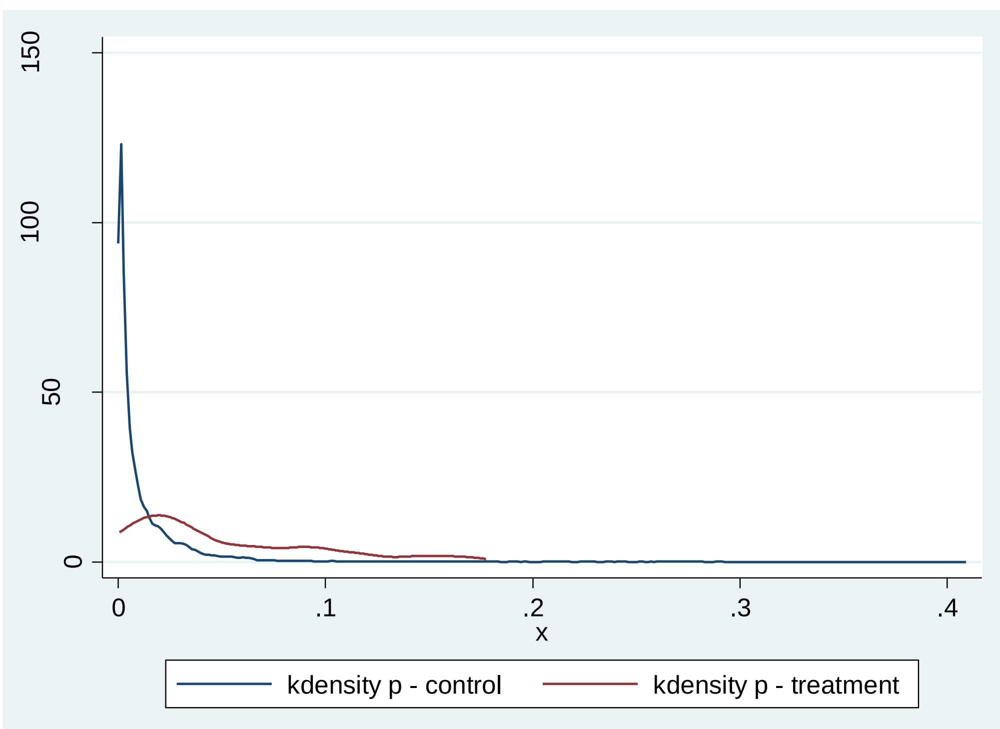
(Table collapsed on quantiles of estimated probabilities)

Group	Prob	Obs_1	Exp_1	Obs_0	Exp_0	Total
1	0.0003	0	0.3	3396	3395.7	3396
2	0.0008	4	1.7	3391	3393.3	3395
3	0.0016	3	4.2	3392	3390.8	3395
4	0.0026	9	7.2	3387	3388.8	3396
5	0.0041	8	11.3	3387	3383.7	3395
6	0.0068	10	18.3	3385	3376.7	3395
7	0.0108	37	29.2	3359	3366.8	3396
8	0.0181	40	47.8	3355	3347.2	3395
9	0.0313	90	80.6	3305	3314.4	3395
10	0.4089	210	210.5	3185	3184.5	3395

```
      number of observations =     33953
            number of groups =        10
     Hosmer-Lemeshow chi2(8) =        13.33
                 Prob > chi2 =         0.1010
```

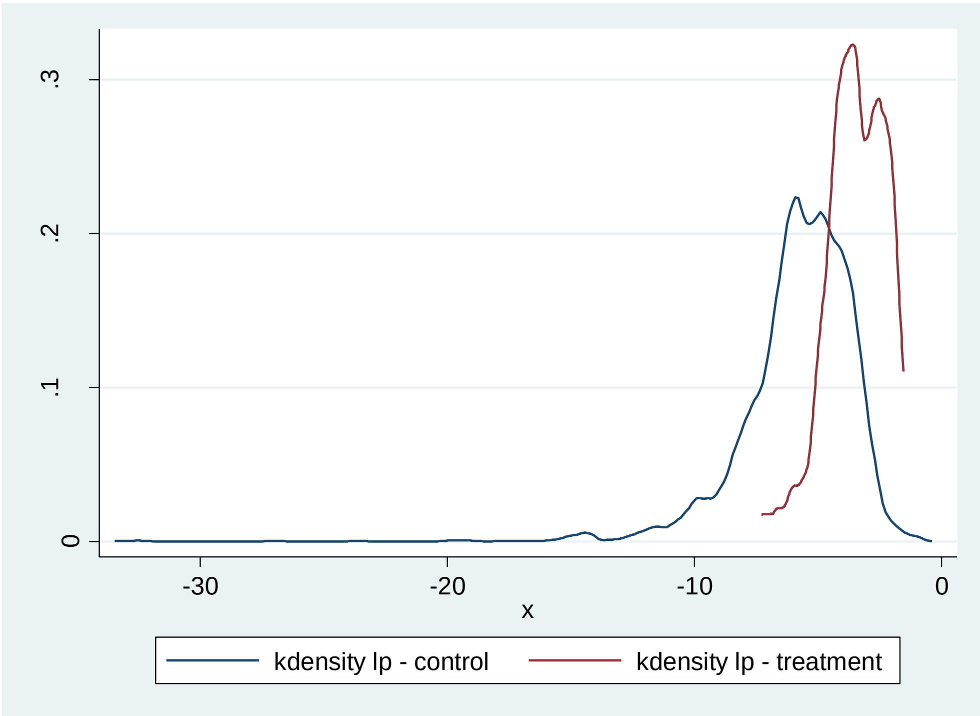
The graph below shows the distribution of the propensity score in the treated and the untreated Census blocks (blocks within one mile of the stations).

The Distribution of Propensity Score



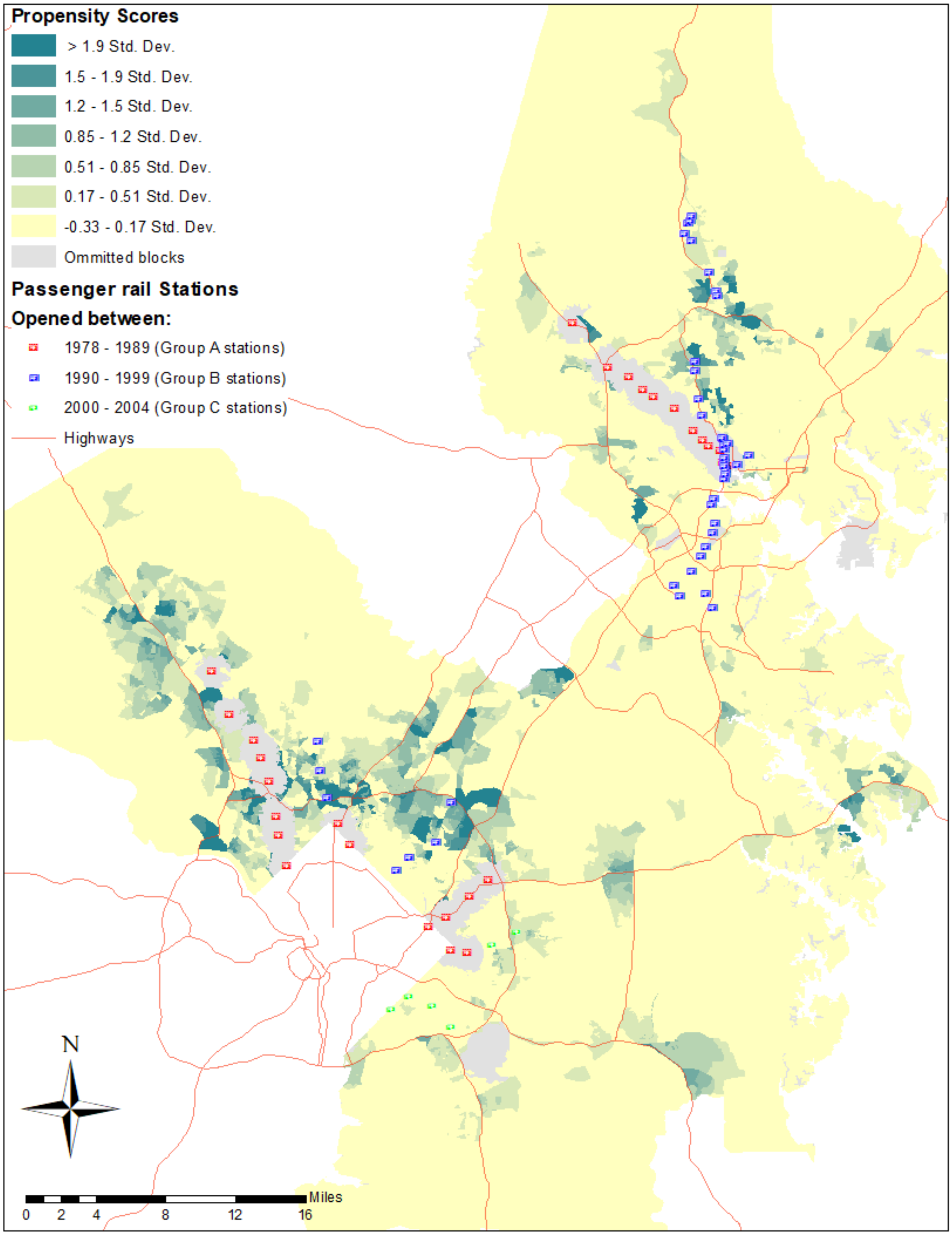
It is recommended, however, to use the log of the odds of the propensity score (also called the linear predictor), rather than the propensity score shown above, since it tends to be more normally distributed (Lunt, 2014). The graph below shows the log odd of the propensity scores used in this study to calculate the regression weights of the PS-weighted negative binomial model.

The Distribution of log odds of Propensity Score



The map below illustrates the estimated propensity scores of the logistic regression explained above.

Map showing the estimated propensity scores for each Census block of the study area except blocks located within a mile radius of group A and C



The Stata program “*propwt*” is used to create the IPT weights. The Stata command is as follow:

propwt *t propensity*, ipt

where *t* is the treatment variable (i.e. the dummy variable indicating whether or not a Census block is within a mile of group B rail stations); *propensity* is the propensity scores variable generated earlier by the logistic regression; and ipt is the calculated inverse probability of treatment.

In order to use the IPT weights in the analysis, they are specified as part of the Stata regression command of the negative binomial model “*xtnbreg*” by adding the syntax [*pweight=ipt*] to the command before any options:

***xtnbreg* outcome covariates [*pweight*=*ipt*], pa**

Note that the population-averaged (pa) option is added to the end of the negative binomial regression model that includes the IPT weights because the random-effect option does not permit for the inclusion of regression weights. The difference between the random-effects and population-averaged estimators are very subtle, however. For continuous or count outcomes, the two approaches are nearly identical. Differences emerge between the two specifications only when analyzing binary outcomes (Hilbe, 2011). The results remained the same when a standard negative binomial model (without weights) was tested using both estimators (population-averaged and random-effects).

A brief explanation of the subtle difference between random-effects and population-average estimators is shown below. See Hilbe (2011) for more detail on the difference between the two estimators.

Random-effects estimators fit the following model

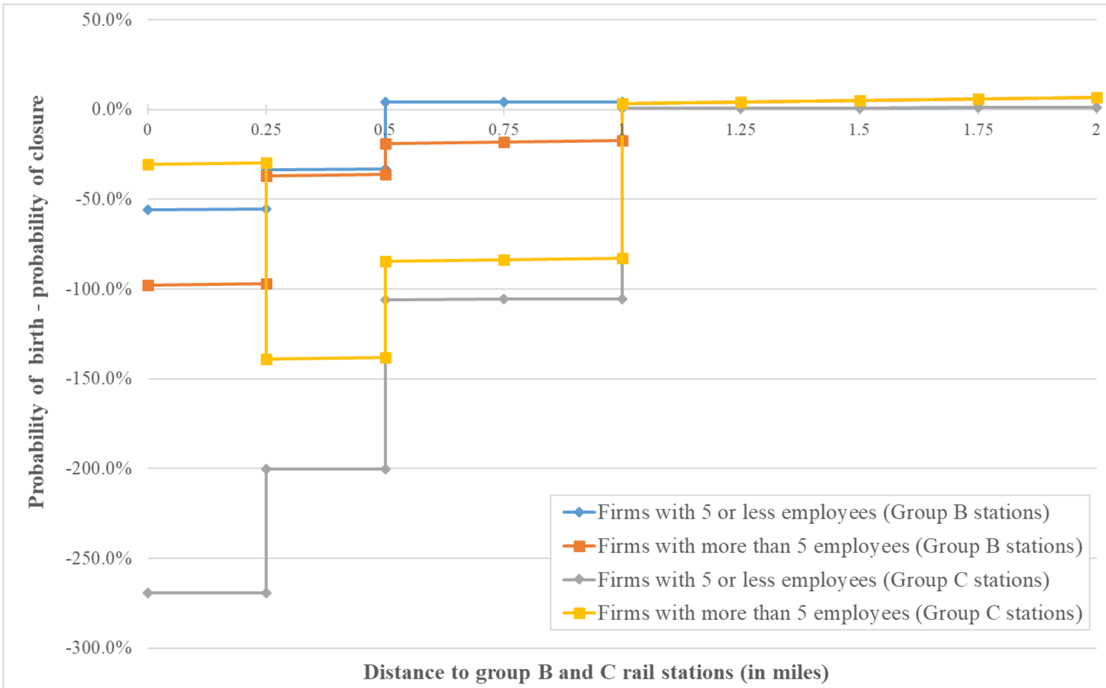
Pr(Yij=1 | Xij, ui) = F(Xij b + ui)

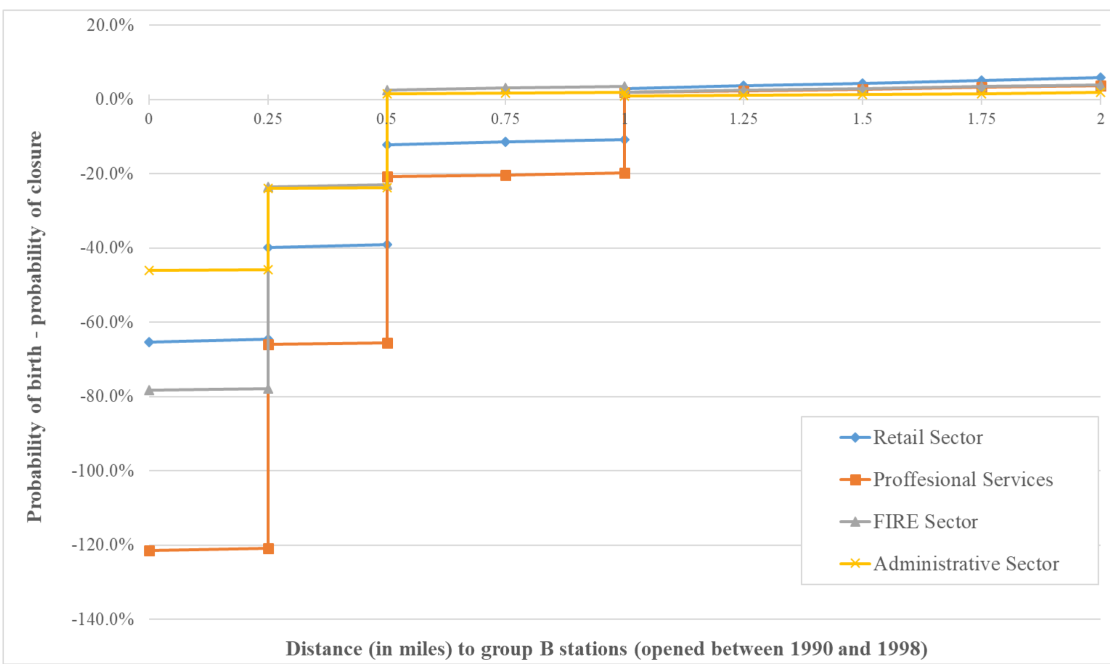
whereas population-average estimators fit the following model:

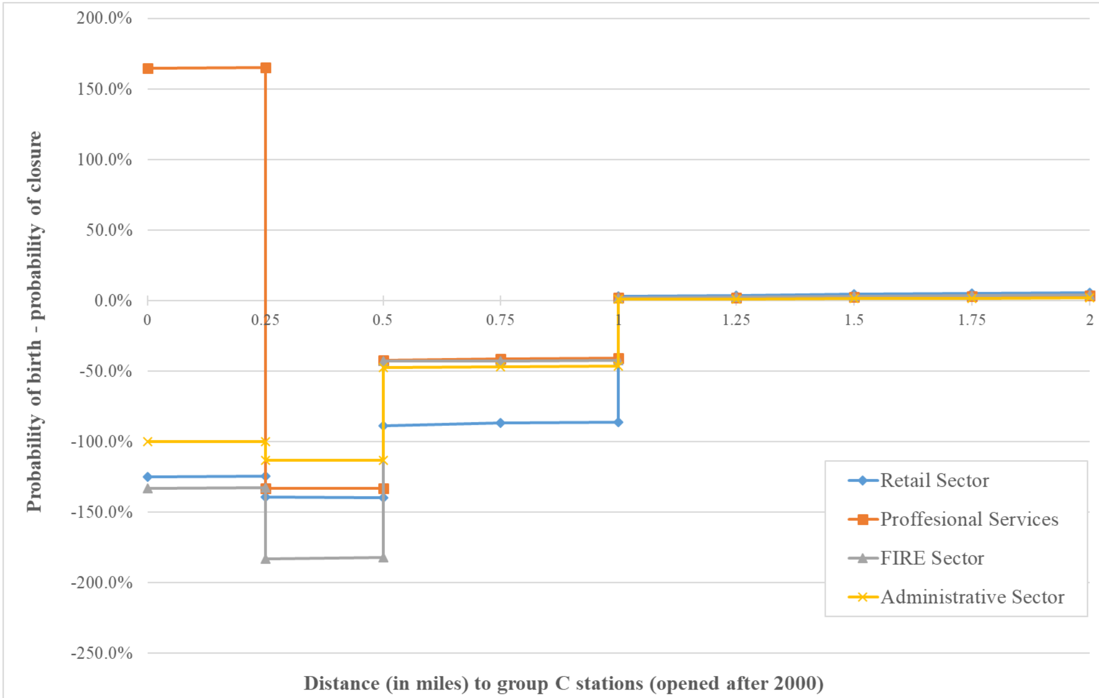
Pr(Yij=1 | Xij) = G(Xij b*)

The subtle point is that b and b* are different population parameters. Even though the estimators appear to be estimating different things, in practice, however, b and b* are often very close. The population-averaged model does not fully specify the distribution of the population but rather specifies a marginal distribution. The random-effect model, on the other hand, fully specify the distribution (ui is given a distribution), which allows the mean of the dependent variable to vary across the subjects.

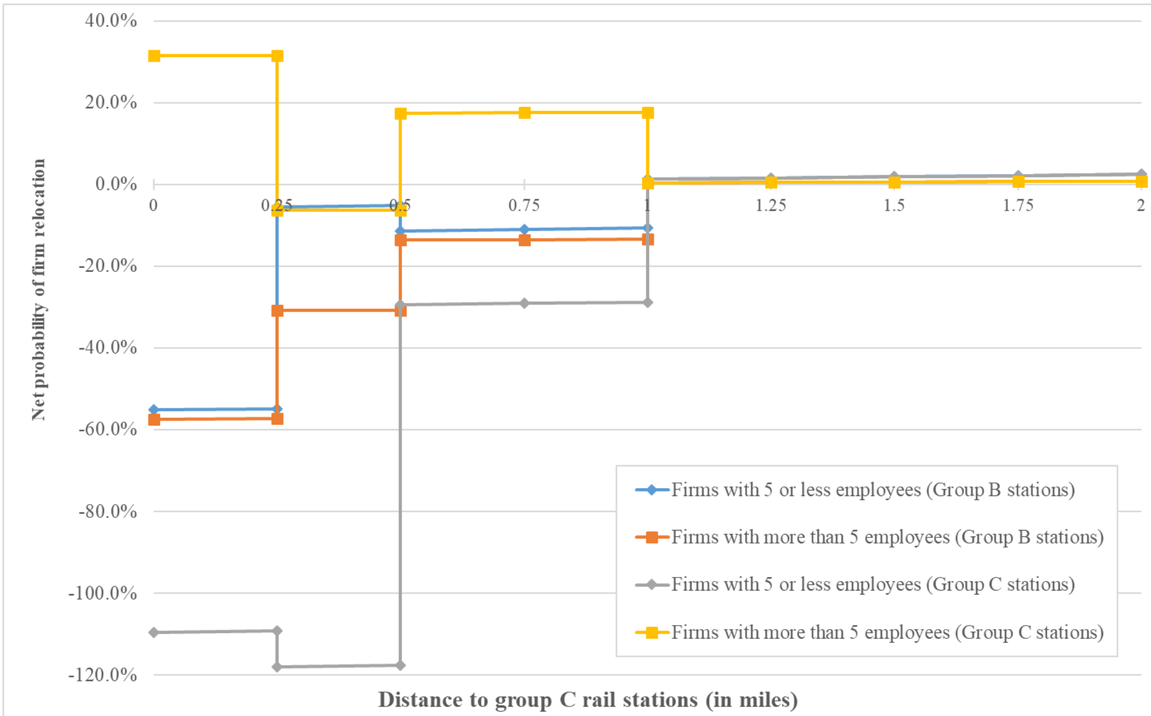
Appendix C: Firm-retainment predicted-effects of station distance variables for group B (opened between 1990 and 1998) and group C stations (opened after 2000), by firm size and industry sector

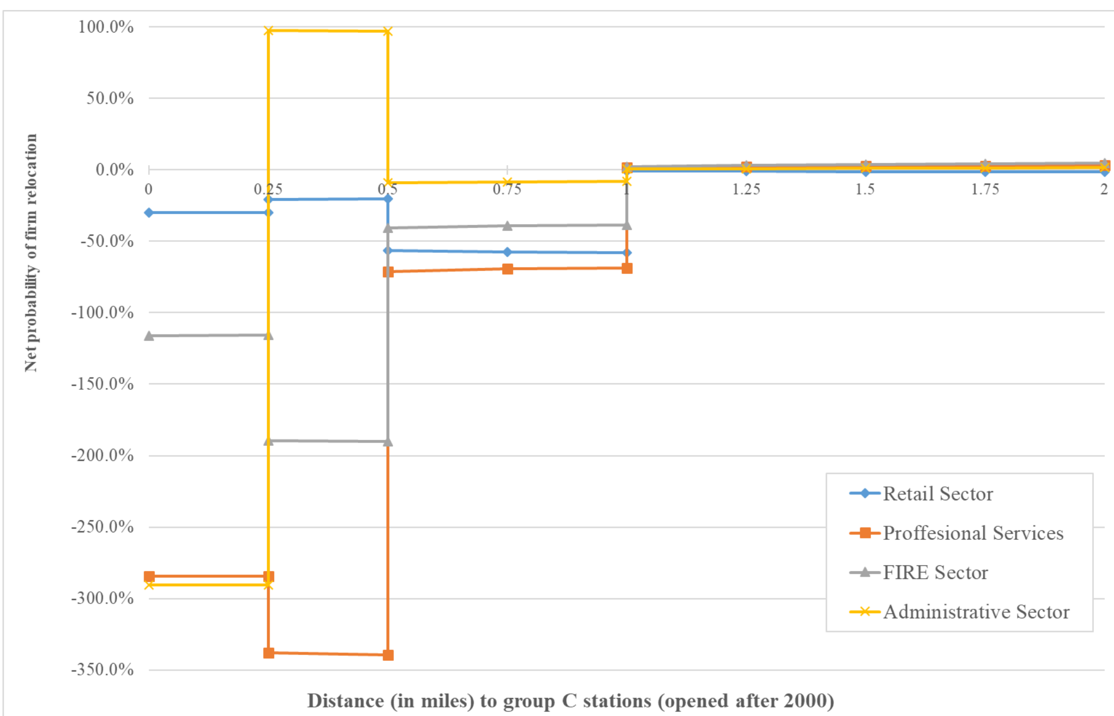





200.0%
150.0%
100.0%
50.0%
0.0%
-50.0%
-100.0%
-150.0%
-200.0%
-250.0%
0
0.25
0.5
0.75
1
1.25
1.5
1.75
2
Probability of birth - probability of closure
Distance (in miles) to group C stations (opened after 2000)
Retail Sector
Proffesional Services
FIRE Sector
Administrative Sector

Appendix D: Net Predicted effects of firm relocation of station distance variables for group B (opened between 1990 and 1998) and group C stations (opened after 2000), by firm size and industry sector





References

Ács, Zoltán J., and Catherine. Armington. 2006. Entrepreneurship, Geography, and American Economic Growth. Cambridge: Cambridge University Press.

American Public Transportation Association. 2017. APTA... 2016 Public Transportation Fact Book. American Public Transportation Association.

An, Youngsoo, Yeonggyeong Kang, and Seungil Lee. 2014. "A Study on the Impact of Soft Location Factors in the Relocation of Service and Manufacturing Firms." International Journal of Urban Sciences 18 (3): 327-339.

Anas, A., R. Arnott, and K. A. Small. 1998. "Urban Spatial Structure." JOURNAL of ECONOMIC LITERATURE 36 (3): 1426-1464.

Audretsch, David B, and Maryann P Feldman. 1996. “R&d Spillovers and the Geography of Innovation and Production.” The American Economic Review 86 (3): 630–40.

Audretsch, David B, and Michael Fritsch. 1994a. "On the Measurement of Entry Rates." Empirica: Journal of Applied Economics and Economic Policy - Sponsored by the Austrian Economic Association and the Austrian Institute of Economic Research 21 (1): 105-113.

Audretsch, David B, and Michael Fritsch. 1994b. "The Geography of Firm Births in Germany." Regional Studies 28 (4): 359-365.

Austin, Peter C. 2011. “An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.” Multivariate Behavioral Research 46 (3): 399–424.

Bacher, Hans Ulrich, and Marius Brülhart. 2013. "Progressive Taxes and Firm Births." International Tax and Public Finance 20 (1): 129-168.

Bang, Heejung, and James M Robins. 2005. “Doubly Robust Estimation in Missing Data and Causal Inference Models.” Biometrics 61 (4): 962–73.

Barnatchez, Keith, Crane Leland, and Decker Ryan. 2017. "An Assessment of the National Establishment Time Series (NETS) Database," Board of Governors of the Federal Reserve System (U.S.) - Finance and Economics Discussion Series 2017-110.

Beaudry, Catherine, and Andrea Schiffauerova. 2009. "Who's Right, Marshall or Jacobs? The Localization versus Urbanization Debate." Research Policy 38 (2): 318-337.

Bell, Andrew, and Kelvyn Jones. 2015. “Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data.” Political Science Research and Methods 3 (01): 133–53.

Bell, Douglas Adrian. 1991. "Office Location — City or Suburbs? Travel Impacts Arising from Office Relocation from City to Suburbs." Transportation: Planning - Policy - Research - Practice 18 (3): 239-259

Belzer, Dena, Sujata Srivastava, and Mason Austin. 2011. Transit and Regional Economic Development. Oakland, CA: Center for Transit-Oriented Development

Bowes, David, and Keith Ihlanfeldt. 2001. "Identifying the Impacts of Rail Transit Stations on Residential Property Values." Journal of Urban Economics 50 (1): 1-25.

Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, and Stürmer T. 2006. “Variable Selection for Propensity Score Models.” American Journal of Epidemiology 163 (12): 1149–56.

Browning, Jon E. 1980. How to Select a Business Site: The Executive's Location Guide. Toronto,on: Mcgraw-Hill.

Brueckner, Jan K. 2011. Lectures on Urban Economics. Cambridge, Mass.: MIT Press.

Cainelli, Giulio, Sandro Montresor, and Giuseppe Vittucci Marzetti. 2014. “Spatial Agglomeration and Firm Exit: A Spatial Dynamic Analysis for Italian Provinces.” Small Business Economics : An Entrepreneurship Journal 43 (1): 213–28.

Cervero, Robert. 2007. “Transit-Oriented Development's Ridership Bonus: A Product of Self-Selection and Public Policies.” Environment and Planning A 39 (9): 2068–85.

Cervero, Robert. 2008. “Effects of Tod on Housing, Parking, and Travel”. Tcrp Report, 128.

Cervero, Robert, and Michael Duncan. 2002. “Transit's Value-Added Effects: Light and Commuter Rail Services and Commercial Land Values.” Transportation Research Record 1805 (1805): 8–15.

Cervero, Robert, and Michael Duncan. 2006. “'Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing?” Journal of the American Planning Association 72 (4): 475–90.

Cervero, Robert, National Research Council (U.S.), Transit Cooperative Research Program, United States, and Transit Development Corporation. 2004. Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects. TCRP report, 102; Report (Transit Cooperative Research Program), 102. Washington, D.C.: Transportation Research Board. http://trb.org/publications/tcrp/tcrp_rpt_102.pdf.

Chatman, Daniel G., and Robert B. Noland. 2011. "Do Public Transport Improvements Increase Agglomeration Economies? A Review of Literature and an Agenda for Research." Transport Reviews 31 (6): 725-742.

Chatman, Daniel G., and Robert B. Noland. 2014. "Transit Service, Physical Agglomeration and Productivity in US Metropolitan Areas." Urban Studies 51 (5): 917-937.

Chatman, Daniel G., Robert B. Noland, and Klein N. 2016. "Firm Births, Access to Transit, and Agglomeration in Portland, Oregon, and Dallas, Texas" Transportation Research Record: Journal of the Transportation Research Board v2598 (201601): 1-10.

Daniel, Mann, Johnson & Mendenhall, Kaiser Engineers, and Baltimore Metropolitan Area Mass Transit Development Steering Committee. 1968. Baltimore Region Rapid Transit System: Feasibility and Preliminary Engineering. Baltimore, Maryland.

Debrezion, Ghebreegziabiher, Eric Pels, and Piet Rietveld. 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): 161-180.

Duncan, Michael. 2011. The Impact of Transit-Oriented Development on Housing Prices in San Diego, CA. Urban Studies 48(1): 101-127.

Fotopoulos, Georgios, and Nigel Spence. 1999. "Spatial Variations in New Manufacturing Plant Openings: Some Empirical Evidence from Greece." Regional Studies 33 (3): 219-229.

Freedman DA, and Berk RA. 2008. “Weighting Regressions by Propensity Scores.” Evaluation Review 32 (4): 392–409.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. The Spatial Economy: Cities, Regions and International Trade. Cambridge, MA: MIT Press.

Giuliano, G. 2004. “Land Use Impacts of Transportation Investments: Highway and Transit.” In: S. Hanson and G. Giuliano, Editors, The Geography of Urban Transportation (third Ed.).

Green, Rodney D., and David M. James. 1993. Rail Transit Station Area Development: Small Area Modeling in Washington, D.C. Armonk, N.Y.: M.E. Sharpe.

Hade EM, and Lu B. 2014. “Bias Associated with Using the Estimated Propensity Score as a Regression Covariate.” Statistics in Medicine 33 (1): 74–87.

Haltiwanger, John C., Lisa M. Lynch, Christopher D. Mackie, and National Research Council (U.S.). 2007. Understanding Business Dynamics: An Integrated Data System for America's Future. Washington, D.C.: National Academies Press.

Hausman, Jerry, Bronwyn H Hall, and Zvi Griliches. 1984. “Econometric Models for Count Data with an Application to the Patents-R & D Relationship.” Econometrica 52 (4): 909–38.

Hayter, Roger. 1997. The Dynamics of Industrial Location: The Factory, the Firm, and the Production System. Chichester: Wiley.

Hess, Daniel Baldwin, and Tangerine Maria Almeida. 2007. “Impact of Proximity to Light Rail Rapid Transit on Station-Area Property Values in Buffalo, New York.” Urban Studies 44 (5-6): 1041–68.

Hilbe, Joseph M, and William H Greene. 2007. “Handbook of Statistics.” Essay. In 7 Count Response Regression Models, 210–52. Elsevier Science & Technology. doi:10.1016/S0169-7161(07)27007-5.

Hilbe, Joseph M. 2011. Negative Binomial Regression. Second edition. Cambridge, UK: Cambridge University Press.

Holl, Adelheid. 2004a. "Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain." Regional Science and Urban Economics 34 (3): 341-363.

Holl, Adelheid. 2004b. "Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal*." Journal of Regional Science 44 (4): 693-712.

Holl, Adelheid. 2004c. "Start-Ups and Relocations: Manufacturing Plant Location in Portugal." Papers in Regional Science: The Journal of the Regional Science Association International 83 (4): 649-668.

Holl, Adelheid. 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.

Jacovidis, Jessica N, Kelly J Foelber, and S. Jeanne Horst. 2016. “The Effect of Propensity Score Matching Method on the Quantity and Quality of Matches.” The Journal of Experimental Education 9 (9): 1–24.

Jacobs, Jane. 1969. The Economy of Cities. New York: Random House.

Joffe, Marshall M, Thomas R Ten Have, Harold I Feldman, and Stephen E Kimmel. 2004. “Model Selection, Confounder Control, and Marginal Structural Models.” The American Statistician 58 (4): 272–79.

Johnson, Peter, and Simon Parker. 1996. "Spatial Variations in the Determinants and Effects of Firm Births and Deaths." Regional Studies 30 (7): 679-688.

Jostarndt, Philipp., and Bernd. Rudolph. 2007. Financial Distress, Corporate Restructuring and Firm Survival: An Empirical Analysis of German Panel Data.

Kilkenny, Maureen, and Thisse Jacques-François. 1999. “Economics of Location: A Selective Survey.” Computers and Operations Research 26 (14): 1369–94.

Knapp, Gerrit-Jan, Brown, Aviva Hopkins, and Lewis, Rebecca. 2015. "Rezoning Montgomery: People, Politics and Place." In Emily Talen, ed., Retrofitting Sprawl: Addressing Seventy Years of Failed Urban Form: 57-79.

Kronenberg, Kristin. 2013. "Firm Relocations in the Netherlands: Why Do Firms Move, and Where Do They Go?" Papers in Regional Science 92 (4): 691-713.

Krugman, Paul. 2011. "The New Economic Geography, Now Middle-Aged." Regional Studies 45 (1): 1-7.

Lasch, Frank, Frank Robert, and Frédéric Le Roy. 2013. "Regional Determinants of ICT New Firm Formation." Small Business Economics: An Entrepreneurship Journal 40 (3): 671-686.

Leuven, E. and Sianesi, B. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html

Liu Chao, Sevgi Erdogan, Ting Ma, and Frederick W. Ducca. 2016. "How to Increase Rail Ridership in Maryland: Direct Ridership Models for Policy Guidance." Journal of Urban Planning and Development 142 (4).

Lunt, Mark. 2014. Propensity Analysis in Stata: Revision: 1.1.

Ma, Ting, Eli Knaap and Gerrit-Jan Knaap. 2014. "Retail Location and Transit: An Econometric Examination of Retail Location in Prince George's and Montgomery County, Maryland. National Center for Smart Growth (NCSG) Report.

Manson Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2017. IPUMS National Historical Geographic Information System: Version 12.0 [Database]. Minneapolis: University of Minnesota. http://doi.org/10.18128/D050.V12.0

Marshall, Alfred. 1890. Principles of Economics. V. 1. London: Macmillan.

Marshall, Alfred. 1964. Principles of Economics: An Introductory Volume. 8th ed. London: Macmillan.

Mathur, Shishir, and Christopher Ferrell. 2009. Effect of Suburban Transit Oriented Development on Residential Property Values. San Jose, CA: Mineta Transportation Institute.

Mejia-Dorantes L., and Lucas K. 2014. "Public transport investment and local regeneration: A comparison of London's Jubilee Line Extension and the Madrid Metrosur." Transport Policy. 35: 241-252.

Melo, P. C., D. J. Graham, and R. B. Noland. 2010. "Impact of Transport Infrastructure on Firm Formation: Evidence from Portuguese Municipalities." Transportation Research Record: Journal of the Transportation Research Board No. 2163, Transportation Research Board of the National Academies, Washington, D.C: 133–143.

Mohammad, Sara I., Daniel J. Graham, Patricia C. Melo, and Richard J. Anderson. 2013. "A Meta-Analysis of the Impact of Rail Projects on Land and Property Values." Transportation Research Part A 50: 158-170.

Morgan, Stephen L, and Jennifer J Todd. 2008. “A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects.” Sociological Methodology 38 (1): 231–81.

Murin, William J. 1971. Mass Transit Policy Planning: An Incremental Approach. Lexington, Mass.: Heath Lexington Books.

Nelson, Arthur C., Eskic Dejan, Ganning Joanna P., Hamidi Shima, Petheram Susan J., Liu Jenny H., and Ewing Reid. 2015. "Office Rent Premiums With Respect To Distance From Light Rail Transit Stations In Dallas And Denver". Urban Studies and Planning Faculty Publications and Presentations. 128.

Neumark, D., J. Zhang, and B. Wall. 2005. "Employment Dynamics and Business Relocation: New Evidence from The National Establishment Time Series." Working Paper Series, no. 11647: ALL.

Nguyen, Cao Y., Kazushi Sano, Tu Vu Tran, and Tan Thanh Doan. 2013. "Firm relocation patterns incorporating spatial interactions". The Annals of Regional Science. 50 (3): 685-703.

O'Sullivan, A. 2003. Where do firms locate? In: Urban Economics, 5th ed, pp. 65-91. Irwin, PA: McGraw-Hill.

Ottaviano, Gianmarco, Takatoshi Tabuchi, and Jacque. -F. Thisse. 2002. "Agglomeration and Trade Revisited," International Economic Review, 43, 101–127.

Pan, Qisheng. 2013. “The Impacts of an Urban Light Rail System on Residential Property Values: A Case Study of the Houston Metrorail Transit Line.” Transportation Planning and Technology 36 (2): 145–69.

Pan, Wei, and Bai Haiyan. 2015. Propensity Score Analysis: Fundamentals and Developments. New York: Guilford Press.

Peek, Gert-Joost, Luca Bertolini, and Hans De Jonge. 2006. "Gaining Insight in the Development Potential of Station Areas: A Decade of Node-Place Modelling in the Netherlands." Planning Practice and Research 21 (4): 443-462.

Pellenbarg, P.H., Wissen, L.J.G. van and van Dijk, Jouke. 2002. Firm relocation: state of the art and research prospects. No 02D31, Research Report, University of Groningen, Research Institute SOM (Systems, Organisations, and Management). http://EconPapers.repec.org/RePEc:gro:rugsom:02d31.

Płaziak, Monika, and Anna Irena Szymańska. 2014. “Role of Modern Factors in the Process of Choosing a Location of an Enterprise.” Procedia - Social and Behavioral Sciences 120: 72–83.

Porter, Michael E. 1990. The Competitive Advantage of Nations. New York: Free Press.

Posner M.A. and Ash A.S. 2012. Comparing Weighting Methods in Propensity Score Analysis. Mimeo, Villanova University.

Puga, Diego. 1999. ‘‘The Rise and Fall of Regional Inequalities, European Economic Review, 43, 303–334.

Renski, Henry. 2008. "New Firm Entry, Survival, and Growth in the United States: A Comparison of Urban, Suburban, and Rural Areas." Journal of the American Planning Association 75 (1): 60-77.

Reynolds, Paul. 1994. "Autonomous Firm Dynamics and Economic Growth in the United States, 1986–1990." Regional Studies 28 (4): 429-442.

Risselada, Anne, Veronique Schutjens, and Frank Van Oort. 2013. "Real Estate Determinants of Firm Relocation in Urban Residential Neighbourhoods." Tijdschrift Voor Economische En Sociale Geografie 104 (2): 136-158.

Robins JM, Hernán MA, and Brumback B. 2000. “Marginal Structural Models and Causal Inference in Epidemiology.” Epidemiology (Cambridge, Mass.) 11 (5): 550–60.

Rosenbaum, Paul R. 1987. “Model-Based Direct Adjustment.” Journal of the American Statistical Association 82 (398): 387–94.

Santos, A., McGuckin, N., Nakamoto, H.Y., Gray, D., and Liss S. 2011. Summary of Travel Trends: 2009 National Household Travel Survey. Washington, D.C.: U.S. Department of Trartation, Federal Highway Administration.

Schuetz, Jenny. 2015. "Do Rail Transit Stations Encourage Neighbourhood Retail Activity?" Urban Studies 52 (14): 2699-2723.

Schumpeter, Joseph. 1934. "The theory of economic development." Cambridge, MA: Harvard University Press.

Smart Location Database, U.S. Environmental Protection Agency. https://www.epa.gov/smartgrowth/smart-location-mapping#SLD

Sutaria, Vinod, and Donald A. Hicks. 2004. "New Firm Formation: Dynamics and Determinants." The Annals of Regional Science: An International Journal of Urban, Regional and Environmental Research and Policy 38 (2): 241-262.

Walls, Donald. 2008. “Understanding Data in the NETS Database,” NETS Documentation.

Wheaton, William C. 2004. “Commuting, Congestion, and Employment Dispersal in Cities with Mixed Land Use.” Journal of Urban Economics 55 (3): 417–38.