A Multilevel Analysis of the Contribution of Individual, Socioeconomic and Geographical Factors on Kindergarten Children's Developmental Health: A Saskatchewan Province-Wide Study

A Thesis Submitted to the College of Graduate Studies and Research

In Partial Fulfillment of the Requirements for the Degree of Master of Science

Collaborative Biostatistics Program, School of Public Health

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

By

Mohsen Soltanifar

March, 2014

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any materials in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Chair of Collaborative Biostatistics Program,

School of Public Health

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5E5

ABSTRACT

In current literature of child public health, a growing number of studies has been dedicated to early childhood development with a focus on child developmental health measured via the teacher completed Early Development Instrument (EDI). Using multilevel modeling as the optimal statistical method to analyze hierarchical EDI data, this study determines the strength of the effect and significance of predictors of children's 5 EDI outcomes, vulnerability, and the multiple vulnerability by taking into account the hierarchy present in its design. In addition, this study conducts an extensive epidemiological review of the risk factors associated with a child's developmental health at each level of the hierarchy, at cross-levels of the hierarchy and their variations across different levels of the hierarchy. This cross-sectional study considered 9045 Saskatchewan children who were ages 4-8 years in the 2008-2009 school years. Individual child characteristics, EDI domains, and vulnerability data were collected by the Ministry of Education teachers in the provincial 2008 EDI project; neighborhood contextual Census data were compiled by SPHERU staff at the University of Saskatchewan. Multilevel linear and logistic models were used to analyze the data. According to the results, individual characteristics, such as being Aboriginal, an ESL learner, male, and being absent from school; neighborhood characteristics such as income inequality; and geographical characteristics such as living in a large city have negative effects on EDI scores and exacerbating the odds of vulnerability. Compounding effects of Aboriginal−special skills, large city−Aboriginal, and large city−neighborhood median income were positive on the above outcomes with considerable either significance or strength, while those of neighborhood income inequality–Aboriginal, and large city–neighborhood income inequality were negative with notable significance and strength. Furthermore, neighborhood contextual variables contribute to a considerable proportion of health outcome variations and the results associated with neighborhood income inequality give further evidence of the income inequality hypothesis. The findings of this study recommend provincial child public health policy makers' extended attention to Aboriginal children, children with ESL status, those children living in neighborhoods with high income inequality and children from Regina.

DEDICATION

This thesis is dedicated to:

God the Holy Spirit, who has patted this child to conduct the research,

Zahra and Borjali, who have raised this child with their best efforts,

Donna and Murray, who have raised this child in the faith,

Maryam, who has helped this child in the early school years,

Brooke, who has been the angel of hope for children in Saskatchewan and Prairies.

ACKNOWLEDGEMENTS

First of all, the author would like to thank Professor Punam Pahwa as the co-supervisor of this thesis. During the time of working on this thesis, the author has been fully beneficiary of her academic support, wise comments and precious advice.

Second, the author very much appreciates Professor Nazeem Muhajarine as the co-supervisor of this thesis. He has fully supported the author in many issues, including research funding, with intelligent comments and invaluable advice.

Third, the author would like to express his gratitude toward his graduate committee members including Professors Mik Bickis and Michael Szafron for their comments and guidance as well as his program course instructors including Professors Juxin Liu, Cindy Feng, and Bonnie Janzen for their taught materials in theoretical statistics, computational biostatistics, and epidemiology, respectively, which have been used frequently in many areas of this thesis.

Finally, the author's special thanks go to his colleagues including Ms. Shan Jin, Ms. Ankona Banerjee, Ms. Meenu Rai and Mr. Prosanta Mondal for their help and support during working on this thesis.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARACY- Australian Research Alliance for Children and Youth

CHASS- Computing the Humanities and Social Sciences

CI- Confidence Interval

ECD- Early Childhood Development

EDI- Early Development Instrument

ESL- English as a Second Language

ICC- Intra-class Correlation Coefficient

MCI- Multiple Challenge Index

OR- Odds Ratio

PPVT- Peabody Picture Vocabulary Test

RIC- Regional Intersection Committee

STEP- Screening Test for Evaluation Preschoolers

TV- Total Variance

UNICEF- United Nations Children’s Fund

VPC- Variance Partition Coefficient

WHO- World Health Organization

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Let the little children come to me, and do not forbid them; for of such is the kingdom of God.
-Luke 18:16

This chapter deals with the rationale and objectives of this study by relating to the current literature of child developmental health and by presenting a brief review of its statistical objectives and epidemiological objectives.

1.1. Rationale

In recent years, studies concerning the socioeconomic status of pre-school children's families and school readiness have received increasing attention due to their long-term impact on a future generations' social and educational success and their overall health issues such as completing high school education, obtaining employment and contributing positively to society as well as their related mental health issues. In particular, for most children with special negative behavior in pre-school years, there is a high probability of maintaining this particular behavior across their life span and an increased risk of antisocial behaviors in adulthood.[1-3]

The concept of "school readiness" defined in section 2.2.1 was first introduced in the literature of the 1990s.[4] It can be considered as the product of the interaction between a child's skills, family environment support and community resources. This concept is defined as a set of children's characteristics, including their cognitive, communication, behavioral and emotional skills, which facilitate their school entrance level learning and adjustment.[4] School readiness is a universal concern and communities at city, provincial, national and international levels design and enact new policies and programs in order to improve the developmental outcomes of children. Furthermore, its associated literature focuses on risk in the frame of early beginning.[5]

In order to measure school readiness, an instrument is needed that provides an affordable, reliable, valid and clear multidimensional assessment of school readiness; and, considers the data in multilevel contextual frames aggregated at neighborhood, city and provincial levels. In addition to possession of such characteristics, the Early Development Instrument (EDI) was shown to be an effective tool in assessing the strength and direction of the relationship between neighborhood level variables and school readiness for kindergarten students; school readiness trend over time; and many variations in school readiness across geographical units.[6-8]

Based on the Saskatchewan provincial wide 2008-2009 EDI data set and Saskatchewan 2006 Census data, this study sets precedence the way for the larger pan-Canadian study of social determinants of children's developmental health which had been proposed in 2012 in terms of number of model hierarchy and study scale.[9] Furthermore, it allows researchers and public health policy makers to identify developmental health disparities among young children in Canada at neighborhood, city, regional and provincial levels.

1.2. Objectives

This research pursues two objectives: the first is statistical and the second is epidemiologic in nature. The statistical objective focuses on a general multilevel technique in order to analyze hierarchical EDI data. An analysis of this data is needed to determine important individual level, neighborhood level and geography level predictors of 5 EDI domains, vulnerability, and the Multiple Challenge Index in terms of strength and significance. Two special cases of a general multilevel model (a linear model and a logistic model) are used. Many studies in the field have examined the relationship of individual characteristics and children's developmental health on city scale.[7,8] Other studies have included neighborhood contextual characteristics in their research.[30-31] By adding geographical factors as the predictors of child's developmental health and enlarging the scale of the study to provincial level, the epidemiologic objectives are to (i) determine the significant determinants of a child's developmental health at the individual, neighborhood and geography level, (ii) discuss within-level and cross-level effect modifications of some key determinants in terms of statistical significance and strength and (iii) specify relative contributions of main determinants at each level to the variations of a child's developmental health outcomes.

CHAPTER 2: LITERATURE REVIEW

Train up a child in the way s/he should go, and when s/he is old s/he will not depart from it.
-Proverbs 22:6

This chapter presents the required study background including child public health, early childhood development, early development instrument, and statistical methods for multilevel data. Firstly, a brief historical background of child public health in the western world is presented and its formal definition is given. Secondly, the concepts of early childhood development and school readiness are defined and the importance of the earlier one in child public health is discussed. Thirdly, the concept of early development instrument and its psychometric properties are discussed. Based on the continuous 5 EDI outcomes, the concepts of vulnerability and multiple vulnerability are defined and some examples of the previous studies in the field are presented. Finally, the chapter ends with introduction of multilevel models, a brief discussion of 2-level model and some examples.

2.1. Child Public Health

2.1.1. Historical Background

Child public health, its related problems and position in overall public health have been brought to attention in a sequence of important events in the Western world and by many individual scientists with varying professional backgrounds and associations with national and international organizations.[11]

In the 19th century UK, city populations grew and living conditions, such as sanitation were poor. Food availability and its quality for children were variable based on the price. Child labor became a commonplace practice.[11] A series of Education Acts aimed to establish universal elementary education and a series of Factory Acts restricted employment of children. Furthermore, childhood mortality rates started declining in the second half of the century.[11] Edward Jenner (1749 - 1823) used a child, as an experimental unit, to test the hypothesis that the risk of catching smallpox can be eliminated in human subjects with cowpox material. The success of his experiment was the start of the vaccination of the children and other useful preventive tools in medical science.[11] As a result of 1854 London Cholera epidemic, a significant contribution to child public health was made by John Snow (1813 - 1858). He asserted that cholera is associated with high mortality in children. Next, Charles Dickens (1812-1870) gave a clear picture of poor children's health in society by publishing his famous novel *Oliver Twist*.[11]

In the first half of the 20th century, British public health interest focused on infant and maternal mortality rates. This interest led to the establishment of maternal and infant welfare services by charitable organizations and city public health departments. In addition, school meals and school health services were established for children. Regular disease statistics were generated by medical inspections as well. In the second half of the 20th century, hospital and health services for children continued to be developed and important developments in child immunizations were cultivated.[11] Later in the 20th century, vaccination and immunization services developed and had a dramatic effect on childhood infectious disease rates. Pre-school child health surveillance services were developed and hospital services for children flourished.[11] Regained interest in child public health and addressing social inequalities impact on it arose after a temporary decline in the early decades of that century. Child health care was revolutionized by David Morely (1923 - 2009) who provided clinics for under five years old children, and charted children's growth.[11] In Sweden, Lennart Kohler (1933-present), the father of modern child public health care, developed the concept of social pediatrics which is defined as a global holistic and multidisciplinary approach to child health. It considers child's health within the context of the family, school, environment, and society. Social pediatrics acts in three areas: (i) child health problems with social causes, (ii) child health problems with social consequences, and (iii) child health care in society. During her work with Aboriginal children in Western Australia, Fiona Stanley (1946-present) is among those contemporary researchers who brought to awareness the importance of environment and living conditions on children's health. In 1977, she helped to establish Western Australian Maternal and Child Health Research Database in order to predict trends in child health and determine the effects of preventive programs. Due to her efforts, in 2002 the Australian Government initiated the Australian Research Alliance for Children and Youth (ARACY) which aimed to advance collaboration and evidence-based action to improve the well-being of children.[11] In the U.S., as the UNICEF third executive director, James P. Grant (1922-1995) led the worldwide campaign of "Child Survival Revolution". This campaign improved immunization, oral rehydration therapy and breastfeeding in children, which saved the lives of at least 25 million children around the globe. He also helped consideration of the convention on the Rights of the Child by the UN General Assembly in 1989. In the UN convention on the Rights of the Child, the right of children to enjoy childhood to the fullest extent possible is recognized.[11]

2.1.2. Operational Definition

Child public health is a multifaceted and complex concept which involves a range of ideals, activities, and academic disciplines. It includes the investigation of health and illness patterns in children and factors that affect their health. It also contains studying the ways in which individuals, professionals, organizations, and societies can modify these factors in order to improve the health and well-being of children:[11]

Definition 2.1.2.1. Child public health is the art and science of promoting health and protecting the well-being of children. This practice includes preventing disease in infants, children, and youth through the skills and organized efforts of professionals, practitioners, their teams, wider organizations, and society as a whole.[11]

Children's health status should be seen in the context of their families, communities, environments and wider social and political settings which influence not only children individually but spheres of activity for child public health practice as well. Hence, by addressing social policy, family relationships, environmental concerns, and community structures, we can be hopeful to improve child public health.[11,12]

2.2. Early Childhood Development (ECD)

2.2.1. Concept

There are many expressions referring to early childhood development such as "early childhood care", "early childhood care and development", "early childhood care and education", "early childhood education", "early childhood intervention", and "early childhood service". Before explaining this concept we need to define the word "development":

Definition 2.2.1.1. Development refers to the process by which humans change both quantitatively and qualitatively as they grow older.[13]

Considering the above definition of "development", we are in a position to define the concept of "early child development":

Definition 2.2.1.2. Early Childhood Development (ECD) is a multifaceted concept from an ecological framework that focuses on a child's outcome and development. A child's development depends on characteristics of the child and the context, such as health, nutrition, protection, care and education.[14]

As the main components in the above definition of early childhood development, the applications of the terms “child” and “context” are explained.

Firstly, for the term “child” we consider dimensions age and domains of development. Regarding age, the definition of ECD covers all children prenatally until age 8 or when the transition to school is complete. Next, regarding domains of development, the terms physical health and motor development, cognitive and language skills, social and emotional functioning, ethical and spiritual development, and sense of national or group identity are considered.[14]

Secondly, as for the term “context” the current theoretical models argue that early child development is influenced by a set of contextual factors. One of the most famous of such models is the Bronfenbrenner’s Developmental Psychology Model which states that contexts are layered from the closet micro system (e.g. family) to the most distant macro system (e.g. international policies).[14] In interacting with their surrounding environment, children make rapid strides in all aspects of development.[15]

Early Childhood Development happens in three main age periods in which child development proceeds with different contextual critical factors and opportunities. These periods are conception to 3 years, 3 to 5 years and 6 to 8 years.[14] The period of “conception to 3 years” is the period of rapid growth in socio-emotional capabilities, mental capabilities, disease prevention and survival ability. Additionally, children experience sufficient physical development. In particular, brain architecture is constructed in a “bottom-up” sequence of development. Sufficient nutrition is necessary to prevent delayed gross and fine motor developments.[14] The period of “3 to 5 years” is a period in which the critical needs of the previous period advance and become wider. Among such needs are protection against violence, abuse, and negligence and expose to educational opportunities in family based, school based and community based programs. It has been proven that participation in such programs has a positive effect on child developmental outcomes.[14] Finally, the period of “6 to 8 years” which is sometimes referred to as the period of “school readiness”, is the period of transition to school in which group learning and socialization opportunities are likely to have the most positive impact on a child’s development. Development during this period is associated with learning, school completion, later skill development, gaining academic qualifications and success in a non-academic field. Ready children, ready families and ready schools are three pillars of child school readiness.[14]

2.2.2. Importance of ECD in Child Public Health

Regarding the concept of Early Childhood Development discussed in the previous section, one may wonder why this concept is important in the field of child public health. The answer of the proposed question constitutes some of key elements including long term biological impacts, aiding in guidance on interacting with children, guidance on curriculum planning in schools, guidance on observing and identifying children’s special needs, and guidance on advocating for and shaping of public policy of child public health. The above impacts will be discussed in the following.

(i) Long term biological impact on child health

Research shows that early childhood life experiences forms human biological in such a way that impacts physical and mental well-being, cognitive abilities and work productivity throughout one’s life. In addition, beyond the early childhood years, healing any of the above conditions becomes increasingly difficult .[16]

Moreover, early childhood development is a multifaceted concept combined from three core scientific concepts: Firstly, the brain is built over time and a major proportion of its circuiting is constructed during the early years of life by the bottom up process in which simple circuits provide basis for more complex ones. If the lower circuits are not wired appropriately, the higher levels of adaptation will be difficult to gain.[17] Secondly, extreme hardships such as stress, poverty, emotional abuse, and malnutrition in early childhood causes physiological interruptions that impact a child's developing brain leading to long-term defects in health, behavior, learning and emotional retreat.[17] Thirdly, neuroplasticity and ability to change behavior decrease over time. As the brain's ability to effectively execute more complex functions increases, its capability to recognize and adapt to new or unexpected challenges decreases. Wired circuits in the brain stabilize over time and being able to make alterations becomes increasingly difficult over time.[17]

(ii) Guidance to interacting with children

Children and adults think and act differently from each other. Children use a different form of language, interact with other people in distinct ways, and apply unique meanings to social events. Unique and unpredictable things cause them to be concerned, weep, or to be happy. Their developmental level and interests are independent of each other.[13] Screaming, running, playing, throwing things, joking and giggling with peers are among their essential needs. In order to smooth communication with children, comfort them, challenge their thinking, and help them to solve problems with peers adults need a deep understanding of how children act and why.[13]

(iii) Curriculum planning in schools

For developing appropriate activities and educational materials for children, a grasp on their thinking and behavior is essential. An understanding of development of children gives teachers and principals the required ideas and background to assess their limitations and design adapted programs and consider flexible strategies to meet their needs.[13] An inappropriate educational curriculum can be the result of overlooking and neglecting children's developmental status. Examples of classrooms that do not reflect knowledge of child development includes those that present young children with difficult, passive and overly abstract academic activities.[13]

(iv) Observing and identifying children's special needs

Observing children is the cornerstone of effective teaching. Careful observations of children's developmental needs constitute the basis of interventions and educational curriculum planning in schools. It leads teachers to identify children with special needs. For example, a child who displays very little motor activity, has language problems, or is rejected by his/her peers needs special attention. Next, focused observation may suggest causes and resolutions of potential problems as well.[13] As an example, a teacher may realize that a particular child does not respond to her/his efforts for social interaction in spite of the fact that there is a common mechanism of interaction for all children regardless of their ethnicity or cultural background. An assessment done by the community social service physicians reveals hearing problem in the child. Hence, they recommend the teacher to consider physical and visual provocations as a better strategy of communication.[13]

(v) Advocating for and shaping public policy

Research and theories on early child development demonstrate that working to improve community services and to influence public health policy has a direct impact on a child's development throughout his/her life time.[18-20] Based on scientific theories of change, grounded in strong evidence and high standard of implementation, investments in early childhood policies and effective intervention programs for vulnerable children increase quality of life prospects and greatest financial benefit for society. In particular, early child development programs help to overcome socioeconomic disparities by extending equal opportunities to all children before they enter primary school and reducing the demand for remedial education interventions targeting young school dropouts.[21]

2.3. Early Development Instrument (EDI)

2.3.1. Introduction

The Early Development Instrument (EDI) was developed as a uniform method to assess children's level of development in their first year of schooling. The instrument was developed in consultation with the Early Years Action Group, Parenting Centers and kindergarten teachers in the Toronto District School Board. After initial testing in 1998-1999, the EDI was refined and implemented in several communities around Canada.[6]

The EDI is a questionnaire containing 104 core questions in five general developmental domains. It is available in both English and French and is completed by the teacher or early childhood educator. Subjects are from ages 4 to 7 and the tool is administered usually in the second half of the kindergarten year.[6] The EDI also contains some additional local or community related questions as well as three sets of context related questions asking about special problems, special skills, and pre-school experience.[6]

The EDI data has uniform and consistent indicators of children's school readiness status aggregated at higher levels such as school, neighborhood, city, province and country. Results based on EDI data analysis enable policy makers to determine required support and identify available resources for children at the individual level, school level, neighborhood level and/or the city level in order to prepare them for the next school year.

The EDI assesses a child's ***school readiness*** in five general domains of child development so that for each there are some subdomains as follows:[6]

i. Physical health and well-being,

i.i. Physical readiness for school day , e.g. child arrives at school hungry,

i.ii. Physical independence , e.g. child has well- coordinated movements,

i.iii. Gross and fine motor skills, e.g. child is able to manipulate objects,

ii. Social competence,

ii.i. Overall social competence, e.g. child is able to get along with other children,

ii.ii. Responsibility and respect, e.g. child accepts responsibility for actions,

ii.iii. Approaches to learning, e.g. child works independently,

ii.iv. Readiness to explore new things, e.g. child is eager to explore new items,

iii. Emotional maturity,

iii.i. Pro-social and helping behaviour, e.g. child helps other children in distress,

iii.ii. Anxious and fearful behaviour, e.g. child appears unhappy or sad,

iii.iii. Aggressive Behaviour, e.g. child gets into physical fights,

iii.iv. Hyperactivity and inattention, e.g. child is restless,

iv. Language and cognitive development,

iv.i. Basic Literacy, e.g. child is able to write own name,

iv.ii. Interest in literacy/numeracy, and memory, e.g. child likes numerical games,

iv.iii. Advanced literacy, e.g. child is able to read sentences,

iv.iv. Basic numeracy, e.g. child is able to count to 20,

v. Communication skills and general knowledge.

Example v.i: child is able to clearly communicate one's own needs and understand others,

Example v.ii: shows interest in general knowledge about the world.

2.3.2. Psychometric Properties

The psychometric properties of the EDI such as reliability, concurrent validity, external validity, and predictive validity have been tested in Canada, United States, Australia, Jamaica and Kosovo. Comparisons of the Canadian results with those of other countries demonstrate that children's patterns of association are similar in all countries tested which provides strong evidence of the EDI's validity across different countries.[6]

Table 2.1 presents the reliability results for the Canadian EDI case.[6] As the table exhibits, in all 5 EDI domains internal reliability calculated via Cronbach alpha were high ranging from 0.84 to 0.96, test-retest reliability were high ranging from 0.82 to 0.94, inter reliability correlations were moderate (0.53) to high (0.80), and parent-teacher agreements were moderate ranging from 0.36 to 0.64.

Table 2.1 Summary of EDI Reliability Tests

	Physical Health and Well-being	Social Competence	Emotional Maturity	Language and Cognitive Development	Communication Skills and General Knowledge
Internal Reliability	0.84	0.96	0.90	0.93	0.94
Test-retest Reliability	0.82	0.92	0.89	0.82	0.94
Inter rater Reliability					
School Teacher-Daycare Teacher	0.69	0.80	0.77	0.72	0.53
School Teacher –Parent	0.36	0.50	0.36	0.64	0.41

The validity test of the EDI is usually considered to be three tests; concurrent validity (testing EDI performance in comparison with other previously validated instruments), external validity (testing degree of association between EDI scores and other measurement tools which have different perspective but measuring similar concepts), and predictive validity (testing EDI ability to predict outcomes).

Table 2.2 presents concurrent validity results of the Canadian EDI and First STEP (Screening Test for Evaluation Preschoolers) score and PPVT (Peabody Picture Vocabulary Test) tools.[6] The First STEP score is a 40-minute testing tool measuring children's cognitive and language abilities.[22] In this test, children's motor skills and socio-emotional skills were measured as well with moderate to high correlations ranging from 0.52 to 0.73 giving a reasonable evidence of the EDI's concurrent validity. Next, PPVT is a 15-minute test of receptive vocabulary and language skill/knowledge with the score being a reliable approximation of a child's IQ.[23] The test results show low to moderate correlations ranging from 0.05 to 0.57. These results show that the EDI has some concurrent validity when measured against a few of other commonly used child development measurement scales. The categorization of correlation values in this study follows Zady's categorization of $0 \leq r \leq 0.49$ as low, $0.50 \leq r \leq 0.69$ as moderate and $0.70 \leq r \leq 1.00$ as high[24].

Table 2.2: Summary of EDI Concurrent Validity Tests (correlation values)

	Physical Health and Well-being	Social Competence	Emotional Maturity	Language and Cognitive Development	Communication Skills and General Knowledge
First STEP (N=122)	0.54	0.65	0.73	0.58	0.52
PPVT (N=1700)	0.05	0.22	0.11	0.26	0.57

Table 2.3 presents the external validity results of the Canadian EDI and parent interviews measurement method in which instead of directly measuring a child's 5 EDI domains, the child's parents were questioned.[6] The results show low positively statistically significant correlations and yield that EDI has poor external validity.

Table 2.3: Summary of EDI External Validity Tests (correlation values)

	Physical Health and Well-being	Social Competence	Emotional Maturity	Language & Cognitive Development	Communication Skills & General Knowledge
Child measured – Parents interviewed	Min 0.15 - Max 0.34	Min 0.21 – Max 0.48	Min 0.21 – Max 0.48	Min 0.15 – Max 0.26	Min 0.15 – Max 0.26

In a longitudinal population based study for exploring the predictive validity of the EDI,[25] it has been shown that the Canadian EDI alone explains 36% of the variance of school achievement and two of its domains Physical health and well-being and Language & cognitive development have contributed to the prediction of school achievement more so than direct school readiness tests. The results of this study present EDI predictive validity as well as other measures of school readiness such as the Lollipop test,[26] which require more time and resources.

2.3.3. Analysis and Interpretation of Results

The EDI data for children can be aggregated from micro levels to macro levels with numerous ways of creating nesting structures. One such example is aggregated data from a 6-level EDI structure. Beginning with individual children as the smallest micro level, they are nested in schools, neighborhoods, cities, regional intersection committees (counties), and provinces (states) as the largest macro level, respectively. In addition, by adding data on race, cultural background, gross domestic product, statistics on education levels, school enrolment, information on socioeconomic status, etc. to the EDI data, a better picture of the size and significance of impacts of different variables on children's school readiness can be obtained.

For each of the five EDI domains the distribution of scores range from 0 to 10, and the site of measurement (e.g. school, neighborhood, city, etc.) is divided into two main categories so that each has two subcategories as follows:

i. On Track: A child whose score is higher than 25% of the site's distribution,

i.i. Very Ready: A child whose score is in the range of 75% −100% of the site distribution,

i.ii. Ready: A child whose score is on the range of 25%−75% of the site distribution,

ii. Not on track: A child whose score is lower than 25% of the site distribution,

ii.i. At risk: A child whose score is in the range of 10%−25% of the site distribution,

ii.ii. Vulnerable: A child whose score in the range of 0 – 10% of the site distribution.

Regarding the above categorization of children, one can define the concept of vulnerable child in the context of school readiness:

Definition 2.3.3.1. A child is called vulnerable (not ready to learn) if for at least one of five EDI domains s/he scores vulnerable in the related site.[6]

In order to draw comparisons between communities of different measurement sites (e.g. schools, neighborhoods, cities, etc.) the researchers consider two quantities of average site percentage in a special category and its associated range interval. As an agreement, to compare average site percentage of two sites A and B, a percentage difference ($\frac{B-A}{A} \times 100\%$) of at least 10% presents enough evidence on significant difference. Wider range interval for average site percentage in a special category of the site A is a sign of a higher variation and degree of inequality across its communities in comparison to the site B.[6] For example, for two neighborhoods A and B with percentages of vulnerable children averages (range interval) of 22% ([10.5%−46.7%]) and 28% ([5.7%−26.5%]) of the site distributions, respectively, neighborhood B has a significantly higher average percentage of vulnerable children than neighborhood A (percentage difference of 27% > 10%) while neighborhood A has a higher degree of inequality across its schools as it has wider range interval.

Referring to the above 5 EDI domains, one can observe that in total there are 16 sub domains. Based on evidence of a mix of poor and average children's scores and exceptions in each domain, a boundary score was determined and if a child had score lower than that point he or she was recognized as experiencing challenge.[27] This special score is called the cut-off point. Table 2.4 presents the list of cut-off points for each of subdomain:

Table 2.4: EDI 16 Subdomain Challenge Cut-off Scores

Domain	Challenge Cut off	% below the challenge cut off in normative sample	Domain	Challenge Cut off	% below the challenge cut off in normative sample
Physical Well being			Language & Cognitive Development		
Physical Readiness for School Day	6.249	3.9	Basic Literacy	7.499	11.0
Physical Independence	9.999	8.9	Interest in Literacy/Numeracy	7.999	15.8
Gross and Fine Motor Skills	6.499	21.8	Advanced Literacy	3.329	19.4
			Basic Numeracy	8.569	14.2
Social Competence					
Overall Social Competence	4.999	8.4			
Responsibility and Respect	4.999	4.7	Communication Skills & Knowledge		
Approach to Learning	4.999	8.1	Communication Skills & Knowledge	6.939	29.0
Readiness to Explore New Things	4.999	3.2			
Emotional Maturity					
Prosocial and Helping Behaviour	4.999	33.5			
Anxious and Fearful Behaviour	4.999	2.1			
Aggressive Behaviour	7.139	7.8			
Hyperactivity and Inattention	5.709	13.1			

The EDI is a useful tool to measure comparisons over time in longitudinal studies.[6] The baselines of such studies are set in two methods: firstly, if the provincial (state) level cut offs and means are available, then they are used to establish the baseline. Secondly, in case of lack of provincial (state) level data, the normative cut offs are used to establish the baseline. The first method is preferred to the second, as that data were gathered provincially in the last three year's period while the second one was gathered nationally in the last five years. In this research we use the cross-sectional approach.

The last definition in this subsection has close connection with vulnerability:

Definition 2.3.3.2. Multiple Challenge Index (MCI) is an indicator of a child experiencing vulnerability in at least 3 EDI domains so that a child with MCI = “yes” is considered to have scores under the “challenge” cut-off points in at least 9 out of 16 EDI subdomains and MCI= “no” if else.[6]

A child with MCI status ‘yes’ is referred as multiple-vulnerable child.

2.3.4. Examples of Previous Studies

In recent years, researchers have shown that the EDI is an effective tool for assessing children’s school readiness and health status through different perspectives, such as studies of individual exclusive determinants, studies of family- neighborhood determinants, and those considering both individual and family-neighborhood determinants in the analysis.

Firstly, the EDI has shown to be an effective tool in assessing children’s school readiness and health status in studies of individual determinants. Muhajarine et al. [28] considered a 1-level logistic model. In this model, the existence of the multiple challenges that children face has been predicted by an individual child’s characteristics, such as holding an Aboriginal status, female status, possessing fewer special skills, and number of special problems with which the child is faced. All of these characteristics were statistically associated with a higher likelihood of being rated as having multiple challenges, with odds ratios of 3.38, 1.91, 2.65, and 2.61, respectively.

Secondly, the EDI has shown to be an effective tool in assessing children’s school readiness and health status in studies of family- neighborhood determinants. As an example, Janus and Duku [29] used the EDI to explore the impact of socioeconomic status, family structure, parent health and parent involvement in literacy development on children’s school readiness gap in a 1-level logistic regression model. The outcome variable in the model was considered to be vulnerability= “1” if the child was diagnosed as vulnerable and vulnerability = “0”, if otherwise. The 15 model predictors included socioeconomic status (4 variables), family structure (3 variables), child health (4 variables), parent health (2 variables), parent involvement (2 variables), and demographics (2 variables). Children coming from low income families, living with smoking parents, or living in families with parents with poor literacy skills had a higher risk of vulnerability with odds ratios of 2.23, 1.25 and 1.29, respectively. As an additional example, Lapointe et al. [7] used the EDI to show the impact of neighborhood contextual factors on children’s school readiness when accounting for their age and gender. There were 13 neighborhood level variables in the 2-level model, and they established that 8, 2, 3, 4, and 6 of them significantly predicted Physical health and well-being, Social competence, Emotional maturity, Language& cognitive development, and Communication skills & general knowledge domains, respectively. As a third example, Cushion et al. [8] used the EDI in repeated measures data to fit a 2-level linear model with 7 predictors. They considered neighborhood poverty index as the principal component outcome of seven correlated neighborhood variables and observed that over time neighborhood poverty index was significantly related to declining scores of the children’s Physical health and well-being but it was non-significantly related to declining scores of the children’s Communication & general knowledge domain.

Thirdly, the EDI has shown to be an effective tool in assessing children’s school readiness and health status in studies considering both individual and family-neighborhood determinants in the analysis. For example, Oliver et al.[30] used the EDI to investigate the relationship between individual and neighborhood socioeconomic characteristics on kindergarten students’ school readiness in a 2-level linear model with 2 individual child variables and 6 neighborhood related variables. The results indicated that a higher family income and speaking English as the maternal language are significantly associated with higher scores in almost all 5 EDI domains. In addition, at the neighborhood level, children in neighborhoods with higher median family income, higher percentage of lone-parent families or higher unemployment rate have lower scores in almost all 5 EDI domains. As a second example, Puchala et al. [31] used the EDI to study the impact of neighborhood contextual factors in addition to individual child factors on school readiness outcomes in a 2-level linear model with 7 individual child predictors, two neighborhood related predictors and one cross-level interaction between child and neighborhood characteristics. At the individual level, children with English as a Second Language (ESL) status, male status, Aboriginal status, or special needs status had significantly lower scores in two domains of Emotional maturity and Communication skills & general knowledge. At the neighborhood level, children from neighborhoods with a higher percentage of employed adults had significantly higher scores in Communication & general knowledge while children from neighborhoods with a lower percentage of the population having changed residences within the previous year, or had higher ethnic diversity had higher EDI scores on Emotional maturity and Communication & general knowledge domains. Finally, the interaction between a child's ESL status and neighborhood ethnic diversity level showed that for neighborhoods with higher diversity, the mean differences of EDI scores in the Emotional maturity and Communication & general knowledge domains between ESL children and non-ESL children were attenuated which supports that neighborhood ethnic diversity buffers school readiness impact in ESL children.

2.4. Statistical Methods for Multilevel Data

2.4.1. Introduction to Multilevel Models

Multilevel models first emerged as a solution to overcome the challenges of integrating micro level and macro level information into a single model in educational statistics, quantitative sociology, and demographical areas of research.[32] This statistical model is a special generalization of the linear regression model in which the variables appear in more than one level. The key concept in this model is “level”. In spite of the five common assumptions of general linear models of existence, linearity, normality, homoscedasticity, and independence of outcomes, a multilevel merely needs the first three assumptions. The last two are violated at each level of its variables.[33] Before introducing this statistical model one needs to review some useful definitions:

Definition 2.4.1.1. Aggregated data refers to data in which the higher level unit is constructed by combining information from the lower level units of which the higher level unit is composed.[34]

One example of aggregated data can be seen in summaries of first grade students' mathematics scores in public and separate school systems. Figure 2.1 shows that students are nested in schools while each level of the hierarchy has its related variables (e.g. gender for the level 1 and school type for the level 2).

Definition 2.4.1.2. Multilevel analysis refers to an analytical approach that is appropriate for aggregated data in which units at a lower (micro) level are nested within units at a higher (macro) level.[34]

Multilevel analysis and contextual analysis were originally used in sociology, but are now used to investigate the effect of collective characteristics on individual level outcomes. These methods of analysis are sometimes considered synonyms of each other, in that their related statistical models both include higher level and lower level predictors in the standard linear regression with lower level individual outcomes. However, multilevel analysis is more general than contextual analysis as its related model allows consideration of the possibility of residual correlation between lower levels units (e.g. individuals) within higher level units (e.g. groups) while contextual analysis lacks such assumption. The multilevel analysis model also allows examination of the relationship between factors associated with group variability and factors associated with it, two important possibilities that the contextual model lacks.

Definition 2.4.1.3. The multilevel model is the statistical model used in multilevel analysis.[34]

The term Multilevel Model has other conventional synonyms in biostatistics literature such as the mixed linear model, the hierarchical linear model, the random effect model, the random coefficient model, the covariance component model, the variance component model and the mixed model.[34]

There are three main types of multilevel models, which depend on the status of the coefficients in the linear regression model.[35] Firstly, the random intercept model, in which intercepts are random variables, and slopes are fixed across groups. The outcome for each individual is predicted by the intercept which varies across groups. This type of model provides information on intra-class correlations, a useful criterion in deciding whether to use multilevel model. Secondly, the random slope model in which intercepts in the model are fixed. In addition, the slopes are random variables which allow them to vary across groups. Thirdly, the random intercepts and slopes model which includes both intercepts and slopes as random variables. This model is the most complex type of the multilevel models. Some more definitions are needed before introductions of concepts of variation partition coefficient and intra-class correlation coefficient.

Definition 2.4.1.4. Individual level variables refer to variables that characterise individuals at the lowest level of aggregated data.[34]

In the context of figure 2.1, examples of individual variables include age, gender, ESL status, etc.

Definition 2.4.1.5. Group level variables refer to variables that characterise groups at higher levels of aggregated data.[34]

Group level variables are sometimes referred to as macro variables, ecological variables, or more generally, contextual variables.[34] As an example of group level variables found in aggregated data of figure 2.1, school average in mathematics is considered to be a group level variable characterising children nested in schools.

Definition 2.4.1.6. Cross-level interaction refers to the interaction between a group variable and a variable in a lower level which is nested in the group variable.[34]

Cross-level interaction can be interpreted as the modification of the effects of the lower level variables by the effects of the higher level variables which are composed of the lower level units. An example of aggregated data in figure 2.1 with individual children nested in schools is the effect of school type (public or separate) modifying the effect of child's gender (male or female) on his or her annual GPA as an outcome variable by the cross-level interaction: school type × gender.

Definition 2.4.1.7. Variance component (σ) refers to the variance between groups (individuals) at a specific level of the hierarchy.[34]

Note that the total variance (TV) of the lowest level outcome variable can be written as the summation of all variance components of all levels. Also, existence of the variance components is a prominent characteristic of multilevel models that distinguishes them from traditional contextual models and population average models.[34] In order to interpret the absolute magnitudes of the variance component at a special level we calculate its related 95% coverage interval - $(-1.96.\sigma , +1.96.\sigma)$ and for reporting them we usually center them around some interpretable value such as the mean of all aggregated data.[36]

Definition 2.4.1.8. Variance partition coefficient (VPC) refers to the proportion of the total variance that lies at a specific level of the model hierarchy.[36]

Note that the VPC of a specific level is calculated via dividing the variance between groups at that level by the total variance of the lowest level outcome variable.[36]

Definition 2.4.1.9. Intra-class correlation coefficient (ICC) refers to the extent to which values of the lowest level outcome variable are similar for individuals belonging to the same group.[36]

Note that the ICC is calculated via dividing the summation of higher level variance components (with reference to the group level for which the individuals belong) by the total variance of the lowest level outcome variable.

For the sake of simplicity and preparation the required background for 3-level model, an example of 2-level model is presented to explain the above concepts and their related mathematical formulas. In chapter three, the extension of 2-level model to 3-level model will be presented. Consider a 2-level model in figure 2.1 in which there are n schools in a city and in the jth school ($1\leq j \leq n$) there are n_j students whose health outcome variables are measured from the range of 0 to 10. As figure 2.1 shows, level-1 of the aggregated data set is represented by students and level-2 of the data set is represented by the schools in which the students are nested.

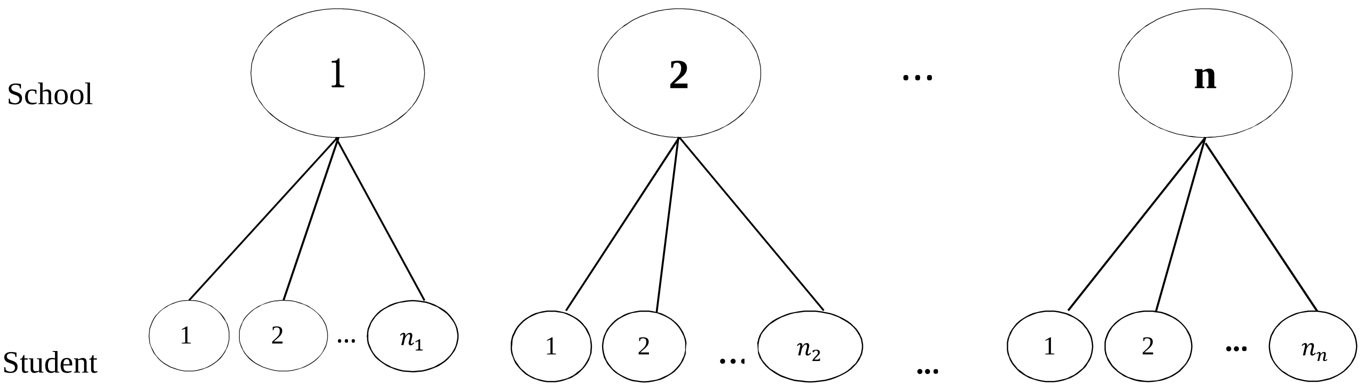

Figure 2.1: Data Structure for a Two-level Hierarchical Model

From a mathematical perspective, a 2-level regression model for the above aggregated data can be written in the form:[39]

$$Y_{ij} = \beta_{0j} + \beta_{1j}.X_{1ij}^{(1)} + \cdots + \beta_{pj}.X_{pij}^{(1)} + \gamma_1.X_{1j}^{(2)} + \cdots + \gamma_q.X_{qj}^{(2)} + e_{ij}\,,\ \ e_{ij} \sim N(0, \sigma_e^2) \qquad (2\text{-}1)$$

where Y_{ij} is the health outcome of the i th student in the j th school, $X_{kij}^{(1)}$'s (1≤k≤p) are level-1 individual independent variables pertaining to the characteristics of the ith student in the jth school and $X_{lj}^{(2)}$'s (1≤l≤q) are level-2 school variables pertaining to characteristics of the jth school. Next, depending on whether the intercept β_{0j}, coefficients $\beta_{kj} (1 \le k \le p)$ or both are random variables, the 2-level model is called a random intercept model, a random slope model or a random intercept and slope model, respectively. The random intercept and slope model is defined by:

$$\beta_{kj} = \beta_k^{(0)} + u_{kj}\ (\beta_0^{(0)} = \beta_0),\ (0 \le k \le p), \qquad (2\text{-}2)$$

in which $\beta_k^{(0)}$ are fixed numbers, u_{kj} are normal random variables with $E(u_{kj}) = 0$, $Var(u_{kj}) = \sigma_{u_k}^2 < \infty$ and the random error term e_{ij} satisfies $E(e_{ij}) = 0$. Note that the model is simplified to random slope model if $u_{0j} = 0$, or it is simplified to random intercept model if $u_{kj} = 0$, (1≤k≤p).

Referring to model (2-1), in order to test fixed effects the typical null hypothesis is:

$$H_0:\ \beta_k^{(0)} = 0\ (\ H_0{:}\,\gamma_l = 0\ )$$

with t- value of t = $\frac{\widehat{\beta_k^{(0)}}}{s.e.(\widehat{\beta_k^{(0)}})}$, $df = (\sum_{i=1}^{n} n_i) - 1$ (t = $\frac{\widehat{\gamma_l}}{s.e.(\widehat{\gamma_l})}$, $df = \mathrm{n} - 1$) where in which the numerator is the maximum likelihood estimate of the coefficient and the denominator is the estimated sampling standard error of the numerator.

In order to test the cluster effect, that is whether we need a multilevel model at all, we compare the 2-level model (2-1) with $u_{kj} \neq 0, (0 \leq \mathrm{k} \leq \mathrm{p})$ with simpler linear regression model (2-1) with $u_{kj} = 0, (0 \leq \mathrm{k} \leq \mathrm{p})$ using the likelihood ratio tests due to the fact that both models are fitted by the maximum likelihood (ML) method.[36] The null and alternative joint hypotheses are written as:

$$H_0{:}\ \sigma_{uk}^2 = 0, (\text{for all } 0 \leq \mathrm{k} \leq \mathrm{p}\ )$$

$$H_1{:}\ \sigma_{uk}^2 > 0, (\text{for some } 0 \leq \mathrm{k} \leq \mathrm{p}).$$

Let L_0, and L_1 be likelihood values for the linear regression model (2-1) and the 2-level model (2-1), respectively. Then the LR test statistics for testing above likelihood ratio test is given by:

$$\mathrm{LR} = (\text{-}2 \log (L_0)\ ) - (\text{-}2 \log (L_1)\ ),$$

which should be compared to a chi-squared distribution with degrees of freedom equal to the number of extra parameters in the 2-level model (2-1), that is, $X_{(p+1,\ 0.95)}^2$.

The statistical power of effects of the multilevel model depends on the number of groups (individuals) in the level to which the effect belongs. In addition, for conducting sufficient powerful research, a large sample size is required, and the number of groups at levels higher than level 1 is more important than the number of individuals at level 1.[37]

We give some interpretations of intercepts and slopes of the equation (2-1). The intercept β_0 in (2-2) measures the overall mean of Y_{ij} across all schools and all students. The mean value of Y_{ij} for school j called random intercept is $\beta_{0j} = \beta_0 + u_{0j}$, and the random school effect u_{0j} is the difference between school j's mean and the overall mean. Next, coefficients β_{kj} in (2-1) are referred to as random slopes having mean $\beta_k^{(0)}$, respectively, and variances σ_{uk}^2, $(1 \leq k \leq p)$, respectively. The random variables u_{kj}, $(1 \leq k \leq p)$ are referred to as level 2-residuals. Schools with high (low) values of level-2 residuals tend to have students with high (low) outcome variable scores. With the model total variance TV, a 95% coverage interval for school effects is given by $(\beta_0 - 1.96\sqrt{TV - \sigma_e^2}, \beta_0 + 1.96\sqrt{TV - \sigma_e^2})$ meaning that 95% of school effects are expected to lie in the range of $\beta_0 - 1.96\sqrt{TV - \sigma_e^2}$ to $\beta_0 + 1.96\sqrt{TV - \sigma_e^2}$. Thus, schools at the 97.5 th percentile of the school distribution are estimated to score $3.92.\sqrt{TV - \sigma_e^2}$ points higher than those schools at the 2.5 th percentile. The random variable e_{ij} is referred to as level-1 residual, and students with high (low) level-1 residual values tend to have higher (lower) outcome variable scores relative to other students from the same school. A 95% coverage interval for student residual errors is given by $(\beta_0 - 1.96.\sigma_e, \beta_0 + 1.96.\sigma_e)$ meaning that within schools, those students at the 97.5 th percentile of the distribution are estimated to score $3.92.\sigma_e$ points higher than those students at the 2.5 th percentile.[36]

Referring to random intercept and slope model (2-1) the total variance (TV) is given by:

$$\mathrm{TV} = \sum_{k=0}^{p} \sigma_{uk}^2 \, {x_{kij}^{(1)}}^2 + 2\sum\sum_{0\leq k_1 \neq k_2 \leq p} \sigma_{uk_1uk_2} {x_{k_1ij}}^{(1)} {x_{k_2ij}}^{(1)} + \sigma_e^2, \quad ({x_{0ij}}^{(1)} = 1)$$

where in which $\sigma_{uk_1uk_2} = \text{cov}\left(u_{k_1j}, u_{k_2j}\right)$, for all ($1\leq k_1, k_2 \leq p$) and $\text{cov}\left(u_{kj}, e_{ij}\right) = 0$, for all ($1\leq k\leq p$). Hence, the variance partition coefficient (VPC) and Intra-class correlation (ICC) are given by:[36]

$$\text{VPC(school)} = 1- \frac{\sigma_e^2}{TV}, \qquad (2\text{-}3)$$

$$\text{VPC(student)} = \frac{\sigma_e^2}{TV},$$

and

$$\text{ICC(student)} = 1- \frac{\sigma_e^2}{TV}. \qquad (2\text{-}4)$$

Note that for random intercept model (2-1), the related values of VPC and ICC take the following simpler form:[36]

$$\text{VPC(school)} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2+\sigma_e^2}, \qquad (2\text{-}5)$$

$$\text{VPC(student)} = \frac{\sigma_e^2}{\sigma_{u0}^2+\sigma_e^2},$$

and

$$\text{ICC(student)} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2+\sigma_e^2}. \qquad (2\text{-}6)$$

One important feature of multilevel modeling is that one can enter the cross-level interactions terms $X_{kij}^{(1)}.X_{lj}^{(2)}$ ($1\leq k\leq p$, $1\leq l\leq q$) to the 2-level model above in order to discuss the effect of level-1 variable $X_{kij}^{(1)}$on Y_{ij} when it changes the values of level-2 variable $X_{lj}^{(2)}$. In this case, the equation (2-1) will have the following general form of the 2-level model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}.X_{1ij}^{(1)} + \cdots + \beta_{pj}.X_{pij}^{(1)} + e_{ij}\ ,\ e_{ij}\sim N(0,\sigma_e^2) \quad (2\text{-}10)$$

$$\beta_{kj} = \gamma_{k0} + \gamma_{k1}.X_{1j}^{(2)} + \cdots + \gamma_{ks_k}.X_{s_kj}^{(2)} + u_{kj} \quad (0 \leq k \leq p),$$

with $\mathbf{s_k + 1}$ fixed effects for each $\boldsymbol{\beta_{kj}}$, and the total number of fixed effects in the 2-level model (2-10) being equal to $\sum_{k=0}^{p}(\mathbf{s_k + 1})$. [39]

2.4.2. Applications of Multilevel Models

The two primary major applications of multilevel models appeared in the fields of educational measurement and sociology in the mid-1980s. There are more applications of these models in other examples of hierarchical data structures in the fields of survey data, repeated measures, twin studies, and meta-analysis as follow.[32]

First of all, multilevel models have application in analysis of survey data that is obtained from nested sampling in heterogeneous subgroups with multiple levels of nesting. Such data is obtained by drawing information from the highest levels of the hierarchy, such as province or state, then the second highest level of hierarchy, such as Regional Intersection Committee (RIC) or county and so forth until the lowest level (i.e. individual). One example of such survey data is the mathematics and statistics scores of grade 1 children of a province (state) with RIC (county), city, neighborhood, school, and individual children as other levels of the hierarchy.[32]

Second, multilevel models have application in analysis of longitudinal data or repeated measures data in which a single outcome variable is measured at a number of fixed time periods (considered as the level 1 of the hierarchy) for a group of individuals (considered as the level 2 of the hierarchy). There is some flexibility in such data sets. For example, each individual outcome variable can be measured at different time points, and there can be missing data. An example of such longitudinal data is monitoring a group of children's sleep patterns on successive nights of one month and measuring and recording the extent to which they coughed each night. Here, children are level 2 of the hierarchy and nights are the level 1.[38]

Third, multilevel models have applications in analyzing twin studies where based on necessity, nature or design of the study the level-2 group size of the study is typically two. Examples of twin studies based on necessity are 2-level hierarchical data sets in which the level

2 groups are constituted of married couples, identical twins, or paired siblings. Examples of twin studies based on this design are 2-level longitudinal studies in which for a group of individuals, the health status of the subject (e.g. blood pressure, weight, etc.) is measured before and after administration of a special drug or undergoing a particular treatment.

Fourth, multilevel models are useful in meta-analysis in which a quantitative analysis of data and results from multiple previous studies on the same scientific problem is conducted. Multilevel modelling yields invaluable information about meta-analysis including 1) estimation of the average effect size across a set of studies, 2) estimation of the variance of the effect-size parameters, 3) possibility of posing and testing a set of linear regression models in order to explain variation in the effect size parameters, 4) estimation of residual variance of each linear regression model effect size parameters, and 5) empirical Bayes estimates of each study's effect. One example of application of multilevel modelling in meta-analysis is a study considering the existence of the effect of teacher expectancy on pupils' IQ as null hypothesis in 19 studies over 20 years, which concluded in contradicting results.[39] The results of the study 2-level model showed that on average experimental students scored about 0.083 standard deviation units higher than controls with significant important variability of true-effect sizes.

CHAPTER 3: METHODS

My fathers planted for me, and I planted for my children.
- Hebrew Proverb

This chapter deals with study objectives, collection and measurement of variables of interest, the 3-level random intercept model and its modelling process. Firstly, the study 3-level random intercept model is introduced and then, the study epidemiological objectives in the frame of three main questions are presented. Secondly, data collection process, number, position and measurement of variables of interest in the model hierarchy are discussed. Finally, the chapter continuous in description of the study 3-level model as the special generalization of the 2-level model in chapter two and it closes with its related step by step modelling process of the data.

3.1. Study Objectives

3.1.1. Statistical Objective

Most used statistical multilevel models in education research are 2-level or at most 3-level models. In this research, an example of applying a 3-level model is given that from mathematical perspective can be written in the form:[39]

$$f\left(E\left(Y_{ijk}^{(t)}\right)\right) = \beta_{0jk}^{(0)} + \sum_{s=1}^{u} \beta_s^{(1)}.X_{sijk}^{(1)} + \sum_{s=1}^{v} \beta_s^{(2)}.X_{sjk}^{(2)} + \sum_{s=1}^{w} \beta_s^{(3)}.X_{sk}^{(3)}, \quad (1 \leq t \leq 7)$$

where $X_{sijk}^{(1)}, X_{sjk}^{(2)},$ and $X_{sk}^{(3)}$ were level-1 (represented by subscript i), level-2 (represented by subscript j), and level-3 (represented by subscript k) variables and for continuous outcome variable $Y_{ijk}^{(t)} \sim \mathrm{N}\left(\mathrm{e}_{\mathrm{ijk}}, \sigma_{\mathrm{e}}^2\right)$, $(1 \leq t \leq 5)$ and *f(x) = x* whilst for binary outcome variable $Y_{ijk}^{(t)} \sim$ Bernouli (π_{ijk}) where $\pi_{ijk} = Pr\left(Y_{ijk}^{(t)} = 1\right)$, $(6 \leq t \leq 7)$ and *f(x) = logit (x).*

Moreover, for the linear model *f(x) =x*, $Y_{ijk}^{(t)}$ $(1 \leq t \leq 5)$ represented one of the five EDI domains outcome variables for the *i* th child in the *j* th neighborhood in the *k* th geographical area, and for the logistic model *f(x) = logit (x)*, $Y_{ijk}^{(t)}(6 \leq t \leq 7)$ represented being recognized as "vulnerable child" and experiencing "multiple challenge", respectively, for the same child. In addition,

$$\beta_{0jk}^{(0)} = \beta_0 + \gamma_{jk}, \quad \text{and} \quad \beta_s^{(t)} = \beta_0^{(t)} + \gamma_s^{(t)} \ (t = 1,2,3)$$

where $\beta_0, \beta_0^{(t)}$ $(t = 1,2,3)$ were fixed numbers and $\gamma_{jk}, \gamma_s^{(t)}$ $(t = 1,2,3)$, were random variables satisfying $E(\gamma_{jk}) = E\left(\gamma_s^{(t)}\right) = 0$, and $Var(\beta_{jk}), Var\left(\gamma_s^{(t)}\right) < \infty, (t = 1,2,3)$ and all j, k, s. Also, the random error term e_{ijk}, satisfied the condition $E(e_{ijk}) = 0$. Similar to the 2-level model, we could consider the cross-level interaction terms such as

$$X_{s_1 ijk}^{(1)} . X_{s_2 k}^{(3)}$$

in the above model in order to discuss the effect of a lower level variable $X_{s_1 ijk}^{(1)}$ on the outcome variable Y_{ijk} with it changes on the values of a higher level variable $X_{s_2 k}^{(3)}$.

3.1.2. Epidemiological Objective

In Saskatchewan, kindergarten children were nested in neighborhoods; the neighborhoods were nested in cities, and, finally, cities were nested in ten provincial regional intersectional committees (RIC's).[10] Hence, any child's health outcome can be considered as a linear function of the individual or level-1characteristics (e.g. age, gender, Aboriginal status, ESL status, etc.), neighborhood or level-2 characteristics (e.g. high, medium or low income status), geographical area or level - 3 characteristics (e.g. city, town or village status), and RIC or level-4 characteristics (e.g. lacking or holding major city status). Considering that there were only three major cities in the province with the minimum population of 35,000 citizens (i.e. Saskatoon, Regina and Prince Albert), and there were ten RIC's, the nested nature of the 4-level hierarchical model was violated. Therefore, level 4 has been removed from the hierarchy and the 3-level hierarchical model has been chosen to be focused on instead. Regarding our mentioned 3-level model, and available EDI data from Saskatchewan Ministry of Education for the school years 2008-2009, three main questions of interest arose:

Question 1: What are the significant determinants of developmental health within each of the constituents and, nested levels, namely: child, neighbourhood, and geographical area?

Examples:

1.1. Do male children have better health outcomes than female children?

1.2. Do Aboriginal children have better health outcomes than non-Aboriginal children?

1.3. Do children within French Immersion program have better health outcomes than others?

1.4. Do native English speaking children have better health outcomes compared to others?

1.5. Does age have positive impact on children's health outcomes?

1.6. What is the role of neighborhood income inequality on children's health outcomes?

1.7. Are children's health outcomes better in major cities than in non-urban areas? (This is called the impact of population density on children's health outcomes).

Question 2: What are the selected determinants that moderate the effects of other determinants significantly at another level?

Examples:

2.1. Are the health outcomes for children absent from school worse in neighborhoods with higher income inequality than in those from lower income inequality?

2.2. Do Aboriginal children have worse health outcomes in neighborhoods with higher income inequality than in those with lower income inequality?

2.3. Do Aboriginal children have better health outcomes in major cities than in non-urban areas?

2.4. How does neighborhood income inequality modify the effect living in a major city has on a child's developmental health?

2.5. How does neighborhood median income modify the effect living in a major city has on a child's developmental health?

Question 3: What are the relative contributions of significant determinants at each level to the variance of developmental health outcome?

Examples:

3.1. For fixed geography characteristics, what are the between neighborhood variations?

3.2. For fixed geography, and neighborhood characteristics, what are the between children variations?

3.3. Which level of variables and individual child's characteristics has the most contribution to variation of a child's health outcomes?

3.2. Data Collection

The first part of this study data consists of cross-sectional EDI data. This cross-sectional data was conducted by teachers and educational assistants at Saskatchewan provincial school divisions during 2008-2009 school years. It included the 5 EDI domains, vulnerability and the Multiple Challenge Index as outcome variables. The informations covered 9045 students from kindergarten to grade two in 418 schools around the province.

The second part of the data set used in this study was collected by the staff at the University of Saskatchewan Social Science Research Libraries (ssrl.usask.ca) and was based on the 2006 neighborhood boundaries taken from Statistics Canada. The data utilizes Saskatchewan neighborhood based list of postal codes for the cities of Saskatoon, Regina, Prince Albert, and non-urban areas.

The third part of the data set was collected by staff at University of Saskatchewan Spatial Initiative division (www.spatial.usask.ca) and was based on 2006 census data at Dissemination Area level available from Computing the Humanities and Social Science (CHASS) data center at the University of Toronto (datacenter.chass.utoronto.ca/census). The data set includes Saskatchewan neighborhood level variables of Gini Index (as the twice of the area between the line of equality and Lorenz curve), Median Income, % of Unemployment for 15+ years old people, % of holders of high school degree, and average value of dwelling.

In order to acquire the final data set of the study, the first and second data sets were merged based on the variable "postal code" and after some minor modifications (including removing subjects whose postal code did not match or at least two of their 5 EDI outcomes were missing), the outcome data set was merged with the third, yielding the final version of this study's data set.

3.3. Framework of Variables

In the first part of this study's data set, there were 160 variables for Saskatchewan children 4-8 years old in the 2008-2009 school year including 104 items comprising the EDI questionnaires. In the second part of the data set, there were 2 variables: neighborhood names and child residential postal codes. In the third part of the data set there were 6 variables including neighborhood names and 5 neighborhood level variables including school type, Gini index, median income, unemployment rate, percentage of high school graduates and average dwelling value. Merging the above three data sets the final data set of the study included 166 variables of which 28 variables were used in the study data analysis. The 28 variables in the study data analysis were divided into three sets of variables namely three hierarchy variables, 18 explanatory variables distributed in three levels and 7 outcome variables distributed into two sets. Table 3.1 summarizes the study variables and their corresponding explanations:

Table 3.1: List of Study Main Variables

Variable	EDI Questionnaire/ Census Explanation
Hierarchy Variables	
Child	The unique EDI ID considered for the child
Neighborhood	Name of child's residential neighborhood
Geographical Area	Name of child's city
Explanatory Variables	
Child Characteristics	
Age	Child's age at the time of survey
Days Absent	Child's number of days absent from school in one year
Number of Special Skills	Child's number of special skills
Number of Special Problems	Child's number of special problems
Gender	Child's Gender (male or female)
Aboriginal	Child's Aboriginal status
Special Needs	Child's requirement of special needs (yes or no)
French/English Immersion	Child's participation in French/English Program (yes or no)
English	English as child's maternal language (yes or no)
Non-Parental Care	Whether child is in non-parental custody (yes or no)
Language/Religion Class	Whether child attends a language/religion class
Neighborhood Characteristics	
School Type	Neighborhood School Type as Public, Separate or Francophone
Gini	Neighborhood Gini Index
Median Income	Neighborhood Median Income in $10,000 (per capita)
Unemployment Rate	Neighborhood unemployment rate for people 15+ years age (%)
High School Diploma	Neighborhood rate of high school diploma holders (%)
Average Dwelling	Neighborhood average value of dwelling in $ 10,000
Geographical Area Characteristics	
Geographical Area	City Name such as Saskatoon, Regina, Prince Albert or Non-urban Areas
Outcome Variables	
Continuous	
Physical Wellbeing	Physical Well-being ranging from 0 to 10
Social Competence	Social Competence ranging from 0 to 10
Emotional Maturity	Emotional Maturity ranging from 0 to 10
Language & Cognitive Dev	Language & Cognitive Development ranging from 0 to 10
Communication & General Kno	Communication & General Knowledge ranging from 0 to 10
Binary	
Vulnerability Status	Vulnerability Status as yes or no
MCI	Multiple Challenge Index as yes or no

Figure 3.1 shows a modified version of Gebbie's general ecological model of the determinants of health for the case of this study using a 3-level model of children's health.[40] It shows that the number of levels (i.e. here three) and the type and number of variables within each level were critical in determining child's health outcomes.

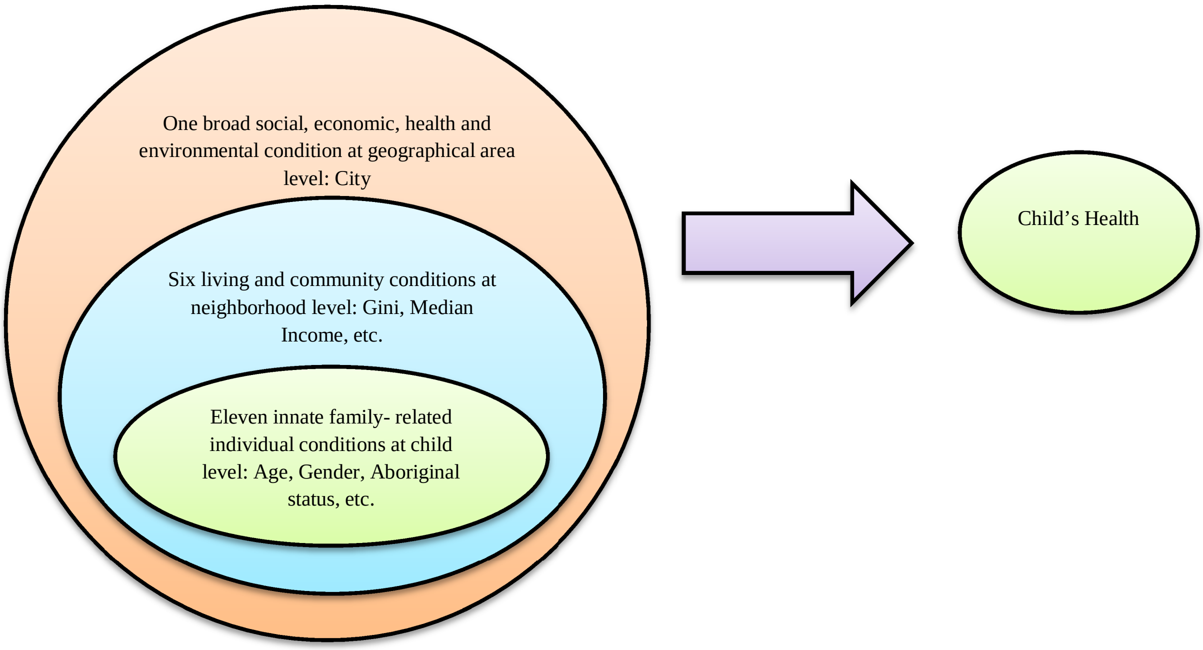

Figure 3.1: A Guide to Thinking About the Three-level Determinants of Child's Health

3.4. Measurements

3.4.1. Explanatory Variables

Referring to Table 3.1, one can observe that explanatory variables were divided into three groups: child characteristics, neighborhood characteristics and geographical area characteristics.

Level −1 variables or individual child characteristics were measured via the EDI 104 questionnaire by school teachers or educational assistants.

Level – 2 variables or neighborhood variables were measured as follows: the school type variable was measured by finding its corresponding school division at the Saskatchewan Ministry of Education Website (http://www.education.gov.sk.ca/school-division); and the variable Gini index was calculated based on median household income (before tax) for the population 15+ years of age. Here, in case of major cities of Saskatoon, Regina and Prince Albert, a Gini Index of zero mostly meant perfect equality while in case of non-urban areas a Gini Index of zero mostly meant only one sample of the Dissemination Area was used for calculating the neighborhood level Gini coefficient (which, as a result, artificially fixes Gini Index as zero). Neighborhoods with Gini Index zero constituted 12.5%, 3.3%, 10.0%, and 71.0% of all neighborhoods in Saskatoon, Regina, Prince Albert and non-urban areas, respectively. A median income of zero meant either at least half of the population 15^{+} years of age residing in that neighborhood had no income, or the original data in the Dissemination Area level was 0. Finally, the average dwelling value was calculated based on owner-occupied private non-farm, non-reserve dwellings and an Average Value of Dwelling of 0 meant that no one owned the dwelling in that neighborhood.

The level −3 geographical area variable was measured via list of child's residential postal codes in the data set and Canada Post's list of city-based postal codes available on the website (http://www.canadapost.ca/ cpotools).

3.4.2. Outcome Variables

Referring to Table 3.1, one can observe that outcome variables were divided into two groups, continuous and binary. The continuous variables were the 5 EDI domains in ranging from 0 to 10 that had been in the EDI 104 questionnaire and were measured by school teachers or educational assistants. The binary variables of vulnerability status and the Multiple Challenge Index were measured by statistical experts at the Saskatchewan Ministry of Education via finding the distribution of each of the 5 EDI domains.[6] For the case of vulnerability status, statistical experts determined whether a child's score had been in the bottom 10% of the site distribution in at least one domain. Furthermore, for the case of multiple vulnerability, they specified whether a child's score had been in the bottom 10% of the site distribution in at least three domains.[6]

3.5. The 3-level model

3.5.1. Statistical Background

In the social and health sciences, having three-level models is not uncommon. Fitting a two-level model to a three-level hierarchy has two negative consequences:[36] Firstly, it causes misattribution of the response variation to the remaining two levels' variables, and in its own order yields to erroneous conclusions about the relative importance of each level on variability of the response variable. For example, if one models a three level hierarchical data of geographical area-neighborhood-child via a two-level model neighborhood-child, then the variability in the response variable attributed to geographical area level variables will be absorbed by neighborhood level variables which cause an overestimation of neighborhood level variations. Secondly, such incorrect two-level modelling of a three-level hierarchy data produces biased standard errors for the coefficients of the explanatory variables in the model, changing their significance and hence yielding one to draw wrong conclusions on the relationship between each explanatory and the response variables. In the case of the above example, the neighborhood level variables can have higher standard errors, and some of them can lose their significance at the standard significance level of $p = 0.05$.

Referring to the three-level model of geographical area-neighborhood-child, there were $K = 4$ cities (Saskatoon, Regina, Prince Albert and non-urban areas) in the province of Saskatchewan and in the k^{th} $(1 \leq k \leq 4)$ geographical area there were J_k neighborhoods (Saskatoon: $J_1 = 65$, Regina: $J_2 = 30$, Prince Albert: $J_3 = 10$, and non-urban areas: $J_4 = 80$) and in the j^{th} neighborhood $(1 \leq j \leq J_k)$ there were $N_{j,k}$ children whose continuous and binary health-related outcome variables had been measured. In major cities of Saskatoon, Regina and Prince Albert the concept of neighborhood was operationalized according to the municipalities definition of neighborhood area within their city. However, in non-urban areas the concept of neighborhoods is not as easily operationalized. Therefore, in this study, any geographical entity outside of the three large cities mentioned (Saskatoon, Regina and Prince Albert) including cities with population less than 35,000, such as towns, villages, and resort villages were operationalized as 'neighborhoods' for non-urban areas. These non-Urban 'neighborhoods,' would tend to differ from the urban neighborhoods in important ways, such as population density, diversity and even urban design and structure. These differences in what is labeled as a 'neighborhood' between those of non-urban areas and the three cities is well understood by the researcher; however, for the purpose of this thesis, and in the interest of geographical inclusion (as opposed to exclusion), it was decided to proceed with the current four classes of 'neighborhoods' rather than excluding (and losing generalizability) about 50% of the population sample in this study.

As Figure 3.2 shows, level-1 of the aggregated data set is represented by the children, level-2 is represented by the neighborhoods in which the children are nested, and level-3 of the data set is represented by the geographical areas in which the neighborhoods are nested.

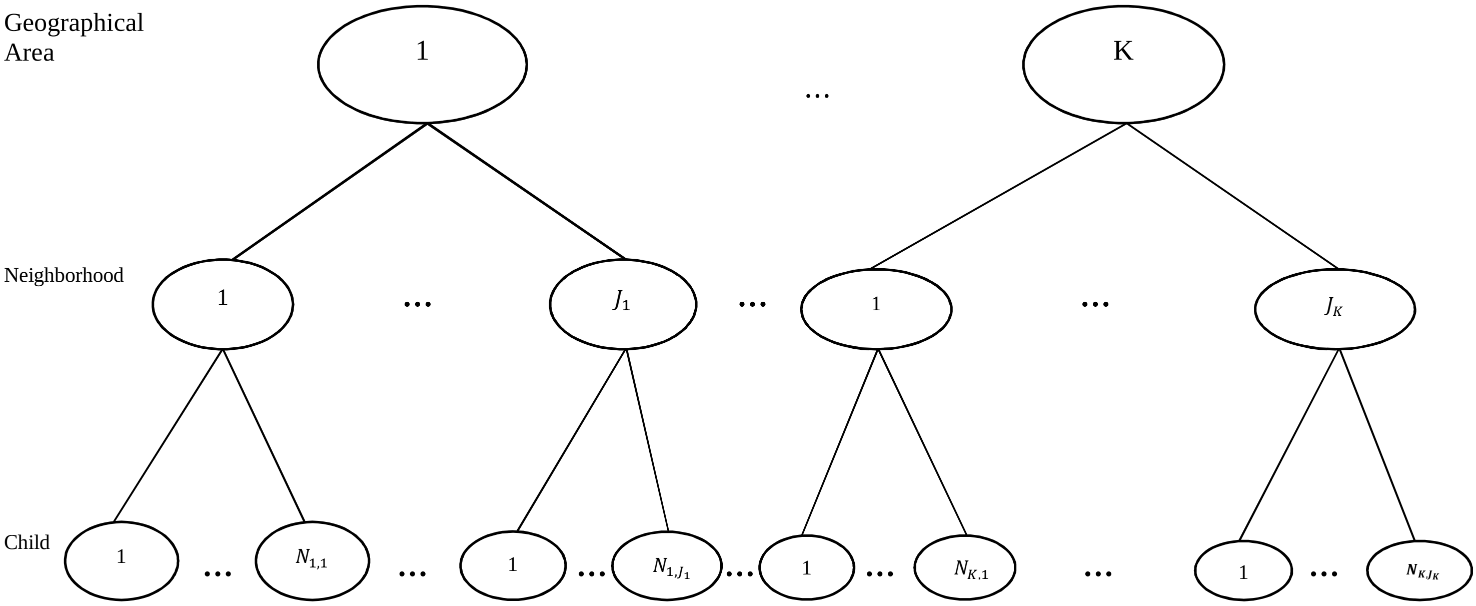

Figure 3.2: Data Structure for a Three-level (Geographical Area-Neighborhood-Child) Model

As an extension of two-level models, in three-level models level-1 intercepts and slopes may be random at levels 2 and 3 and level-2 intercepts and slopes may be random at level-3. From mathematical perspective, 3-level regression model for the above aggregated data was written in the form:[39]

Level 1: $$f\left(E\left(Y_{ijk}^{(n)}\right)\right) = \beta_{0jk} + \sum_{s=1}^{p} \beta_{sjk} X_{ijk}^{(s)}, \qquad (1 \leq n \leq 7) \qquad (3-1)$$

Level 2: $$\beta_{sjk} = \gamma_{s0k} + \sum_{t=1}^{q_s} \gamma_{stk} X_{jk}^{(t)} + e_{sjk}^{(2)} \qquad (s = 0, \dots, p)$$

Level 3: $$\gamma_{stk} = \delta_{st0} + \sum_{u=1}^{r_{st}} \delta_{stu} X_{k}^{(u)} + e_{stk}^{(3)} \quad (s = 0, \dots, p,\ t = 0, \dots, q_s)$$

where for continuous outcome variable of the five EDI domains $Y_{ijk}^{(n)} \sim \mathrm{N}\left(e_{ijk}^{(1)}, \sigma_{e(1)}^{2}\right), (1 \leq n \leq 5)$, *f(x)* = *x* was used while for binary outcome variable of vulnerability status or Multiple

Challenge Index $Y_{ijk}^{(n)} \sim$ Bernouli (π_{ijk}) where $\pi_{ijk} = Pr\left(Y_{ijk}^{(n)} = 1\right), (6 \le n \le 7)$, *f(x) = logit (x)* was used. In addition, with $q_s + 1$ fixed effects for each β_{sjk}, and $r_{st} + 1$ fixed effects for each γ_{stk}, the total number of fixed effects in the 3-level model (3-1) was $\sum_{s=0}^{p} \sum_{t=0}^{q_s+1} (r_{st} + 1)$. In this study, $p = 11,\ q_0 = 6, q_s \neq 0(\text{for some } s = 1, \dots, p), r_{00} = 1,$ and $r_{st} \neq 0$ $(\text{for some } s = 1, \dots, p, t = 1, \dots, q_s)$ were found. Note that the values $q_s, r_{st}(s, t > 0)$ referred to the cross-level interactions coefficients and $\sigma_{e(1)}^2 = Var\left(e_{ijk}^{(1)}\right) < \infty, \sigma_{e(2)}^2 = Var\left(e_{sjk}^{(2)}\right) < \infty$, and $\sigma_{e(3)}^2 = Var\left(e_{stk}^{(3)}\right) < \infty$. Similar to the two-level model, depending on whether the intercepts $\beta_{0jk}, \gamma_{s0k}$ or coefficients $\beta_{sjk}, \gamma_{stk}$ or both are random variables, the three-level model will be called the random intercept model, the random slope model or the random intercept and slope model, respectively. In addition, the random intercept and slope model is defined as:

$$\beta_{sjk} = \beta_s^{(0)} + u_{sjk}, \quad \text{and} \quad \gamma_{stk} = \gamma_{st}^{(0)} + v_{stk}(\text{s}, \text{t} > 0) \qquad (3-2)$$

where $\beta_s^{(0)}, \gamma_{st}^{(0)}$ are fixed numbers and u_{sjk}, v_{stk}, are random variable satisfying $E(u_{sjk}) = E(v_{stk}) = 0,$ and $Var(u_{sjk}) = \sigma_{u_s}^{\ 2} < \infty, Var(v_{stk}) = \sigma_{v_{st}}^{\ 2} < \infty$, for all j, k, s. Also, the model is simplified to random slope model if $u_{0jk} = v_{s0k} = 0$, and it is simplified to random intercept model if $u_{sjk} = v_{stk} = 0\ (s, t > 0)$. In this study, the random intercept model was used.

Let in the equation $(3-1)$ the two continuous variable “U” and dichotomous variable “V” with their associated interaction term “U*V” appear in the following form:

$$f(E(Y)) = \beta_U.U + \beta_V.V + \beta_{U*V}.(U * V) + \cdots. \qquad (3-3)$$

Then, for the linear case of $f(t) = t$, for each k units increase in the variable “U”, the outcome difference of “V=1” versus “V=0” is given by “$\beta_{U*V}.k$". Furthermore, for the logistic case of

$f(t) = logit(t)$, for each k units increase in the variable “U”, the outcome odds ratio of “V=1” versus “V=0” changes by “$(exp(\beta_{U*V}.k) - 1) \times 100\%$" . These conclusions were applied in the linear and continuous outcome related results and their interpretations.

Testing fixed effects and cluster effects is the generalization of the two-level models discussed in section 2.4.1. For example, referring to the model (3-1), in order to test the fixed effects the typical null hypothesis is:

$H_0: \beta_{sjk} = 0$, $H_0: \gamma_{stk} = 0$, and $H_0: \delta_{stu} = 0$,

with t-values of t= $\frac{\widehat{\beta_{sjk}}}{s.e(\widehat{\beta_{sjk}})}$ (df=$\sum_{i=1}^{K}\sum_{j=1}^{J_i} N_{i,j} - p - 1$), t= $\frac{\widehat{\gamma_{stk}}}{s.e.(\widehat{\gamma_{stk}})}$ (df=$\sum_{i=1}^{K} J_i - q_s - 1$), and t= $\frac{\widehat{\delta_{stu}}}{s.e.(\widehat{\delta_{stu}})}$ (df=K-$r_{st} - 1$), respectively, in which the numerator is the maximum likelihood estimate of the coefficient and the denominator is the estimated sampling standard error of the numerator. Next, in order to test the cluster effect we compare the 3-level model (3-1) with $u_{sjk}, v_{stk} \neq 0$ $(s = 0, \dots, p,\ t = 0, \dots, q_s)$ with the 3-level model (3-1) with $u_{sjk}, v_{stk} = 0$ $(s = 0, \dots, p,\ t = 0, \dots, q_s)$ using likelihood ratio tests. The null and alternative hypotheses are written as:

$H_0: {\sigma_{u_s}}^2 = {\sigma_{v_{st}}}^2 = 0$, (for all$s = 0, \dots, p,\ t = 0, \dots, q_s$),
$H_1: {\sigma_{u_s}}^2$, or ${\sigma_{v_{st}}}^2 > 0$, (for some$s = 0, \dots, p,\ t = 0, \dots, q_s$).

Attributing L_0 and L_1 to the likelihood values of the linear regression model (3-1) and the 3-level regression model (3-1), respectively, the LR statistics for testing above hypothesis is given by:

$$\text{LR} = (-2 \log (L_0)) - (-2 \log (L_1)),$$

which should be compared to a chi-squared distribution with degrees of freedom equal to the number of extra parameters in the 3-level model (3-1), that is, $X^2_{(\sum_{s=0}^{p}(q_s+1),\ 0.95)}$. Moreover, one can test the super-cluster effects (necessity of the highest hierarchy level) by testing the null hypothesis stating there are no geographical area effects by comparing the three-level geographical area-neighborhood-child model to the two-level neighborhood-child model. The null and alternative hypotheses and the LR statistics are modified forms of those discussed above for the cluster test, and the details are omitted.

The interpretation of intercepts and slopes of the equation (3-1) is presented regarding whether the outcome variable Y_{ijk} is continuous or binary.

Firstly, for the continuous outcome variable Y_{ijk}, the intercept δ_{000} measures the overall mean of Y_{ijk} across all geographical areas, neighborhoods and children. The mean value of Y_{ijk} for geographical area *k* called random intercept is $\gamma_{00k} = \delta_{000} + e^{(3)}_{00k}$. The random geography effect $e^{(3)}_{00k}$ is the difference between geographical area *k's* overall mean and the overall mean. The mean value of Y_{ijk} for neighborhood *j* nested in geographical area *k* called random intercept is $\beta_{0jk} = \gamma_{00k} + e^{(2)}_{0jk}$. The random neighborhood effect $e^{(2)}_{0jk}$ is the difference between such neighborhood's overall mean and the overall mean of geographical area *k*. Furthermore, the coefficients $\beta_{sjk}, \gamma_{stk}$ in the equation (3-2) are called random slopes having means ${\beta_s}^{(0)}, \gamma^{(0)}_{st}$, respectively, and variances ${\sigma_{u_s}}^2, {\sigma_{v_{st}}}^2$, respectively. The random variables $e^{(1)}_{ijk}$ are referred to as level-1 residuals, the random variables $e^{(2)}_{sjk}$ are referred to as level 2-residuals and the random variables $e^{(3)}_{stk}$ are referred to as level-3 residuals. A 95% coverage interval for geography effects is given by $(\delta_{000} - 1.96 * \sigma_{e(3)}, \delta_{000} + 1.96 * \sigma_{e(3)})$, meaning that 95% of geographical area effects are expected to lie in such range. Consequently, those geographical areas at the 97.5th

percentile of the geography distribution are estimated to score 3.92*$\sigma_{e(3)}$ points higher than those geographical areas at the 2.5 th percentile. Next, within geographical area *k*, a 95% coverage interval for neighborhood effects is given by the interval ($\gamma_{00k} - 1.96 * \sigma_{e(2)}, \gamma_{00k} + 1.96 * \sigma_{e(2)}$) meaning that within geographical areas 95% of neighborhood effects are expected to lie in such range. Consequently, those neighborhoods at the 97.5th percentile of the neighborhood distribution are estimated to score 3.92*$\sigma_{e(2)}$ points higher than those neighborhoods at the 2.5 th percentile. Finally, within geographical area *k*, neighborhood *j*, a 95% coverage interval for child effects is given by the interval ($\beta_{0jk} - 1.96 * \sigma_{e(1)}, \beta_{0jk} + 1.96 * \sigma_{e(1)}$), meaning that 95% of child related effects are expected to lie in such a range. Consequently, those children at the 97.5th percentile of the distribution are estimated to score 3.92*$\sigma_{e(1)}$ points higher than those children at the 2.5 th percentile. The Variance Partition Coefficient (VPC) values for geographical area level, neighborhood level and child level are given by: [36]

$$\text{VPC(geographical area)} = \frac{\sigma^2_{e(3)}}{\sigma^2_{e(1)}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}, \qquad (3-4)$$

$$\text{VPC(neighborhood)} = \frac{\sigma^2_{e(2)}}{\sigma^2_{e(1)}+\sigma^2_{e(2)}+\sigma^2_{e(3)}},$$

$$\text{VPC(child)} = \frac{\sigma^2_{e(1)}}{\sigma^2_{e(1)}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}.$$

In addition, the following three different pairing of children are possible:

$$\text{ICC(same geographical area-same neighborhood-same child)} = 1, \qquad (3-5)$$

$$\text{ICC(same geographical area-same neighborhood)} = \frac{\sigma^2_{e(2)}+\sigma^2_{e(3)}}{\sigma^2_{e(1)}+\sigma^2_{e(2)}+\sigma^2_{e(3)}},$$

$$\text{ICC(same geographical area)} = \frac{\sigma^2_{e(3)}}{\sigma^2_{e(1)}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}.$$

Note that for two children living in two different geographical areas, ICC=0, which is trivial.

Secondly, for the binary outcome variable Y_{ijk}, the intercept δ_{000} measures the overall log odds of the outcome $\{Y_{ijk} = 1\}$ across all geographical areas, neighborhoods and children. The log odds value of $\{Y_{ijk=1}\}$ for geographical area *k* called random intercept is $\gamma_{00k} = \delta_{000} + e_{00k}^{(3)}$. The random geography effect $e_{00k}^{(3)}$ is the difference between geographical areas k 's log odds and the overall log odds. The log odds value of $\{Y_{ijk} = 1\}$ for neighborhood *j* nested in geographical area *k* called random intercept is $\beta_{0jk} = \gamma_{00k} + e_{0jk}^{(2)}$. The random neighborhood effect $e_{0jk}^{(2)}$ is the difference between such neighborhood's log odds and the overall log odds of geographical area *k*. Furthermore, the coefficients $\beta_{sjk}, \gamma_{stk}$ in the equation (3-2) are called random slopes having means $\beta_s{}^{(0)}, \gamma_{st}^{(0)}$, respectively, and variances $\sigma_{u_s}{}^2, \sigma_{v_{st}}{}^2$, respectively. The random variable $e_{ijk}^{(1)}$ is referred to as level-1 residual. Each category of this residual has type I extreme-value distribution and since the standard logistic distribution can be written as the difference of two type I extreme-value random variables, it follows that the level-1 variance is the variance of standard logistic distribution given by $\frac{\pi^2}{3}$.[41,42] Similar to the continuous case, the random variable $e_{sjk}^{(2)}$ is referred to as level 2-residual and the random variable $e_{stk}^{(3)}$ is referred to as level-3 residual. A 95% coverage interval for geography effects is given by $(\delta_{000} - 1.96 * \sigma_{e(3)}, \delta_{000} + 1.96 * \sigma_{e(3)})$ meaning that 95% of geography effects are expected to lie in such a range. Consequently, those geographical areas at the 97.5th percentile of the geography distribution are estimated to have log odds of outcome 3.92*$\sigma_{e(3)}$ points higher than those geographical areas at the 2.5 th percentile. Next, within geographical area *k*, a 95% coverage interval for neighborhood effects is given by the interval $(\gamma_{00k} - 1.96 * \sigma_{e(2)}, \gamma_{00k} + 1.96 * \sigma_{e(2)})$ meaning that 95% of neighborhood log odds of outcome are expected to lie in such a range. Consequently, within geographical areas, those neighborhoods at the 97.5th percentile of the neighborhood distribution are estimated to have log odds of outcome 3.92*$\sigma_{e(2)}$ points higher than those neighborhoods at the 2.5 th percentile. The VPC values for geography-level, neighborhood-level and child-level are given by:

$$\text{VPC(geographical area)} = \frac{\sigma^2_{e(3)}}{\frac{\pi^2}{3}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}, \qquad (3-6)$$

$$\text{VPC(neighborhood)} = \frac{\sigma^2_{e(2)}}{\frac{\pi^2}{3}+\sigma^2_{e(2)}+\sigma^2_{e(3)}},$$

$$\text{VPC(child)} = \frac{\frac{\pi^2}{3}}{\frac{\pi^2}{3}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}.$$

Furthermore, following three different pairing of children are possible:

$$\text{ICC(same geographical area-same neighborhood)} = \frac{\sigma^2_{e(2)}+\sigma^2_{e(3)}}{\frac{\pi^2}{3}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}, \quad (3-7)$$

$$\text{ICC(same geographical area)} = \frac{\sigma^2_{e(3)}}{\frac{\pi^2}{3}+\sigma^2_{e(2)}+\sigma^2_{e(3)}}.$$

3.5.2. Modelling Process

The following steps were followed:

Step 1: We assumed that each of the 5 EDI domains (as the continuous outcome) followed a normal distribution, and each of the dichotomous outcome - vulnerability status and MCI followed a binomial distribution. Standard model building approach was used to select variables for multivariable models.[43-45] A three-level linear regression for each of the continuous outcome and a three-level logistic regression for each of the dichotomous outcome were utilized. Various bivariable models were fitted by taking one independent variable at a time.

Step 2: Now, based on the analysis conducted on step 1, independent variables with $p < 0.20$, and of biological and scientific importance were considered as candidate variables for the final multivariable multilevel model.[43]

Step 3: Multivariable multilevel linear and logistic regressions were conducted to determine significant independent variables for each of the 5 EDI outcome domains, vulnerability status and the Multiple Challenge Index.

Step 4: All statistically significant ($p < 0.05$) and biologically- scientifically important variables (regardless of their statistical significance) were kept in the final multivariable multilevel model.

Step 5: The importance of each variable in the final model was determined by whether that variable had been statistically significant ($p < 0.05$) in at least one of the 5 EDI multivariable multilevel linear regression models, two multivariable multilevel logistic regression models or it had been a biologically and scientifically important variable.

Step 6: The significance of each of 2-factor interaction terms were tested by including it individually in the final main multivariable multilevel model. In this study, merely 2-interaction terms of within-level type and cross-level type were considered. Considering standard model building strategies, all interaction terms were included as candidates in the final model if they had a p-value of < 0.20 for at least the multivariable multilevel model associated to one of the 5 EDI domains or one of 2 binary outcomes. All interaction terms were retained in the final model if they had a p-value of < 0.10 for at least the multivariable multilevel model associated to one of the 5 EDI domains or one of the two binary outcomes. In addition, confounding needed to be checked for in some cases as well. Confounding was checked for if the interaction was non-significant.

Step 7: The model was essentially based on the Chunk wise Method,[46] composed of four chunks (sets) of variables. These four chunks were the individual chunk (11 individual child variables), the neighborhood chunk (6 neighborhood variables), the geography chunk (1 geographical area variable) and the interaction chunk (5 within-level and 23 cross-level interactions).

Step 8: Model fit was evaluated the using Pearson chi-square test for each of the five final multilevel linear models and the Hosmer- Lemeshow test for each of the final two multilevel logistic models.

Step 9: Due to our sociology interest, a three level model was considered. However, we tested cluster effects and super cluster effects using an LR test, as well.

3.6. Software

In this study, SPSS software was used to collect EDI data. Other related data were gathered via Excel, and after conversion to SPSS files they were merged with primary EDI data contained in SPSS file. The final data set was in SPSS format and simultaneously was converted to SAS and STATA formats for operational data analysis.

The statistical analysis of this study was performed with SAS 9.3, STATA 11, and SPSS 17 software packages.

CHAPTER 4: RESULTS

The ones that matter the most are the children.
-Native American Proverb

This chapter presents the results of the study. Firstly, a general descriptive analysis of the outcome variables and predictors of interest is given and the univariate analyses of the 3-level linear and logistic models are conducted. Secondly, with special attention to Aboriginal status and geographical area, a brief statistical inference of means of 5 EDI domains, proportions of 2 binary outcomes and their relative differences are discussed. Thirdly, multilevel linear analysis with focus on the main effects, within-level and cross-level interactions is conducted answering the first two main study epidemiological questions. The answer of the third question is given by the following calculations of specific hierarchy level VPCs and ICCs. Fourthly, the chapter presents the related multilevel logistic analysis which is similar to the linear one above. Finally, it concludes with a brief discussion of the model statistically significant variables and fitness.

4.1. Descriptive Analysis

Descriptive analysis was divided into three parts, including Pearson inter-correlation of the 5 EDI outcomes, basic statistics for dependent variables, and basic statistics for independent variables.

Table 4.1 presents the correlations between the five EDI domains and the correlation between "gold standard" normative sample in parenthesis. The "gold standard" sample is an inclusive data set created in 1999 with Canada wide EDI data of 116,860 children at the age of 5 and included gender, all five EDI domains and children with no special needs.[47] As observed, all correlations in this study were significant ($p < 0.0001$) and were moderate except the high correlation between Social competence and Emotional maturity which is consistent with previous studies.[25]

In addition, correlation results from this sample were in the range of medium to high and similar to that of the correlations with the normative sample. The difference percentage (i.e. $\frac{A-B}{B} \times 100\%$) range from 1.5% in the correlation between Social competence and Emotional maturity to 15.8% for the correlation between Emotional maturity and Communication & general knowledge.

Table 4.1: Observed Bivariate Pearson Correlations, (correlations in the normative "gold standard"), and ((n-values)) for 5 EDI Outcome Variables

	Social Competence	Emotional Maturity	Language & Cognitive Development	Communication & General Knowledge
Physical-Well being	0.621* (0.590) ((9025))	0.513* (0.490) ((8993))	0.574* (0.530) ((8984))	0.641* (0.610) ((9024))
Social Competence		0.802* (0.790) ((8994))	0.635* (0.590) ((8985))	0.637* (0.570) ((9024))
Emotional Maturity			0.495* (0.460) ((8954))	0.521* (0.450) ((8991))
Language & Cognitive Development				0.673* (0.620) ((8986))

Note: * $P < 0.0001$

Finally, in each column from top down the linear or quadratic trend in the observed results is similar to the same trend in that of the correlations with the normative sample.

Table 4.2 presents descriptive statistics (mean, standard errors) for the 5 EDI domains and the "gold standard" normative sample in parenthesis. As observed from top down, the results for mean values followed the same decreasing-increasing-decreasing trend of normative sample mean values. Next, in all 5 EDI domains, Saskatchewan mean values were less than national normative sample results with difference percentage ranging from 1.2% for Social competence to 5.2% for Language & cognitive development. Furthermore, as observed from top-down, the results for standard errors followed the same increasing-decreasing-increasing trend of normative sample standard errors suggesting the consistency of this study data with that of the normative national sample. Finally, 30.0% of Saskatchewan children were deemed 'vulnerable', while 6.74% were deemed 'multiple-vulnerable' (as defined by Multiple Challenge Index).

Table 4.2: Observed Basic Statistics for Dependent Variables (n=9045)

Variable	Category	Mean ± s.e.	n%± s.e.%
Continuous			
Physical Well-Being		8.52 ± 0.0161 (8.79±0.0031)*	
Social Competence		8.19 ± 0.0201 (8.29±0.0051)	
Emotional Maturity		7.91 ± 0.0165 (8.05±0.0044)	
Language & Cognitive Development		7.92 ± 0.0212 (8.36±0.0053)	
Communication & General Knowledge		7.47 ± 0.0279 (7.73±0.0057)	
Binary			
Vulnerability Status	Yes		30.00±0.4818
	No		70.00±0.4818
Multiple Challenge Index (MCI)	Yes		6.74±0.2636
	No		93.26±0.2636

Note: * refers to corresponding gold standard statistics.

Table 4.3 presents the basic descriptive statistics for all hierarchical independent variables.

Firstly, the total sample included 9045 children nested in 185 neighborhoods, in which 65 neighborhoods were in Saskatoon, 30 neighborhoods were in Regina, 10 neighborhoods were in Prince Albert and finally 80 were nested in non-urban areas.

Secondly, referring to the child characteristics, all of children in this study were between 4.5 and 8 years old, with an average age of 5.69 years and an average of 3.80 days absent from school in the 2008-2009 school year. The percentage of boys and girls in the study were almost equal (male: 50.65%; female: 49.35%; standard error=0.53%). Non-Aboriginal children constituted almost five times the number of Aboriginal children in the study sample (Non-Aboriginal: 83.20%; Aboriginal: 16.80%). Most children in the study were native English Speakers (Non-ESL: 95.83%; ESL: 4.17%), and the percentage of children who had non-parental care was almost equal to those who had parental care (Non-parental care: 44.84%; Parental Care: 43.53%; standard error = 0.46%). The majority of children did not attend a language or religion class (Non-attendance: 68.71%; Attendance: 15.75%).

Thirdly, referring to neighborhood characteristics, the majority of children were enrolled in public schools (public: 69.43%; Non-public: 30.57%), and on average, 56.57% of their neighborhood community held at least a high school diploma. Children in the study were living in neighborhoods that had income inequality ranging from perfect equality (Gini = 0) to high inequality (Gini = 0.679); mean income inequality could be described as low (Gini = 0.127). In addition, neighborhood-level median income per capita ranged from 0 to $46640, with mean value of $ 25020. Neighborhood average value of dwelling ranged from 0 to $ 34271 with mean value of $ 12305.

Finally, referring to geographical area variable, almost half (47.53%) of all Saskatchewan children in this study were living in the three biggest cities in the province: Saskatoon (24.25%), Regina (21.29%), and Prince Albert (1.99%).

Table 4.3: Basic Statistics for Independent Variables (n=9045)

Variable	Category	n%	Mean ± s.e.	**Min**	**Max**
Child Characteristics					
Age			5.69 ± 0.0060	4.57	7.94
Days Absent			3.80 ± 0.1026	0.00	60.00
Number of Special Skills			0.44 ± 0.0200	0.00	7.00
Number of Special Problems			0.29 ± 0.0150	0.00	9.00
Gender	Female	49.35			
	Male	50.65			
Aboriginal Status	Aboriginal	16.80			
	Non-aboriginal	83.20			
Requirement of Special Needs	Yes	4.00			
	No	96.00			
Attendance at French/English Immersion School	Yes	12.13			
	No	87.87			
English as Maternal Language	Yes	95.83			
	No	4.17			
Non-Parental Care	Yes	44.84			
	No	43.53			
	Unspecified	12.63			
Attendance at Language/Religion Class	Yes	15.75			
	No	68.71			
	Unspecified	15.53			
Neighborhood Characteristics					
School Type	Public	69.43			
	Separate	29.71			
	Francophone	0.86			
Gini Index.			0.127± 0.0010	0	0.679
Median Income in $ 10,000 (per capita)			2.502± 0.0089	0	4.664
Unemployment Rate for People 15+ Years of Age (%)			5.399± 0.0413	0	36.660
% of People with at Least High School Education			56.570± 0.1055	8.115	95.745
Average Value of Dwelling in Real* $10,000.			12.305± 0.0637	0	34.271
Geographical Area Characteristics					
Geographical Area	Saskatoon	24.25			
	Regina	21.29			
	Prince Albert	1.99			
	Non-urban Areas	52.47			

Note: * The price that has been modulated from a nominal price by removing the effects of its general level changes over time such as inflation.

4.2. Univariate Analysis

Tables 4.4a and 4.4b present the association between selected covariates and 7 outcome variables: the 5 EDI domains and 2 derived variables of vulnerability and the multiple challenge index.

Referring to linear regression model, firstly, child level variables of “age”, “days absent from school”, “number of special skills”, “number of special problems”, “gender”, “Aboriginal status” and “non-parental care” were significant with each of the 5 EDI domains $(p < 0.001)$. The variable “English as second language” was significant with four of the EDI domains $(p < 0.05)$ and for the variable “Requirement of special needs” was significant with only two of the EDI domains of Social competence and Emotional maturity. The variables “French/English Immersion school attendance”, and “Language/Religion class attendance” were significant with two of the EDI domains of Physical well-being and Communication & general knowledge $(p < 0.05)$. Second, neighbourhood level variables “School Type” (with public school as its reference category), “Gini Index”, and “Rate of high school degree holders” were moderately significant with only one EDI domain of either Physical well-being, or Social competence or Language & cognitive development $(p < 0.10)$, while other variables such as “median income”, “unemployment rate”, and “average value of dwelling” had no significant associations with any of the 5 EDI domains. Finally, the variable “geographical area” (with non-urban areas as its reference category) was moderately significant with two of the EDI domains with Saskatoon moderately significant in Physical well-being and Language & cognitive development, Regina moderately significant in Social competence, Emotional maturity and Language & cognitive development and Prince Albert moderately significant in Social competence and Emotional maturity, $(p < 0.10)$.

Referring to logistic regression model, firstly, child level variables “days absent from school”, “number of special skills”, “number of special problems”, “gender”, “Aboriginal status”, “French/English immersion school attendance”, and “non-parental care” were significant for both binary outcomes $(p \leq 0.001)$, while other variables such as “age”, “English as maternal language” and “language/religion class attendance” were significant in one of the two binary outcomes $(p < 0.001)$. Secondly, all neighbourhood level variables except the variable “School Type” had no significant associations with either of the two binary outcomes. Finally, at the level of the geographical area, only “Regina” was significantly associated with both binary outcomes $(p < 0.005)$.

In selecting candidate variables for inclusion in the multivariable multilevel model building process, two criteria were considered: first, if a variable has a p-value < 0.200 for at least one of the 5 EDI domains or for either of the two binary outcomes, then the variable was retained as a candidate variable in all seven models in order to maintain consistency of inclusion in all the models. Second, of the neighborhood level variables “Median Income”, “Unemployment Rate”, and “Average value of Dwelling” though they were not statistically significant, they were retained in the modelling process due to their importance of revealing social contextual aspect of this study.

Table 4.4a. Univariate Analysis: Association Between each of the Covariates and 5 EDI domains, Vulnerability Status, and the Multiple Challenge Index (MCI)

Model	Linear					Logistic	
	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge	Vulnerability Status	MCI
Independent Variable							
Child Characteristics							
Age	0.304* (< 0.001)† [0.211,0.396]‡	0.216 (<0.001) [0.100,0.333]	0.133 (0.006) [0.038,0.228]	0.564 (<0.001) [0.430,0.685]	0.405 (<0.001) [0.244,0.566]	−0.411 (<0.001) [-0.555,-0.267]	−0.113 (0.384) [-0.369,0.141]
Days Absent	−0.045 (<0.001) [-0.050,-0.040]	−0.043 (<0.001) [-0.049,-0.036]	−0.031 (<0.001) [-0.036,-0.026]	−0.063 (<0.001) [-0.069,-0.056]	−0.069 (<0.001) [-0.077,-0.060]	0.065 (<0.001) [0.057,0.074]	0.047 (<0.001) [0.037,0.056]
Number of Special Skills	0.263 (<0.001) [0.232,0.295]	0.290 (<0.001) [0.250,0.330]	0.216 (<0.001) [0.184,0.249]	0.443 (<0.001) [0.402,0.484]	0.550 (<0.001) [0.496,0.605]	−0.471 (<0.001) [-0.545,-0.397]	−0.941 (<0.001) [-1.193,-0.683]
Number of Special Problems	−0.745 (<0.001) [-0.779,-0.710]	−1.055 (<0.001) [-1.098,-1.013]	−0.740 (<0.001) [-0.776,-0.704]	−1.003 (<0.001) [-1.049,-0.958]	−1.439 (<0.001) [-1.499,-1.381]	1.379 (<0.001) [1.282,1.477]	1.216 (<0.001) [1.126,1.306]
Gender(Male)	−0.416 (<0.001) [-0.477,-0.356]	−0.809 (<0.001) [-0.884,-0.735]	−0.801 (<0.001) [-0.861,-0.741]	−0.614 (<0.001) [-0.693,-0.535]	−0.884 (<0.001) [-0.992,-0.782]	0.824 (<0.001) [0.727,0.921]	0.907 (<0.001) [0.724,1.089]
Aboriginal Status	−0.063 (<0.001) [-0.091,-0.035]	−0.101 (<0.001) [-0.136,-0.066]	−0.070 (<0.001) [-0.099,-0.042]	−0.107 (<0.001) [-0.143,-0.070]	−0.102 (<0.001) [-0.151,-0.054]	0.978 (<0.001) [0.756,1.204]	0.928 (<0.001) [0.552,1.303]
Requirement of Special Needs	−0.004 (0.636) [-0.018,0.011]	−0.026 (0.005) [-0.044,-0.008]	−0.019 (0.013) [-0.034,-0.004]	−0.015 (0.113) [-0.035,0.04]	−0.021 (0.106) [-0.046,0.004]	0.019 (0.085) [-0.003,0.042]	0.017 (0.130) [-0.005,0.039]
French/English Immersion School Attendance	0.028 (0.008) [0.007,0.048]	0.020 (0.126) [-0.006,0.046]	−0.000 (0.976) [-0.021,0.022]	0.026 (0.063) [-0.001,0.053]	0.045 (0.013) [0.010,0.081]	−0.301 (<0.001) [-0.454,-0.147]	−0.961 (<0.001) [-1.328,-0.594]
English as Second Language	0.185 (0.020) [0.029,0.340]	−0.004 (0.138) [-0.344,0.048]	−0.171 (0.036) [-0.332,-0.011]	−0.511 (<0.001) [-0.715,-0.307]	-2.058 (<0.001) [-2.325,-1.191]	0.632 (<0.001) [0.409,0.854	0.315 (0.112) [-0.074,0.703]
Non-parental care	−0.003 (<0.001) [-0.004,-0.001]	−0.004 (<0.001) [-0.005,-0.003]	−0.002 (<0.001) [-0.003,-0.001]	−0.004 (<0.001) [-0.006,-0.003]	−0.006 (<0.001) [-0.008,-0.004]	0.005 (<0.001) [0.003,0.007]	0.005 (0.001) [0.002,0.007]
Language/Religion Class Attendance	−0.001 (0.005) [-0.002,-0.000]	−0.001 (0.127) [-0.002,0.000]	−0.001 (0.332) [-0.002,0.001]	−0.001 (0.263) [-0.002,0.001]	−0.033 (<0.001) [-0.005,-0.002]	0.003 (<0.001) [0.002,0.005]	−0.000 (0.744) [-0.003,0.002]

Note: * indicates the estimated coefficient in the 1-level model, † indicates the p-value and ‡ indicates the 95% CI.

Table 4.4b. Univariate Analysis: Association Between each of the Covariates and 5 EDI Domains, Vulnerability Status, and the Multiple Challenge Index (MCI)

Model	Linear					Logistic	
	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge	Vulnerability Status	MCI
Independent Variable							
Neighborhood Characteristics							
Separate School	0.031 (0.408) [-0.042,0.104]	−0.091 (0.051) [-0.183,0.000]	−0.026 (0.505) [-0.101,0.050]	−0.057 (0.243) [-0.153,0.039]	0.035 (0.587) [-0.092,0.128]	0.017 (0.763) [-0.094,0.128]	0.082 (0.413) [-0.114,0.278]
Francophone School	0.879 (<0.001) [0.537,1.222]	0.357 (0.104) [-0.073,0.787]	−0.064 (0.721) [-0.415,0.287]	0.149 (0.518) [-0.303,0.601]	0.012 (0.968) [-0.582,0.607]	−0.531 (0.080) [-1.127,0.064]	−0.417 (0.493) [-1.607,0.774]
Medium Gini	−0.267 (0.002) [-0.435,-0.100]	−0.070 (0.537) [-0.294,0.153]	−0.177 (0.052) [-0.356,0.001]	−0.084 (0.484) [-0.320,0.151]	−0.303 (0.031) [-0.578,-0.027]	0.373 (0.001) [0.154,0.594]	0.179 (0.336) [-0.186,0.545]
High Gini	−0.019 (0.827) [-0.192,0.153]	0.080 (0.485) [-0.146,0.305]	0.042 (0.655) [-0.143,0.227]	0.111 (0.357) [-0.126,0.349]	0.059 (0.681) [-0.228,0.344]	0.121 (0.296) [-0.106,0.350]	−0.043 (0.812) [-0.403,0.315]
Median Income in $10,000	0.049 (0.379) [-0.060,0.168]	0.007 (0.921) [-0.135,0.150]	−0.041 (0.576) [-0.152,0.071]	−0.027 (0.725) [-0.173,0.124]	0.095 (0.320) [-0.092,0.282]	−0.011 (0.899) [-0.173,0.150]	−0.056 (0.684) [-0.306,0.213]
Unemployment % of 15+ Years People	0.007 (0.497) [-0.012,0.026]	−0.007 (0.545) [-0.031,0.016]	−0.009 (0.363) [-0.028,0.010]	−0.002 (0.903) [-0.027,0.023]	−0.007 (0.676) [-0.039,0.025]	0.002 (0.892) [-0.025,0.029]	−0.018 (0.571) [-0.079,0.044]
% People with at Least High School Degree	0.004 (0.356) [-0.005,0.014]	0.010 (0.150) [-0.003,0.027]	0.002 (0.663) [-0.008,0.013]	0.017 (0.016) [0.003,0.030]	0.006 (0.515) [-0.012,0.023]	−0.008 (0.293) [-0.023,0.007]	−0.021 (0.116) [-0.047,0.005]
Average Value of Dwelling in $ 10,000	0.004 (0.584) [-0.011,0.020]	0.005 (0.677) [-0.018,0.028]	−0.003 (0.755) [-0.019,0.014]	−0.002 (0.866) [-0.027,0.014]	0.014 (0.384) [-0.017,0.044]	−0.012 (0.370) [-0.039,0.015]	−0.016 (0.403) [-0.054,0.022]
Geographical Area Characteristics							
Saskatoon	−0.143 (0.082) [-0.303,0.018]	0.013 (0.894) [-0.181,0.207]	−0.018 (0.832) [-0.187,0.150]	0.273 (0.008) [0.072,0.0475]	−0.131 (0.328) [-0.393,0.131]	0.191 (0.080) [-0.023,0.405]	−0.166 (0.334) [-0.502,0.171]
Regina	−0.127 (0.230) [-0.335,0.080]	−0.298 (0.019) [-0.548,-0.049]	−0.193 (0.082) [-0.411,0.025]	−0.458 (0.001) [-0.717,-0.199]	−0.303 (0.077) [-0.639,0.033]	0.437 (0.001) [0.170,0.703]	0.534 (0.005) [0.159,0.910]
Prince Albert	−0.114 (0.570) [-0.507,0.279]	0.438 (0.071) [-0.038,0.914]	0.345 (0.100) [-0.066,0.757]	−0.198 (0.435) [-0.694,0.299]	0.380 (0.247) [-0.264,1.025]	−0.103 (0.710) [-0.644,0.439]	−0.363 (0.437) [-1.277,0.552]

Note: * indicates the estimated coefficient in the 1-level model, † indicates the p-value and ‡ indicates the 95% CI.

4.3. Statistical Inference of Means, Proportions and Difference Percentages

Table 4.5 presents the average (95% CI) of the 5 EDI domains for selected child, neighborhood, and geographical area level variables.

First of all, referring to child level variables, girls had significantly higher average values in all 5 EDI domains than boys, and the difference percentages ranged from 4.9% in Physical well-being to 12.4% in Communication and general knowledge. Secondly, Aboriginal children had significantly lower average values in all 5 EDI domains than non-Aboriginal children, and the difference percentages ranged from 8.8% in Emotional maturity to 19.4 % in Communication and general knowledge. Figure 4.1 presents these results for Aboriginal children and non-Aboriginal children. It also shows a decreasing trend for Aboriginal children's scores from left to right while the trend for Non-aboriginal children is not monotonic.

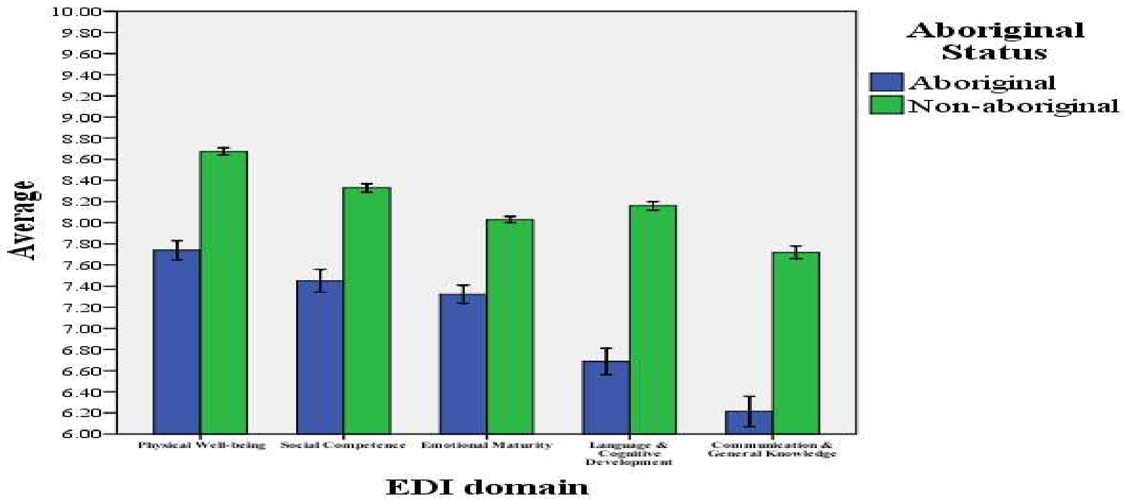

Figure 4.1. Average (95%CI) of 5 EDI Domains by Aboriginal Status

Thirdly, native English speaking children had higher average values in 4 EDI domains than ESL children and the difference percentages ranged from 1.7% in Social competence to 35.9% in Communication and general knowledge. Furthermore, in the domains of Language & cognitive development and Communication & general knowledge such differences were significant. Fourthly, children who had a parental care had higher average values in all 5 EDI domains than those with non-parental care and the difference percentages ranged from 0.1% in Emotional maturity to 6.9% Communication & general knowledge. In all three domains of Physical well-being, Language & cognitive development and Communication & general knowledge such differences were significant.

Next, referring to neighborhood level variables, as the neighborhood income inequality increased average child EDI outcome followed a quadratic (decreasing-increasing) statistically significant trend in all 5 EDI domains. Secondly, children in neighborhoods with Francophone schools had higher average values in all 5 EDI domains than children living in neighborhoods with public and separate schools. Such difference was significant in the two domains of Physical well-being and Social competence.

Table 4.5a: Average (95%CI) of 5 EDI Domains by Child Level, Neighborhood Level and Geographical Area level Variables

	Physical Well-being	Social Competence	Emotional Maturity	Language &Cognitive Development	Communication & General Knowledge
Gender					
Girls	8.72 [8.68,8.77]	8.59 [8.54,8.64]	8.31 [8.27,8.35]	8.22 [8.17,8.28]	7.91 [7.84,7.99]
Boys	8.31 [8.27,8.36]	7.79 [7.73,7.85]	7.52 [7.41,7.57]	7.62 [7.56,7.68]	7.04 [6.96,7.12]
Difference (%)	+4.9**	+10.2**	+10.5**	+7.9**	+12.4**
Race					
Aboriginal	7.74 [7.65,7.83]	7.45 [7.34,7.56]	7.32 [7.24,7.41]	6.69 [6.56,6.81]	6.21 [6.07,6.36]
Non-Aboriginal	8.67 [8.64,8.71]	8.33 [8.29,8.37]	8.03 [8.00,8.06]	8.16 [8.12,8.20]	7.72 [7.66,7.78]
Difference (%)	−10.7**	−10.6**	−8.8**	−18.0**	−19.4**
Maternal Language					
English	8.51 [8.48,8.54]	8.19 [8.15,8.23]	7.91 [7.88,7.95]	7.94 [7.89,7.98]	7.57 [7.51,7.62]
ESL	8.70 [8.56,8.84]	8.05 [7.85,8.24]	7.74 [7.59,7.90]	7.44 [7.22,7.67]	5.57 [5.24,5.90]
Difference (%)	−2.2**	+1.7	+2.2	+ 6.7**	+35.9**
Care					
Parental	8.62 [8.58,8.67]	8.28 [8.22,8.34]	7.94 [7.89,7.98]	8.13 [8.08,8.19]	7.78 [7.70,7.85]
Non-parental	8.47 [8.42,8.52]	8.17 [8.11,8.23]	7.93 [7.89,7.98]	7.79 [7.73,7.86]	7.28 [7.19,7.36]
Difference (%)	+1.8**	+1.3	+0.1	+4.4**	+6.9**

Note: Income inequality categories are based on 2 tertiles of Gini Index with ith tertile $T_i = 0.100, 0.159,$ respectively. ** refers to significant estimates at 5% or lower.

Table 4.5b: Average (95%CI) of 5 EDI Domains by Child Level, Neighborhood Level and Geographical Area Level Variables

	Physical Well-being	Social Competence	Emotional Maturity	Language &Cognitive Development	Communication & General Knowledge
Neighborhood Income Inequality					
Low (1st tertile)	8.59 [8.54,8.64]	8.19 [8.13,8.25]	7.95 [7.90,8.00]	7.96 [7.89,8.02]	7.54 [7.46,7.63]
Average (2ed tertile)	8.34 [8.28,8.40]	8.02 [7.94,8.09]	7.75 [7.69,7.81]	7.68 [7.60,7.77]	7.16 [7.05,7.26]
High(3rd tertile)	8.60 [8.55,8.65]	8.35 [8.28,8.41]	8.02 [7.96,8.07]	8.10 [8.03,8.17]	7.69 [7.59,7.78]
Neighborhood School Type					
Public	8.50 [8.46,8.54]	8.20 [8.15,8.25]	7.92 [7.88,7.96]	7.92 [7.87,7.97]	7.45 [7.38,7.51]
Separate	8.53 [8.47,8.59]	8.14 [8.06,8.21]	7.88 [7.82,7.95]	7.90 [7.82,7.97]	7.52 [7.42,7.62]
Francophone	9.45 [9.27,9.63]	8.66 [8.29,9.03]	7.94 [7.63,8.26]	8.23 [7.86,8.59]	7.63 [7.03,8.23]
Geographical Area					
Saskatoon	8.54 [8.48,8.61]	8.31 [8.24,8.39]	7.97 [7.90,8.03]	8.25 [8.17,8.33]	7.55 [7.44,7.66]
Regina	8.43 [8.36,8.50]	7.93 [7.84,8.02]	7.78 [7.71,7.86]	7.54 [7.44,7.64]	7.26 [7.13,7.38]
Prince Albert	8.47 [8.23,8.71]	8.60 [8.37,8.84]	8.26 [8.06,8.45]	7.69 [7.42,7.96]	7.95 [7.60,8.30]
Non-urban Areas	8.54 [8.50,8.59]	8.21 [8.16,8.27]	7.92 [7.88,7.97]	7.93 [7.87,7.98]	7.50 [7.43,7.57]

Note: Income inequality categories are based on 2 tertiles of Gini Index with ith tertile $T_i = 0.100,\ 0.159,$ respectively. ** refers to significant estimates at 5% or lower.

Finally, referring to geography level variable, the following results were concluded. Firstly, Saskatoon children had higher average values in all 5 EDI domains than Regina children, but the difference was not statistically significant in Physical well-being. However, Saskatoon children had higher average values in only 2 EDI domains (Physical well-being and Language &

cognitive development) than Prince Albert children, but the difference in the first domain was not statistically significant. In addition, Saskatoon children had higher average values in all 5 EDI domains than non-urban areas children and such difference was statistically significant only in the Language & cognitive development domain. Second, Regina children had lower average values in all 5 EDI domains than all other children; but the differences were not statistically significant in the two domains of Physical well-being and Communication & general knowledge. Thirdly, Prince Albert children had statistically significantly higher average values than non-urban areas children in 3 EDI domains of Social competence, Emotional maturity and Communication & general knowledge. Figure 4.2 summarizes the above results.

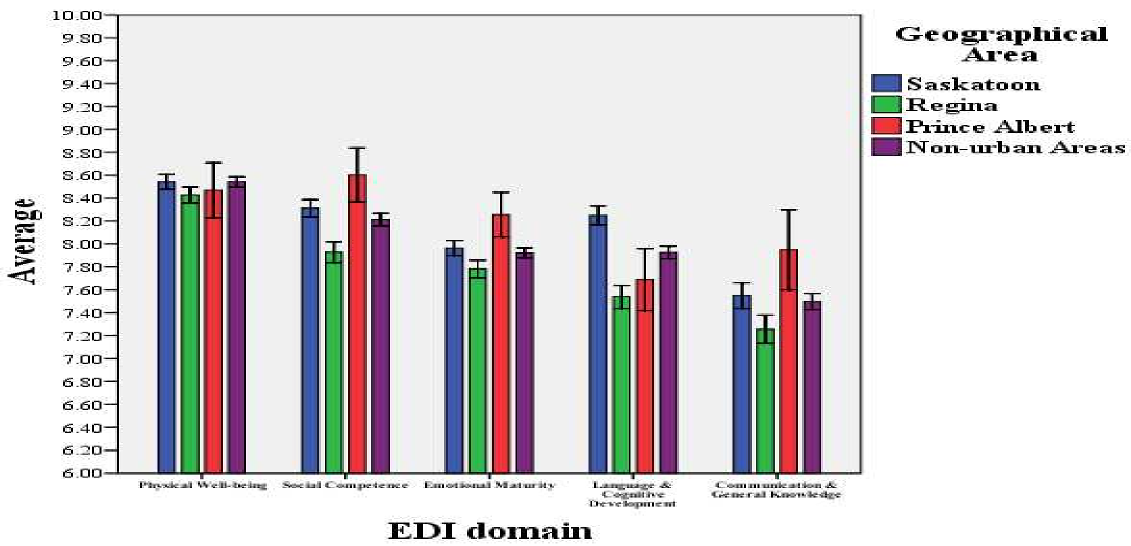

Figure 4.2. Average (95% CI) of 5 EDI Domains by Geographical Area

Tables 4.6a and 4.6b present proportions of vulnerable status (95% CI) by child, neighborhood, and geographical area level variables (%) for each EDI domains.

First of all, referring to child level variables, proportion of girls who were deemed vulnerable was lower all 5 EDI domains than boys. The difference percentages ranged from 38.0% in Physical well-being to 65.8% in Emotional maturity. Secondly, Aboriginal children had significantly higher vulnerable status proportions in all 5 EDI domains than non-Aboriginal children with the difference percentages ranging from 99.9% in Emotional maturity to 231.1% in Language & cognitive development. Thirdly, native English speaking children had significantly lower vulnerable status proportions in 2 EDI domains than ESL children. Difference percentages ranged from 37.2% in Language & cognitive development to 63.3% in Communication & general knowledge. Furthermore, in the first 3 EDI domains of Physical well-being, Social competence, and Emotional maturity native English speaking children had higher vulnerable status proportions than ESL children but such differences were not statistically significant. Fourthly, children who had parental care had significantly lower vulnerable status proportions in 4 EDI domains than those with non-parental care with percentage changes ranging from 19.9% in the Social competence to 36.8% in Language & cognitive development. However, in the domain of Emotional maturity, an inverse trend was present.

Secondly, referring to neighborhood level variables, as the neighborhood income inequality increased, the proportion of vulnerable status children in each of the 5 EDI domains followed a quadratic (increasing-decreasing) statistically significant trend. Second, children in neighborhoods with public schools had higher vulnerable status proportions in Physical well-being, Language & cognitive development and Communication & general knowledge than other children. In addition, children in neighborhoods with separate schools had higher vulnerable status proportions in Social competence and Emotional maturity domains than other children.

Table 4.6a: Proportion of Children Scoring ‘low’ in Each of 5 EDI Domains by Child Level, Neighborhood Level and Geographical Area Level Variables (%)

	Physical Well-being Low Status	Social Competence Low Status	Emotional Maturity Low Status	Language&Cognitive Development Low Status	Communication & General Knowledge Low Status
Gender					
Girls	9.38 [8.52,10.24]	6.96 [6.21,7.71]	5.85 [5.16,6.54]	8.68 [7.85,9.51]	11.14 [10.21,12.06]
Boys	15.14 [14.10,16.17]	15.10 [14.06,16.13]	17.13 [16.03,18.22]	14.13 [13.12,15.14]	19.14 [18.00,20.28]
Difference (%)	−38.0**	−53.9**	−65.8**	−38.6**	−41.8**
Race					
Aboriginal	25.96 [23.75,28.17]	19.43 [17.43,21.43]	19.79 [17.78,21.80]	26.46 [24.22,28.69]	29.62 [27.23,31.93]
Non-Aboriginal	9.55 [8.89,10.22]	9.39 [8.73,10.05]	9.90 [9.22,10.57]	8.45 [7.82,9.08]	12.30 [11.55,13.04]
Difference (%)	+171.8**	+106.9**	+99.9**	+231.1**	+132.7**
Maternal Language					
English	12.50 [11.80,13.20]	11.09 [10.42,11.76]	11.65 [10.97,12.33]	11.19 [10.52,11.86]	14.03 [13.29,14.77]
ESL	8.92 [6.01,11.82]	10.84 [7.67,14.01]	10.90 [7.71,14.09]	17.66 [13.90,21.95]	38.21 [33.25,43.17]
Difference (%)	+40.1**	+2.3	+6.9	−37.2**	−63.3**
Care					
Parental	10.46 [9.51,11.40]	9.74 [8.83,10.85]	11.98 [10.98,11.98]	8.67 [7.80,9.54]	11.74 [10.75,12.73]
Non-parental	13.79 [12.83,14.45]	12.17 [11.26,13.05]	11.20 [10.32,12.08]	13.71 [12.75,14.67]	18.11 [16.94,19.08]
Difference (%)	−24.1**	−19.9**	+6.9	−36.8**	−35.2**
Neighborhood Income Inequality					
Low (1st tertile)	10.76 [9.73,11.79]	10.32 [9.31,11.34]	10.85 [9.81,11.89]	10.42 [9.39,11.44]	14.07 [12.91,15.23]
Average (2ed tertile)	15.24 [13.90,16.58]	13.69 [12.40,14.97]	13.90 [12.61,15.19]	14.18 [12.88,15.48]	18.11 [16.67,19.55]
High(3rd tertile)	11.28 [10.12,12.4]	9.44 [8.36,10.52]	10.10 [8.99,11.22]	9.99 [8.88,11.10]	13.70 [12.43,14.97]

Table 4.6b. Proportion of Children Scoring 'low' in Each of 5 EDI Domains by Child Level, Neighborhood Level and Geographical Area Level Variables (%)

	Physical Well-being Low Status	Social Competence Low Status	Emotional Maturity Low Status	Language&Cognitive Development Low Status	Communication & General Knowledge Low Status
Neighborhood School Type					
Public	12.62 [11.80,13.44]	10.74 [9.97,11.50]	11.34 [10.56,12.13]	11.63 [10.83,12.42]	15.52 [14.63,16.42]
Separate	11.85 [10.63,13.08]	12.01 [10.78,13.24]	12.22 [10.98,13.47]	11.14 [9.94,12.33]	14.47 [13.14,15.80]
Francophone	1.30 [0.00,3.83]	6.49 [0.99,12.00]	5.19 [0.24,10.15]	6.58 [1.01,12.15]	12.99 [5.48,20.50]
Geographical Area					
Saskatoon	11.65 [10.31,13.00]	10.29 [9.01,11.56]	10.90 [9.59,12.21]	9.05 [7.85,10.25]	14.09 [12.63,15.65]
Regina	14.57 [12.99,16.51]	14.01 [12.46,15.56]	14.31 [12.74,15.88]	15.72 [14.08,17.35]	18.69 [16.94,20.43]
Prince Albert	12.22 [7.44,17.01]	6.67 [3.02,10.31]	7.22 [3.44,11.00]	11.67 [6.98,16.36]	12.78 [7.90,17.65]
Non-urban Areas	11.67 [10.76,12.58]	10.42 [9.55,11.29]	10.90 [10.01,11.79]	10.80 [9.92,11.69]	14.37 [13.37,15.37]

Note: Income inequality categories are based on 2 tertiles of Gini Index with ith tertile $T_i = 0.100,\ 0.159,$ respectively. ** indicates significant estimates at 5% or lower.

Finally, referring to geography level variable the following results were reached. Firstly, Saskatoon children had the lowest vulnerable status proportions in 2 EDI domains, Physical well-being and Language & cognitive development but such differences were not statistically significant. Secondly, Regina children had the highest vulnerable status proportions in all 5 EDI domains but such differences were not statistically significant in Physical well-being, Language & cognitive development and Communication & general knowledge domains. Thirdly, Prince Albert children had the lowest vulnerable status proportions in the 3 EDI domains of Social competence, Emotional maturity, and Communication & general knowledge but such differences were not statistically significant.

Table 4.7 presents proportions (95% CI) of selected child, neighborhood, and geographical area level variables (%) with vulnerability and multiple vulnerability, respectively.

First of all, referring to child level variables, girls had significantly lower proportions of vulnerability and multiple vulnerability than boys. The difference percentages ranged from 41.4% in vulnerability to 55.2% in multiple vulnerability. Secondly, Aboriginal children had significantly higher proportions of both vulnerability and multiple vulnerability than non-Aboriginal children with difference percentages ranging from 106.4% in vulnerability to 211.0% in multiple vulnerability. Figure 4.3 presents these results. Thirdly, native English speaking children had 37.5% lower proportion of vulnerability than ESL children, which is significant. Fourthly, children with parental care had significantly lower vulnerability and multiple vulnerability proportions than those with non-parental care. The difference percentages ranged from 19.0% in vulnerability to 24.0% in multiple vulnerability.

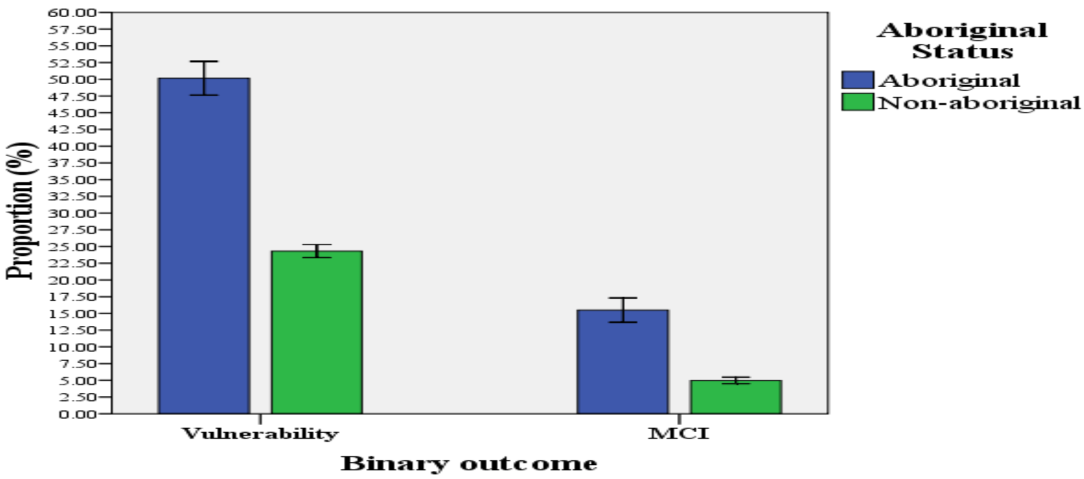

Figure 4.3. Proportion (95% CI) of Aboriginal Children and non-Aboriginal Children with Vulnerability and Multiple Challenge Index

Secondly, referring to neighborhood level variables as the neighborhood income inequality increased the proportion of children with vulnerability status and the proportion of children with multiple vulnerability followed a quadratic (increasing-decreasing) statistically significant trend.

Table 4.7: Proportion (95% CI) of Child Level, Neighborhood Level and Geographical Area Level Variables (%) with Vulnerability Status and Multiple Challenge Index

Variable	Vulnerability Status	MCI
Gender		
Girls	20.95 [19.76,22.18]	4.15 [3.56,4.17]
Boys	36.11 [34.72,37.50]	9.28 [8.44,10.12]
Difference (%)	-41.2^{**}	-55.2^{**}
Race		
Aboriginal	50.18 [47.61,52.65]	15.49 [13.67,17.31]
Non-Aboriginal	24.31 [23.34,25.29]	4.98 [4.49,5.47]
Difference (%)	$+106.4^{**}$	$+211.0^{**}$
Maternal Language		
English	27.89 [26.94,28.84]	6.66 [6.13,7.19]
ESL	44.59 [39.53,49.66]	8.63 [5.77,11.48]
Difference (%)	-37.5^{**}	-22.8
Care		
Parental	25.32 [24.05,26.73]	5.52 [4.82,6.23]
Non-parental	31.26 [29.97,32.54]	7.26 [7.00,8.48]
Difference (%)	-19.0^{**}	-24.0^{**}

Variable	Vulnerability Status	MCI
Neighborhood Income Inequality		
Low (1st tertile)	26.59 [25.11,28.06]	6.46 [5.64,7.28]
Average (2ed tertile)	33.59 [31.83,35.35]	8.56 [7.51,9.60]
High(3rd tertile)	26.25 [24.63,27.87]	5.32 [4.49,6.15]
Neighborhood School Type		
Public	28.50 [27.83,29.62]	6.72 [6.10,7.34]
Separate	29.25 [27.53,30.97]	6.89 [5.93,7.84]
Francophone	16.88 [8.52,25.25]	3.85 [0.00,9.11]
Geographical Area		
Saskatoon	28.15 [26.27,30.04]	4.88 [3.98,5.78]
Regina	34.13 [32.01,36.25]	9.09 [7.80,10.37]
Prince Albert	24.44 [18.17,30.72]	4.44 [1.43,7.46]
Non-urban Areas	26.77 [25.51,28.03]	6.74 [6.23,7.26]

Note: Income inequality categories are based on 2 tertiles of Gini Index with i^{th} tertile $T_i = 0.100, 0.159,$ respectively. ** indicates significant estimates at 5% or lower.

Next, the proportion of children in Francophone school type neighborhoods with vulnerability was significantly lower than other school type neighborhoods. In addition, the proportion of children with vulnerability in separate school type neighborhoods had no significant difference than the proportion of children with vulnerability in public school type neighborhoods. Similar conclusions were found from the previous result when vulnerability was replaced with multiple vulnerability.

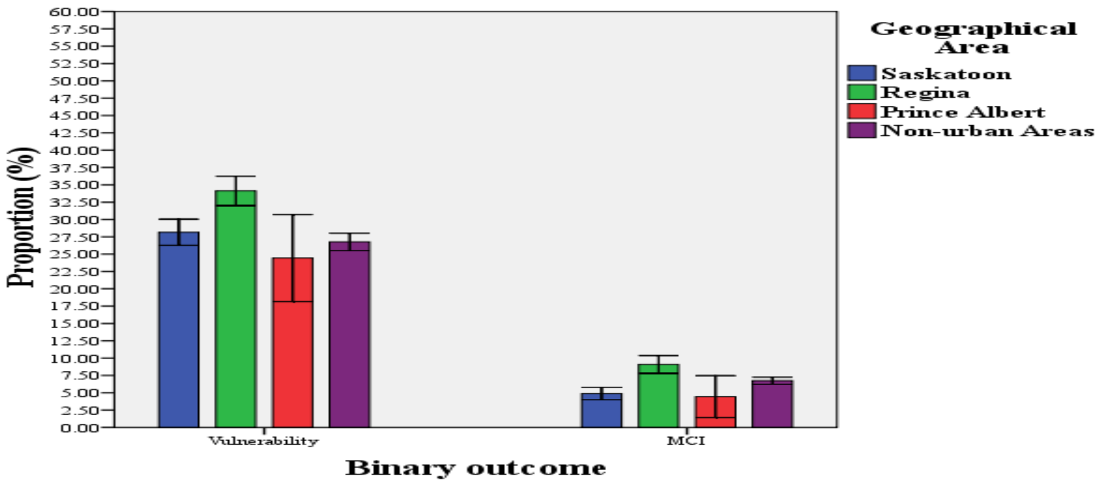

Figure 4.4. Proportion (95% CI) of Geographical Area with Vulnerability and Multiple Challenge Index

Finally, referring to the geographical area level variable the following results were concluded. Firstly, Saskatoon had the second largest proportion of vulnerable whilst it had the second lowest proportion of multiple vulnerable. Secondly, Regina had significantly the highest proportion of vulnerable children, and significantly the highest proportion of multiple vulnerable children. Thirdly, Prince Albert had the lowest proportion of vulnerable children and the lowest proportion of multiple vulnerable children. Finally, Non-urban Areas had the second lowest proportion of vulnerable children and the second highest proportion of multiple vulnerable children. Figure 4.4 summarizes the above observations.

4.4. Multilevel Linear Analysis

Tables 4.8a, 4.8b, and 4.8c, present the multilevel generalized linear model used in this study, which estimates main effects, within-level interaction and cross-level interaction terms, respectively, at the child, neighborhood, and geography level variables. In the following subsections, the effects of main effects, within-level interaction and cross-level interactions will be presented.

4.4.1. Main Effects

Table 4.8a. depicts $\widehat{\beta}$ $(p - value)$ results for the main effects of 5 EDI domains based on multilevel linear regression model including the child level variables of 'number of special problems', 'requirement of special need', 'French/English Immersion school attendance', 'ESL status', 'Language/Religion class attendance' and the neighborhood level variables of 'Unemployment % of people 15+years of age', '% of People with High School Diploma' and 'Average value of dwelling'.

First of all, having special problems had a significantly negative impact on all 5 EDI outcomes and each additional problem significantly decreased the scores associated with each EDI outcome, ranging from 0.633 units for Emotional maturity to 1.336 units for Communication & general knowledge.

Secondly, special needs requirements had a negative weak impact on all 5 EDI outcomes, and in two domains of Language & cognitive development and Communication & general knowledge such negative impact was significant.

Thirdly, attending at French or English Immersion school had no statistically significant impact on the developmental health outcomes. In addition, such effect was weak (max 0.016 units) and, is practically negligible.

Fourthly, native English speaking children had higher average scores in 4 EDI domains than ESL children and in the two domains of Language & cognitive development and Communication & general knowledge such differences were significant. In particular, for the later domain, the native English speaking children had 2.143 units difference from the ESL children. However, ESL children had significantly higher scores than native English speaking children in Physical well-being such difference was 0.147 units.

Fifthly, attending a language or religion class had a very weak impact (only 0.001 units) on all 5 EDI outcomes and, its effect is practically negligible.

Sixthly, a higher neighborhood unemployment rate had negative impact on all 5 EDI domains, and in the two domains of Social competence and Emotional maturity, such a negative impact was significant. However, such effect was weak (max 0.014 units) and, its effect is practically negligible.

Seventhly, despite the fact that a higher neighborhood percentage of high school educated people had a positive impact on 4 EDI domains, such effect was very weak (max 0.008 units), and it is practically negligible.

Table 4.8a. Main Effects $\widehat{\beta}$ $(p - value)$ Based on Multilevel Linear Regression Model for 5 EDI Domains.

Dependent Variable	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge
Independent Variables					
Child Characteristics					
Age	+0.268* (0.056)†	+0.567** (0.001)	+0.400** (0.005)	+0.426** (0.016)	+0.429* (0.063)
Days absent	−0.027** (<0.001)	−0.029** (<0.001)	−0.021** (<0.001)	−0.043** (<0.001)	−0.032** (<0.001)
Number of Special Skills	+0.179** (<0.001)	+ 0.207** (<0.001)	+0.156** (<0.001)	+ 0.312** (<0.001)	+0.414** (<0.001)
Number of Special Problems	−0.689** (<0.001)	−0.963** (<0.001)	−0.633** (<0.001)	−0.901** (<0.001)	−1.336** (<0.001)
Gender(Male)	−0.826* (0.096)	−0.101 (0.866)	−0.228 (0.651)	−1.670** (0.007)	−1.302 (0.111)
Aboriginal Status	−0.622** (<0.001)	−0.426** (<0.001)	−0.472** (<0.001)	−0.838** (<0.001)	−0.853** (<0.001)
Requirement of Special Needs	−0.011 (0.223)	−0.017 (0.140)	−0.012 (0.181)	−0.022* (0.063)	−0.026* (0.089)
French/English Immersion School Attendance	+0.004 (0.664)	−0.001 (0.954)	−0.012 (0.181)	+0.001 (0.932)	+0.016 (0.277)
English as Second Language	+0.147** (0.050)	−0.110 (0.226)	−0.095 (0.214)	−0.634** (<0.001)	−2.143** (<0.001)
Non-Parental Care	−0.001 (0.420)	−0.002 (0.281)	−0.001 (0.405)	−0.003 (0.143)	−0.003 (0.216)
Language/Religion class Attendance	−0.001** (0.004)	−0.001 (0.246)	−0.001 (0.268)	+0.001 (0.229)	−0.001* (0.092)
Neighborhood Characteristics					
Separate School	−0.016 (0.699)	−0.224** (<0.001)	−0.089** (0.021)	−0.236** (<0.001)	−0.120* (0.080)
Francophone School	+0.835** (<0.001)	+0.484* (0.063)	−0.013 (0.951)	+0.378 (0.161)	+1.040** (0.003)
Medium Gini	−0.190 (0.708)	+0.891 (0.146)	+0.906* (0.098)	+1.207** (0.036)	+0.662 (0.422)
High Gini	+0.380 (0.451)	+0.216 (0.723)	−0.021 (0.968)	+1.070* (0.058)	+1.401* (0.087)

Table 4.8 a (Continued)

Median Income in $ 10,000.	-0.100^{**} (0.050)	-0.161^{**} (0.009)	-0.067 (0.213)	-0.038 (0.520)	-0.143^{*} (0.087)
Unemployment % of People 15+ Years of Age	-0.002 (0.707)	-0.014^{*} (0.040)	-0.012^{**} (0.044)	-0.009 (0.197)	-0.009 (0.352)
% of People with at Least High School Degree	-0.001 (0.699)	$+0.008^{**}$ (0.039)	$+0.007^{*}$ (0.051)	+0.005 (0.148)	+0.001 (0.907)
Average Value of Dwelling in Real $ 10,000	-0.004 (0.652)	+0.001 (0.951)	-0.002 (0.796)	-0.017 (0.138)	-0.007 (0.646)
<u>Geography Characteristics</u>					
Prince Albert City	-1.554 (0.142)	-0.027 (0.983)	+0.602 (0.595)	-0.431 (0.722)	-1.098 (0.524)
Regina City	-0.340 (0.623)	-0.872 (0.296)	-0.771 (0.301)	-0.977 (0.213)	-0.994 (0.377)
Saskatoon City	-1.102^{**} (0.011)	-0.940^{*} (0.070)	-0.487 (0.290)	-0.524 (0.290)	-1.548^{**} (0.027)

Note: † Indicates p-values. ** indicates the coefficient estimates significant at 5% level or lower, and * indicates the level of significance at 10% or lower.

Finally, although a higher neighborhood average value of dwelling had a negative impact on 4 EDI domains, such effect was weak (max 0.017 units), and it is practically negligible.

4.4.2. Within-Level Interactions

Table 4.8b depicts within-level interactions $\widehat{\beta}\ (p - value)$ based on multilevel linear regression model for 5 EDI domains. The child level variables of ‘Aboriginal status’, ‘gender’, and ‘gender’ modify the effects of other child level variables, such as ‘number of special skills’, ‘non-parental care’, and ‘age’ , respectively. In addition, the neighborhood level variable of ‘Gini income inequality’ modifies the effect of the other neighborhood level variable, i.e., ‘Median Income’.

First of all, an increased number of skills had a significantly positive effect on all 5 EDI outcomes and, in addition, for each additional skill the outcome difference between Aboriginal children and non-Aboriginal children significantly decreased in all 5 EDI domains ranging from 0.046 units in Emotional maturity to 0.223 units in the Language & cognitive development. Therefore, scores of children who have more skills regardless whether they were of Aboriginal status, caught those of non-Aboriginal children, but the associated gap widened.

Figure 4.5 shows the predicted means (95% CI) for the 5 EDI domains by number of special skills and Aboriginal status. As observed, in all 5 EDI domains the predicted mean of Aboriginal children is significantly lower than that of non-Aboriginal children within the lowest number of skills; however, among children with higher number of skills such differences become non-significant. At the highest number of skills their magnitude decrease and in 3 EDI domains the effects become positive, meaning that Aboriginal children had better outcomes compared to others.

Secondly, non-parental care status had negatively weak effect on all 5 EDI outcomes. Furthermore, comparing non-parental care status to parental care status, the outcome difference between males versus females had a weak association (max 0.001 units decrease in 4 EDI domains) and, therefore, is negligible.

Thirdly, child's male status had negative effect on all 5 EDI outcomes; such effect was significant in 2 EDI domains of Physical well-being and Language & cognitive development. However, for each additional year of age, the outcome difference of males versus females increased within range of 0.078-0.093 units for the 2 EDI domains of Social competence and Emotional maturity. Hence, in those EDI domains, for older children, not only did males have lower scores than females, but the gap widened as well.

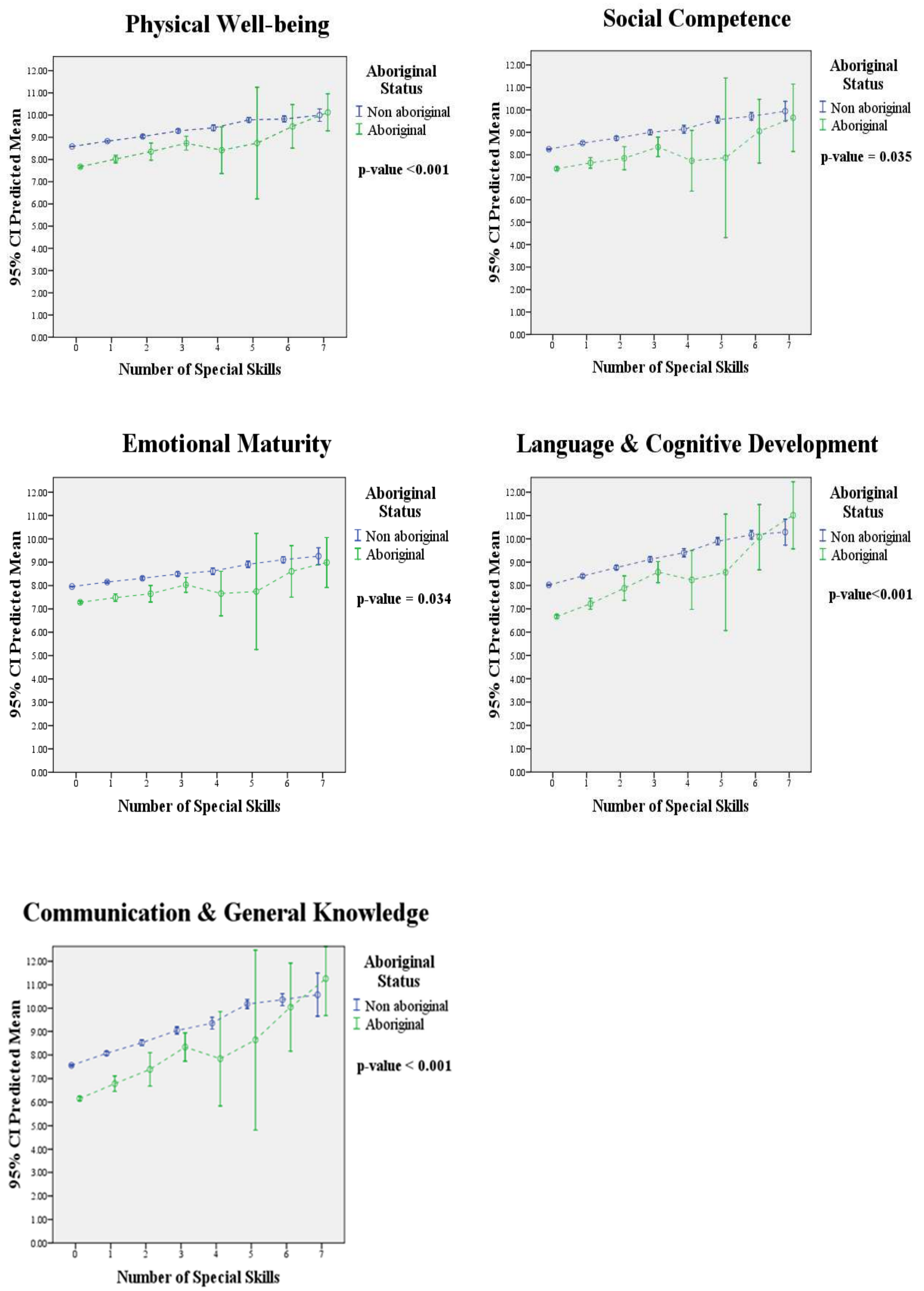

Figure 4.5. Predicted Means (95% CI) of 5 EDI Domains by Number of Special Skills and Aboriginal Status

Finally, higher neighborhood median income and higher neighborhood income inequality had compounding negative effects in 4 EDI domains. Specifically, for each additional $10,000 increase in neighborhood median income, the outcome difference of children living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality increased (increased) with a range of 0.155-0.309 units (range of 0.012-0.460 units) for 4 EDI domains. Consequently, by increasing neighborhood median income, not only did children in neighborhoods with medium (high) level income inequality have lower health scores than those in neighborhoods with low level income inequality, but the gap between them widened as well.

Table 4.8b: Within-Level Interactions $\widehat{\beta}\ (p-value)$ Based on Multilevel Linear Regression Model for 5 EDI Domains

Dependent Variable	Physical Well-being	Social Competence	Emotional Maturity	Language &Cognitive Development	Communication & General Knowledge
Independent Variable					
Within Level Interactions					
Within Level One					
Aboriginal*Number of Skills	$+0.105^{**}$ (<0.001)†	$+0.055^{**}$ (0.035)	$+0.046^{**}$ (0.034)	$+0.223^{**}$ (<0.001)	$+0.215^{**}$ (<0.001)
Gender *Non-parental care	-0.001 (0.512)	-0.001 (0.127)	-0.001 (0.296)	-0.001 (0.468)	+0.001 (0.530)
Gender*Age	+0.090 (0.296)	-0.093 (0.376)	-0.078 (0.373)	$+0.210^{*}$ (0.054)	+0.112 (0.432)
Within Level Two					
Medium Gini*Median Income	+0.129 (0.477)	-0.206 (0.347)	-0.309 (0.114)	-0.287 (0.163)	-0.155 (0.600)
High Gini*Median Income	-0.139 (0.492)	-0.012 (0.959)	+0.011 (0.960)	-0.313 (0.168)	-0.460 (0.163)

Note: † indicates are p-values. ** indicates the coefficient estimates significant at 5% level or lower, while * indicates level of significance at 10% or lower.

4.4.3. Cross-Level Interactions

Table 4.8c depicts estimated cross-level associations $\widehat{\beta}\ (p-value)$ based on multilevel linear regression model for 5 EDI domains .The neighborhood level variables of 'Gini Income Inequality', 'Gini Income Inequality' , and 'School Type' modify the effect of child level variables of 'Absenteeism days', 'Aboriginal status', and 'absenteeism days', respectively. In addition, the geography variables of ‘major city’, and ‘major city’ modify the effect of child level variables of 'Aboriginal status’ and ‘Non-parental care’, respectively. Furthermore, neighborhood level variables “Gini Income Inequality’ and ‘Median Income’ modify the effect of geography level variables ‘major city’ and ‘major city’, respectively.

First of all, days absent from school had a significantly negative weak effect on all 5 EDI domains and, furthermore, for each additional week absent from school outcome differences of children living in neighborhoods with medium (high) income inequality compared with children living in neighborhoods with low income inequality increased within a range of 0.091-0.182 units (0.084-0.119 units) in the 3 EDI domains of Physical well-being, Language & cognitive development, and Communication & general knowledge. Therefore, not only did higher neighborhood income inequality have adverse negative compounding effects with days absent from school, but the gap widened as well. Also, for each additional week of absence from school, the outcome difference of children studying at separate (Francophone) schools versus those studying at public schools significantly decreased (increased) within a range of 0.063-0.224 units (0.014-0.105 units) in 4 EDI domains. Thus, regarding more days absent from school, the gap between children at separate schools and children at public schools narrowed, while the gap widened between those of Francophone schools and those of public schools.

Secondly, Aboriginal status had a significantly negative effect on all 5 EDI domains and, in addition, the outcome differences of aboriginal children versus non-Aboriginal children in neighborhoods with medium (high) income inequality increased within a range of 0.008-0.293 units (0.050-0.137 units) in 5 (4) EDI domains. Therefore, higher neighborhood income inequality had detrimental effects on Aboriginal children’s scores and widened the gap between their scores compared with those of non-Aboriginal children. Figure 4.6 shows Predicted Means (95% CI) of 5 EDI domains by neighborhood income inequality and Aboriginal status. As observed, the decrease in average EDI outcome in Aboriginal children is slightly sharper than that of non-Aboriginal children. This trend is clear for the three domains of Social competence, Language & cognitive development and Communication & general knowledge. However, the lines are almost parallel due to non-significant interactions and we kept them in the model due to their sociological importance.

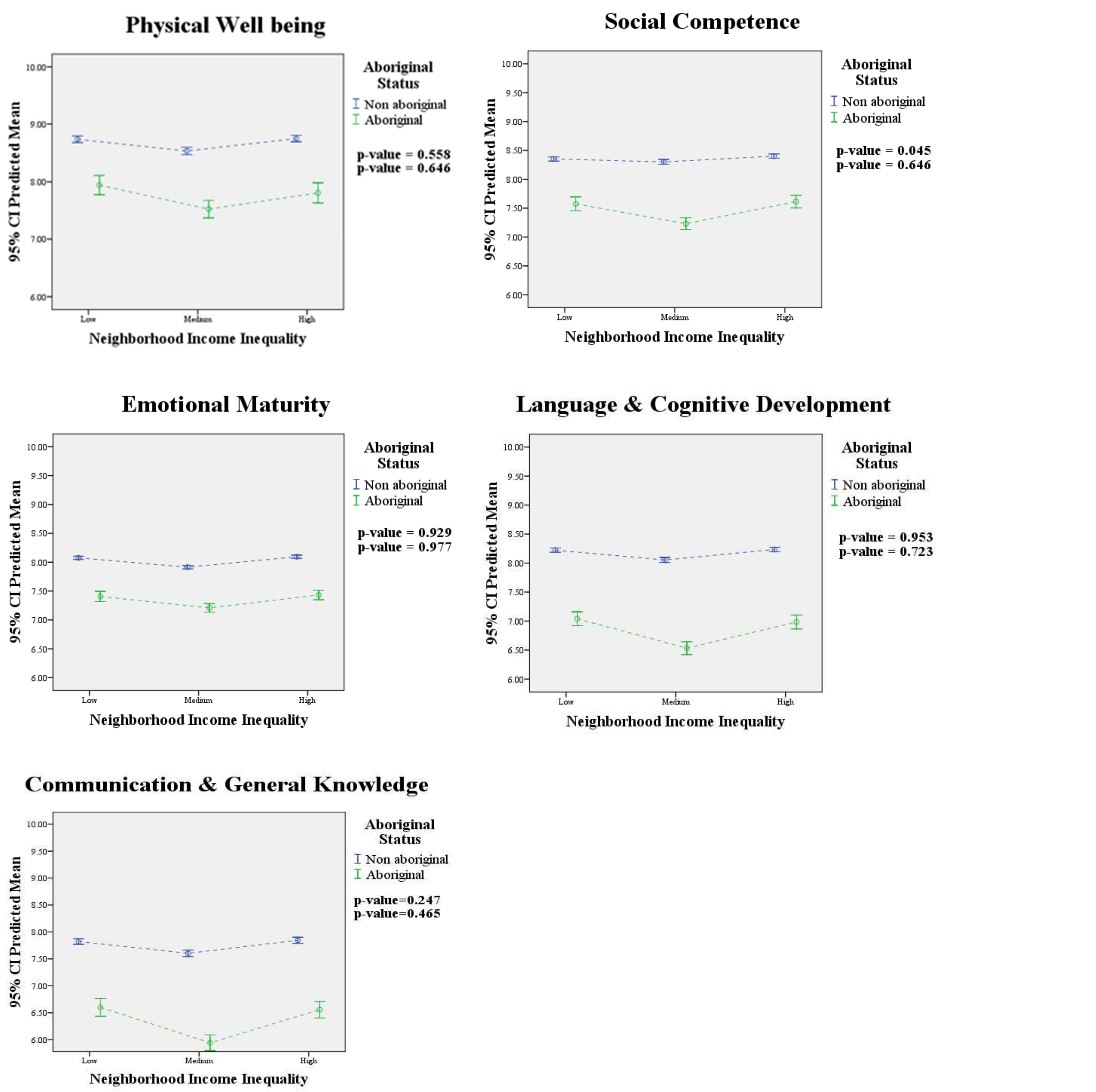

Figure 4.6. Predicted Means (95% CI) of 5 EDI Domains by Neighborhood Income Inequality and Aboriginal Status

Thirdly, non-parental care had a negatively weak effect on all 5 EDI domains, and comparing children without parental care to those with parental care, the outcome differences of those living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality had a weak change of max 0.006 units in 5 EDI domains. Therefore, the compound effect of child parental care status and residential neighborhood income inequality was negligible. Similarly, comparing children without parental care to those with parental care, the outcome differences of children living in cities (Prince Albert, Regina, or Saskatoon) versus those living in non-urban areas had a very weak change of max 0.007 units in the 5 EDI domains, and hence, compounding effect of child parental care status and major city is negligible.

Fourthly, as observed above, Aboriginal status had a significantly negative effect on all 5 EDI outcomes. However, child's major city can exacerbate or mitigate such negative effect. Specifically, for children living in Prince Albert, the outcome differences of Aboriginal ones versus non-Aboriginal ones increased within a range of 0.008-0.311 units compared with non-urban area children in 4 EDI domains. In contrast, for children living in Regina, the outcome differences of Aboriginal ones versus non-Aboriginal ones decreased within a range of 0.278-0.476 units compared with non-urban area children in 4 EDI domains. Also, for children living in Saskatoon, the outcome differences of aboriginal ones versus non-Aboriginal ones decreased and reversed within a range of 0.167-0.480 units compared with non-urban area children in all 5 EDI domains. Consequently, in comparing living in Prince Albert, which had a detrimental effect on outcome differences between Aboriginal children versus non-Aboriginal children, living in Regina and Saskatoon had a positive buffering effect on children.

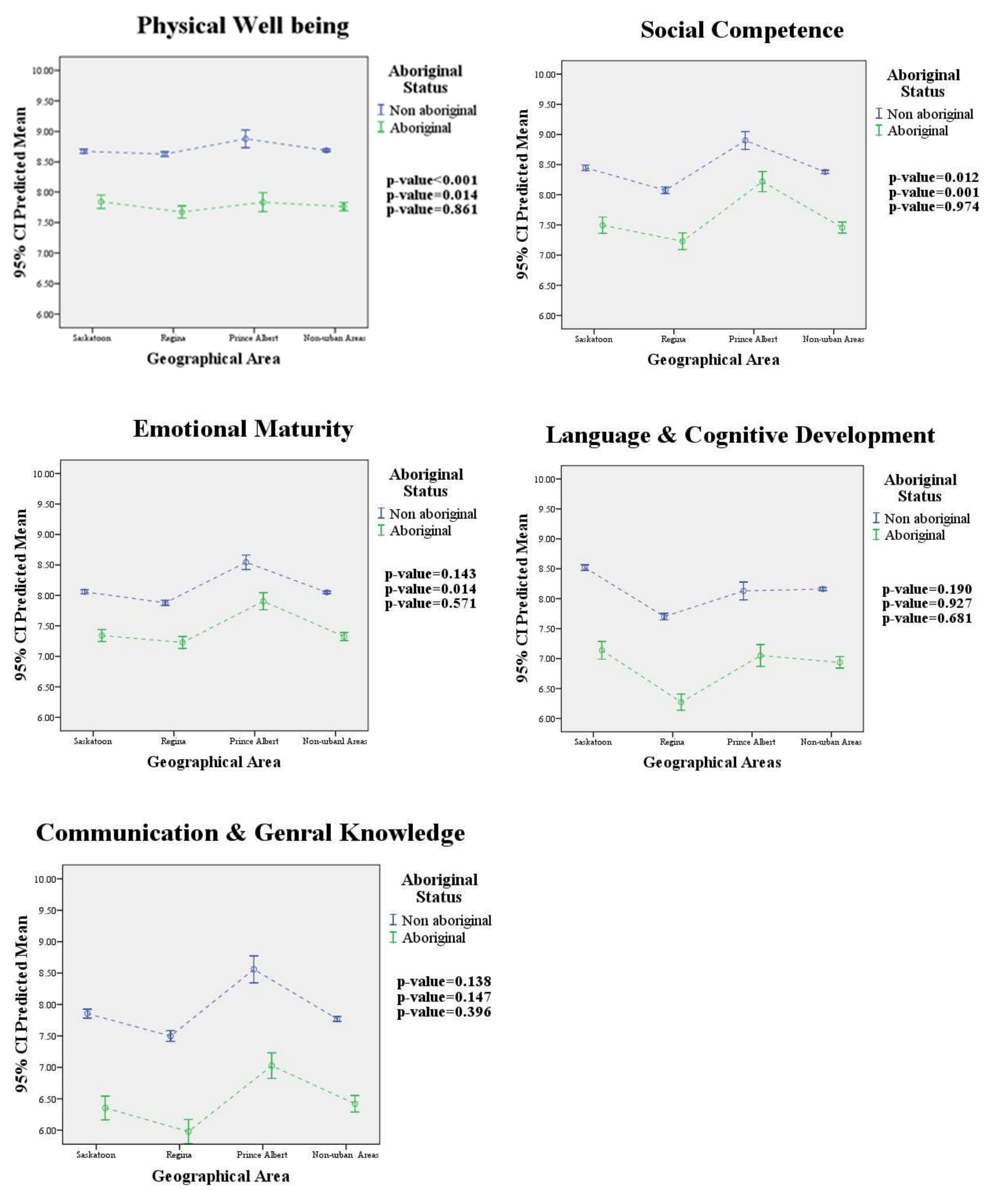

Figure 4.7. Predicted Means (95% CI) of 5 EDI Domains by Geographical Area and Aboriginal Status

Figure 4.7 shows predicted means (95% CI) of the 5 EDI domains by geographical area and Aboriginal status. As observed, in all domains, Regina children had the lowest predicted mean values for both Aboriginal and non-Aboriginal groups, while Prince Albert children had the highest predicted mean values for such groups.

Fifthly, children living in Prince Albert had lower scores compared with non-urban areas children in 4 EDI domains, and, in addition, compared to non-urban areas children, the outcome differences of those living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality increased within range of 0.496-1.194 units (0.132-0.863 units) in all 5 EDI domains. Also, children living in Regina had lower scores compared to non-urban areas children in all 5 EDI domains. However, compared to non-urban areas children, the outcome differences of those living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality increased within range of 0.569-0.745 units (0.138-0.348 units) in 3 EDI domains. Such negative compounding effects of higher neighborhood income inequality-major city existed for the city of Saskatoon, too. Children living in Saskatoon had lower scores compared to non-urban areas children in all 5 EDI domains, and in addition, compared to non-urban areas children, the outcome differences of those living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality increased within range of 0.059-0.381 units (0.085-0.160 units) in 4 EDI domains (2 EDI domains). Figure 4.8 presents predicted means (95% CI) of the 5 EDI domains by geography and neighborhood income inequality. As observed, compared to non-urban areas children, children living in Prince Albert had the highest outcome fluctuations as neighborhood income inequality changes in all 5 EDI domains; however, Saskatoon children had the lowest outcome fluctuations as neighborhood income inequality changed in all 5 EDI domains.

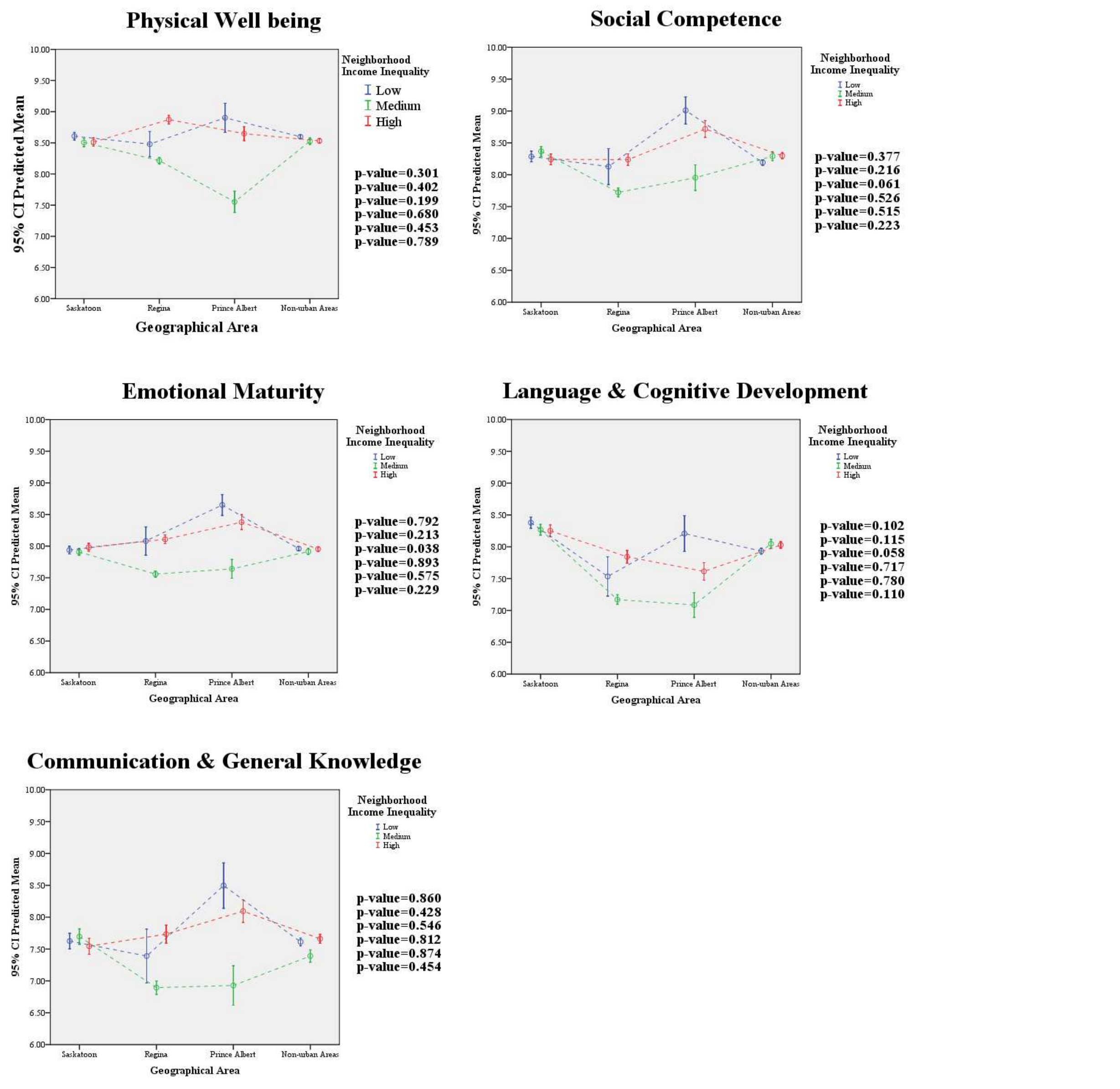

Figure 4.8. Predicted Means (95% CI) of 5 EDI Domains by Geographical Area and Neighborhood Income Inequality

Finally, although children living in Prince Albert had lower scores compared to non-urban area children in 4 EDI domains, the outcome differences of Prince Albert children versus non-urban areas children decreased within a range of 0.324-0.855 units for each additional $10,000 increase in neighborhood median income in those 4 EDI domains. Also, despite the fact that children living in Regina had lower scores compared to non-urban areas children in all 5 EDI domains, the outcome differences of Regina children versus non-urban areas children decreased within a range of 0.156-0.471 units for each additional $10,000 in neighborhood median income, for all 5 EDI domains. Finally, for the city of Saskatoon, despite of the fact that children living in this city had lower scores compared with non-urban areas children in all 5 EDI domains, the outcome differences of Saskatoon children versus non-urban areas children decreased within a range of 0.191-0.678 units for each additional $10,000 in neighborhood median income in all 5 EDI domains. In 4 EDI domains, such buffering compounding effects were significant.

Table 4.8c: Cross-Level Interactions $\widehat{\beta}$ $(p - value)$ Based on Multilevel Linear Regression Model for 5 EDI Domains

Dependent Variable	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge
Independent Variable					
Cross-Level Interactions					
Medium Gini*Days Absent	-0.013^{**} $(0.024)^{\dagger}$	-0.011 (0.106)	-0.006 (0.282)	-0.022^{**} (0.003)	-0.026^{**} (0.008)
High Gini*Days Absent	-0.012^{**} (0.027)	+0.001 (0.801)	+0.002 (0.711)	-0.017^{**} (0.013)	-0.017^{*} (0.062)
Medium Gini*Aboriginal	-0.070 (0.558)	-0.293^{**} (0.045)	-0.010 (0.929)	-0.008 (0.953)	-0.230 (0.247)
High Gini*Aboriginal	-0.052 (0.646)	-0.102 (0.646)	+0.003 (0.977)	-0.050 (0.723)	-0.137 (0.465)
Separate School*Days Absent	$+0.009^{**}$ (0.038)	$+0.017^{**}$ (0.002)	$+0.014^{**}$ (0.002)	$+0.032^{**}$ (<0.001)	$+0.015^{**}$ (0.045)
Francophone School*Days Absent	+0.015 (0.306)	+0.010 (0.565)	+0.006 (0.694)	+0.002 (0.911)	-0.066^{**} (0.008)
Medium Gini*Non-parental care	+0.001 (0.546)	$+0.006^{**}$ (<0.001)	$+0.003^{**}$ (0.023)	+0.001 (0.556)	-0.001 (0.422)
High Gini*Non-parental care	$+0.003^{**}$ (0.024)	$+0.004^{**}$ (0.004)	$+0.003^{**}$ (0.018)	-0.001 (0.436)	+0.001 (0.566)
Prince Albert*Aboriginal	+0.039 (0.861)	-0.008 (0.974)	-0.128 (0.571)	-0.115 (0.681)	-0.311 (0.396)
Regina*Aboriginal	$+0.288^{**}$ (0.014)	$+0.476^{**}$ (0.001)	$+0.293^{**}$ (0.014)	-0.013 (0.927)	+0.273 (0.147)

Table 4.8 c (Continued)

Saskatoon*Aboriginal	+0.480** (<0.001)	+0.327** (0.012)	+0.167 (0.123)	+0.175 (0.190)	+0.260 (0.138)
Prince Albert*Non-parental care	+0.003 (0.143)	+0.007** (0.026)	+0.004* (0.093)	+0.007** (0.021)	+0.002 (0.527)
Regina*Non-parental care	+0.001 (0.159)	−0.002 (0.381)	−0.001 (0.403)	+0.001 (0.839)	−0.000 (0.921)
Saskatoon*Non-parental care	+0.001 (0.159)	−0.000 (0.779)	+0.001 (0.111)	+0.005** (<0.001)	+0.001 (0.515)
Prince Albert*Medium Gini	−0.679 (0.199)	−1.194* (0.061)	−1.180** (0.038)	−1.147* (0.058)	−0.496 (0.546)
Regina*Medium Gini	−0.353 (0.402)	−0.627 (0.216)	−0.569 (0.213)	−0.745 (0.115)	−0.541 (0.428)
Saskatoon* Medium Gini	−0.214 (0.301)	−0.220 (0.377)	−0.059 (0.792)	−0.381 (0.102)	+0.059 (0.860)
Prince Albert*High Gini	−0.132 (0.780)	−0.697 (0.223)	−0.612 (0.229)	−0.863 (0.110)	−0.576 (0.454)
Regina* High Gini	+0.333 (0.453)	−0.348 (0.515)	−0.269 (0.575)	−0.138 (0.780)	+0.114 (0.874)
Saskatoon*High Gini	+0.086 (0.680)	−0.160 (0.526)	+0.030 (0.893)	−0.085 (0.717)	+0.080 (0.812)
Prince Albert*Median Income	+0.650* (0.052)	+0.335 (0.405)	+0.091 (0.798)	+0.324 (0.393)	+0.855 (0.115)
Regina*Median Income	+0.156 (0.524)	+0.374 (0.204)	+0.381 (0.151)	+0.380 (0.165)	+0.471 (0.235)
Saskatoon*Median Income	+0.435** (0.003)	+0.432** (0.014)	+0.191 (0.221)	+0.433** (0.009)	+0.678** (0.004)
Constant	+8.007** (<0.001)	+6.359** (<0.001)	+6.929** (<0.001)	+7.199** (<0.001)	+9.130** (<0.001)
p-value for LR test	<0.001	<0.001	<0.001	<0.001	<0.001
Number of Observations	8027	8025	7994	8007	8024
Pearson χ^2/df	1.740	2.573	1.814	2.696	4.704

Note: † indicates the p-values. ** indicates the coefficient estimates significant at 5% level or lower, while * indicates level of significance at 10% or lower.

4.4.4. Child-level, Neighborhood-level and Geographical Area-level Variances

Table 4.9 presents variance, variance partition coefficient and intra-class coefficients at geography, neighborhood, and child levels for all 5 EDI domains.

Table 4.9: Geographical Area Level, Neighborhood Level, Child Level Variance, VPC, and ICC Values of Multilevel Linear Model

	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge
$\sigma^2_{e(3)}$	0.075	0.092	0.077	0.095	0.124
$\sigma^2_{e(2)}$	0.411	0.494	0.456	0.440	0.664
$\sigma^2_{e(1)}$	1.275	1.550	1.293	1.602	2.096
VPC(geographical area)	0.043	0.043	0.042	0.044	0.043
VPC (neighborhood)	0.233	0.231	0.250	0.206	0.230
VPC (child)	0.724	0.726	0.708	0.750	0.727
ICC(geographical area)	0.043	0.043	0.042	0.044	0.043
ICC (neighborhood)	0.276	0.274	0.292	0.250	0.273

First of all, as the hierarchy level increased, the level associated variance given by all variables in that level decreased in all 5 EDI domains. One reason of such trend can be attributable to decreasing the number of subjects at levels from 9045 children in level-1 to 185 neighborhoods in level-2 and finally to only 4 geographical area in level-3 of the hierarchy. Partition of total variance by child, neighborhood and geographical area level is shown in Figure 4.9.

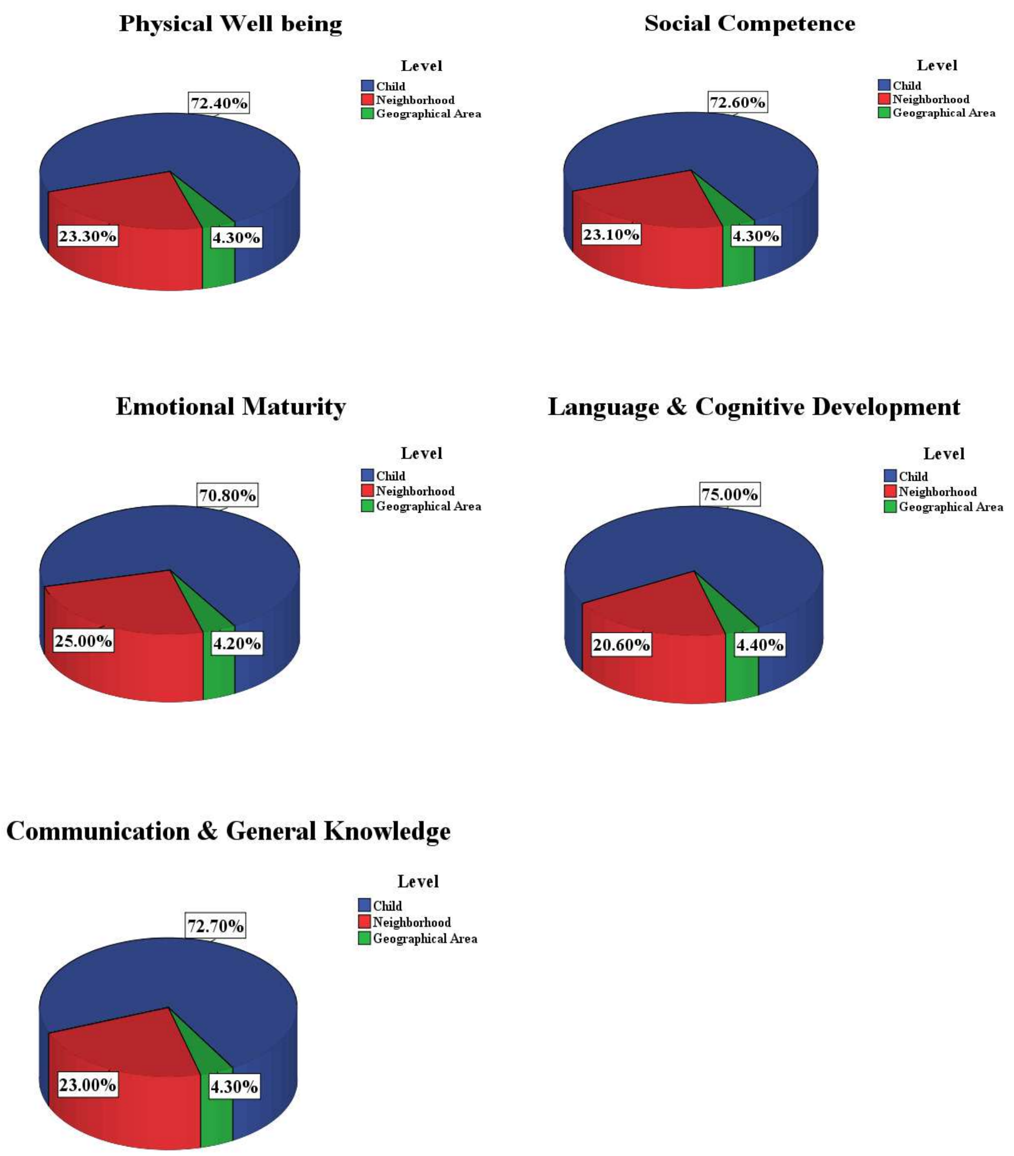

Figure 4.9: Variance Partition by Child, Neighborhood and Geographical Area Level of 5 EDI Domains

Secondly, geographical areas at the 97.5 th percentile of geography distribution were estimated to score 1.074 units higher than geographical areas at the 2.5 th percentile in the Physical well-being domain. For the other 4 EDI domains from left to right, such corresponding values were 1.189 units, 1.088 units, 1.208 units and 1.380 units, respectively. Next, within geographical areas, neighborhoods at the 97.5 th percentile of neighborhood distribution were estimated to score 2.513 units higher than neighborhoods at the 2.5 th percentile in Physical well-being. For the other 4 EDI domains from left to right, such corresponding values were 2.755 units, 2.647 units, 2.600 units and 3.194 units, respectively. Finally, within neighborhoods, children at the 97.5 th percentile of distribution were estimated to score 4.426 units higher than children at 2.5 th percentile in Physical well-being. For the other 4 EDI domains from left to right, such corresponding values were 4.880 units, 4.457 units, 4.962 units and 5.675 units, respectively.

Thirdly, in regards to the VPC statistics, 4.2%-4.4% of the variations in the 5 EDI outcomes lied between geographical areas; 20.6%-25.0% lied within geographical areas between different neighborhoods, while the remaining 70.8%-75.0% lied within neighborhoods between children. About three-quarters of all 5 EDI outcomes variations were attributable to the children themselves.

Finally, looking at ICC statistics, the correlation between two children living in the same geographical area but different neighborhoods was within the range of 0.042-0.044 in all 5 EDI domains, and the correlation between two children living in the same geographical area and same neighborhood was within the range of 0.250-0.292 in all 5 EDI domains. Therefore, children living in the same neighborhoods had much more (approximately 6 times) similar health outcomes in all 5 EDI domains than children living in adjacent neighborhoods.

4.5. Multilevel Logistic Analysis

Table 4.10 presents multilevel generalized logistic model estimates for individual child, neighborhood, and geography level main effects, within-level interaction and cross-level interaction terms. In the following subsections the effects of each type will be dealt.

4.5.1. Main Effects

Similar to the multilevel linear case, Table 4.10 depicts the $\widehat{\beta}\ (p - value)$ results for main effects of two binary outcomes based on multilevel logistic model including child level variables of ‘number of special problems’, ‘special needs requirement ’, ‘French/English immersion school attendance’, ‘ESL status’ , ‘Language/Religion class attendance’ and the neighborhood level variables of ‘Unemployment % of People 15+ years of age, ‘% of People with a High School Diploma’ and ‘Average value of dwelling ‘ as main predictors without interaction terms.

First of all, having special problems had a significantly exacerbating effect on both child’s vulnerability and multiple vulnerability and each additional problem significantly multiplied child’s odds of vulnerability and multiple vulnerability by 4.023 and 3.384, respectively.

Secondly, special needs requirement had an exacerbating weak impact on both child’s vulnerability and multiple vulnerability, and such status multiplied child’s odds of vulnerability and multiple vulnerability by 1.022 and 1.057, respectively.

Thirdly, attending at a French or English immersion school had a buffering impact on child’s vulnerability and multiple vulnerability, respectively. Furthermore, it decreased the child’s odds of vulnerability and multiple vulnerability by 3.7% and 60.2%, respectively.

Fourthly, ESL status had a significantly intensifying effect on both child’s vulnerability status and multiple vulnerability. In addition, the odds of vulnerability and multiple vulnerability in ESL children were 2.643 times and 1.886 times of the associated odds of native English speaking children, respectively.

Table 4.10: Main Effects, Within-Level Interactions and Cross-Level Interactions $\widehat{\beta}\ (p - value)$ Based on Multilevel Logistic Regression Model for Binary Outcomes of Vulnerability and MCI

Variable	Vulnerability Status	MCI	Variable	Vulnerability Status	MCI
Child Characteristics					
Age	-0.765^{**} $(0.014)^{\dagger}$	-0.974^{*} (0.083)	Days Absent	$+0.040^{**}$ (<0.001)	$+0.046^{**}$ (<0.001)
Number of Special Skills	-0.431^{**} (<0.001)	-0.867^{**} (<0.001)	Number of Special Problems	$+1.392^{**}$ (<0.001)	$+1.292^{**}$ (<0.001)
Gender(Male)	+0.244 (0.817)	+0.249 (0.891)	Aboriginal	$+1.052^{**}$ (<0.001)	$+0.783^{**}$ (0.006)
Requirement of Special Needs	+0.022 (0.498)	$+ 0.055^{*}$ (0.082)	French/English Immersion School Attendance	-0.038 (0.699)	-0.921^{**} (<0.001)
English as Second Language	$+0.972^{**}$ (<0.001)	$+0.624^{**}$ (0.014)	Non-parental care	$+0.005$ (0.150)	$+0.007$ (0.338)
Language/Religion Class Attendance	$+0.002^{**}$ (0.007)	+0.001 (0.749)			
Neighborhood Characteristics					
Separate School	+0.127 (0.144)	$+0.287^{*}$ (0.069)	Francophone School	-1.981^{**} (0.002)	-2.961^{*} (0.054)
Medium Gini	-0.230 (0.755)	-0.593 (0.685)	High Gini	-0.752 (0.289)	-0.874 (0.515)
Median Income in $10,000	$+0.150^{*}$ (0.092)	$+ 0.366^{**}$ (0.044)	Unemployment % of People15+ Years of age	+0.004 (0.718)	+0.001 (0.982)
%of People with at Least High School Degree	-0.003 (0.594)	-0.015 (0.174)	Average Value of Dwelling in Real $ 10,000	-0.003 (0.859)	+0.012 (0.697)
Geographical Area Characteristics					
Prince Albert	+0.127 (0.946)	+1.152 (0.773)	Regina	+1.269 (0.191)	+2.658 (0.120)
Saskatoon	$+1.256^{**}$ (0.043)	$+2.458^{**}$ (0.033)			
Within Level Interactions					
Aboriginal*Number of Skills	-0.095^{*} (0.092)	-0.177^{**} (0.022)	Gender*Non-parental Care	$+0.001$ (0.555)	-0.000 (0.932)
Gender*Age	+0.088 (0.634)	+0.108 (0.736)	Medium Gini*Median Income	+0.061 (0.818)	+0.089 (0.863)

Table 4.10 (Continued)

High Gini*Median Income	+0.319 (0.254)	+0.247 (0.642)			
Cross Level Interactions					
Medium Gini*Days Absent	+0.023** (0.046)	−0.014 (0.388)	High Gini*Days Absent	+0.017 (0.125)	+0.008 (0.573)
Medium Gini*Aboriginal	+0.270 (0.459)	−0.041 (0.858)	High Gini*Aboriginal	+0.285 (0.446)	+0.020 (0.926)
Separate School*Days Absent	−0.023** (0.010)	−0.031** (0.010)	Francophone School*Days Absent	+0.028 (0.324)	+0.019 (0.808)
Medium Gini*Non-parental care	−0.006** (0.022)	−0.005 (0.229)	High Gini*Non-parental care	−0.005** (0.045)	−0.007 (0.168)
Prince Albert*Aboriginal	+0.260 (0.575)	+0.126 (0.903)	Regina*Aboriginal	−0.564** (0.014)	−0.383 (0.265)
Saskatoon*Aboriginal	−0.314 (0.118)	−0.936** (0.009)	Prince Albert*Non-parental care	−0.002 (0.653)	+0.002 (0.812)
Regina*Non-parental care	+0.002 (0.513)	−0.002 (0.700)	Saskatoon*Non parental-car	−0.000 (0.877)	−0.004 (0.337)
Prince Albert*Medium Gini	+1.992** (0.016)	+1.252 (0.423)	Regina*Medium Gini	+0.681 (0.241)	-0.345 (0.713)
Saskatoon*Medium Gini	+0.210 (0.472)	+0.459 (0.425)	Prince Albert*High Gini	+1.607** (0.039)	+0.336 (0.836)
Regina*High Gini	−0.039 (0.948)	−0.835 (0.403)	Saskatoon*High Gini	−0.123 (0.667)	−0.026 (0.962)
Prince Albert*Median Income	−0.804 (0.225)	−1.007 (0.493)	Regina*Median Income	−0.495 (0.139)	−0.555 (0.375)
Saskatoon*Median Income	−0.497** (0.022)	−1.085** (0.008)	Constant	+0.916** (<0.001)	+0.855** (<0.001)
p-value for LR Test	<0.001	<0.001			
Number of Observations	8028	8037			
Hosmer & Lemeshow Goodness of Fit Test(p-value)	0.2971	0.1104			

Note: † indicates the p-values. ** indicates the coefficient estimates significant at 5% level or lower, while * indicates level of significance at 10% or lower.

Fifthly, attending a language or religion class had a very weak impact (max odds 1.002) on child's vulnerability status and multiple vulnerability. Its effect is practically negligible.

Sixthly, a higher neighborhood unemployment rate had a weak impact on both child's vulnerability and multiple vulnerability and for each 5% increase in unemployment, odds of vulnerability and multiple vulnerability were multiplied by 1.020 and 1.005, respectively. Seventhly, a higher neighborhood percentage of high school educated people had a weak buffering impact on both child's vulnerability and multiple vulnerability so that each 5% increase of high school degree holders decreased the odds of vulnerability status and multiple vulnerability by 2.5% and 7.3%, respectively.

Finally, a higher neighborhood average value of dwelling had a weak buffering impact on child's vulnerability, so that each increase of $10,000 to the value of dwelling decreased odds of vulnerability by only 1%.

4.5.2. Within-Level Interactions

Similar to the multilevel linear model, Table 4.10 depicts within-level interactions $\widehat{\beta}\ (p - value)$ based on multilevel logistic regression model for two binary outcomes. The child level variables of 'Aboriginal status', 'gender', and 'gender' modify the effects of other child level variables 'number of special skills', 'non parental care', and 'age', respectively. In addition, the neighborhood level variable of 'Gini income inequality' modifies the effect of the other neighborhood level variable 'Median Income'.

First of all, increased number of skills had a significantly buffering effect on both child's vulnerability and multiple vulnerability and, in addition, Aboriginal status exacerbated such buffering effect, so that each additional skill significantly decreased odds ratio of vulnerability and multiple vulnerability in Aboriginal children versus non-Aboriginal children by 9.1% and 16.2%, respectively. Figure 4.10 shows the predicted probability (95% CI) of vulnerability and multiple vulnerability by number of special skills and Aboriginal status. As observed, in both outcomes, the predicted probability of Aboriginal children with the lowest number of skills is significantly higher than non-Aboriginal children with the lowest number of skills; however, by increasing the number of skills such differences become non-significant, and, then at the highest number of skills their magnitude diminish and for the multiple vulnerability it vanishes.

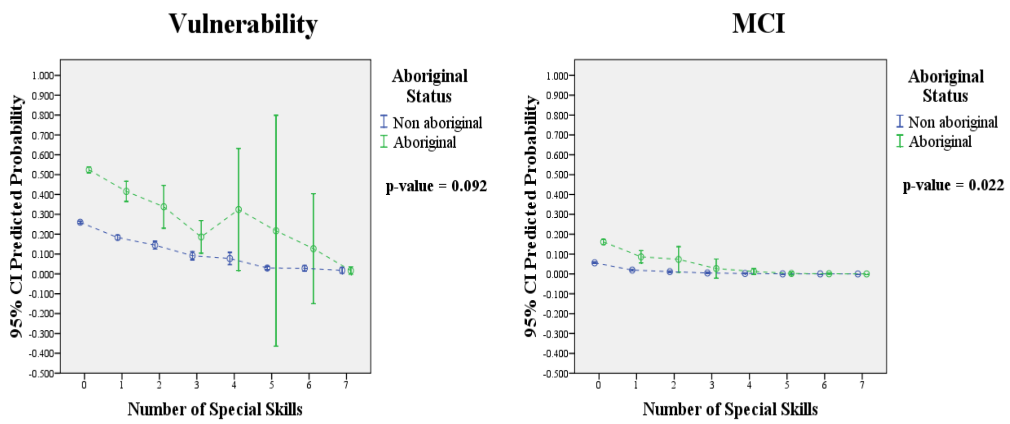

Figure 4.10. Predicted Probabilities (95% CI) of Vulnerability and MCI Outcomes by Number of Special Skills and Aboriginal Status

Secondly, non-parental care status had an exacerbating weak effect on both child’s vulnerability and multiple vulnerability and in addition, male status had no interaction with it.

Thirdly, male status had an exacerbating effect on both child’s vulnerability status and multiple vulnerability and each additional year of age increased odds ratio of vulnerability and multiple vulnerability in males versus females by 9.2% and 11.4%, respectively.

Finally, higher neighborhood median income and higher neighborhood income inequality had a compounding exacerbating impact on both child’s vulnerability and multiple vulnerability. In fact, for each additional $10,000 increase in neighborhood median income, odds ratio of vulnerability and odds ratio of multiple vulnerability in children living in neighborhoods with medium(high) income inequality versus those children living in neighborhoods with low income inequality increased by 6.3% (37.6%) and 9.3% (28.0%), respectively.

4.5.3. Cross-Level Interactions

Similar to the multilevel linear model, Table 4.10 depicts cross-level interactions $\widehat{\beta}\ (p-value)$ based on multilevel logistic regression model for two binary outcomes. The neighbor hood level variables of ‘Gini Income Inequality’, ‘Gini Income Inequality’ , and ‘School Type’ modify the effect of child level variables of ‘days absent from school’, ‘Aboriginal Status’, and ‘days absent from school’, respectively. In addition, the geography variables of ‘major city’, and ‘major city’ modify the effect of child level variables of 'Aboriginal Status’ and ‘Non parental care’, respectively. Furthermore, neighborhood level variables “Gini Income Inequality’ and ‘Median Income’ modify the effect of geography level variables ‘major city’ and ‘major city’, respectively.

First of all, days absent from school had a significantly exacerbating effect on both child’s vulnerability and multiple vulnerability. In addition, for each additional week absent from school, odds ratio of vulnerability for children living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality increased by 17.5% (12.6%). However, the results for the multiple vulnerability outcome were inconclusive. Also, for such additional weeks absent from school, odds ratio of vulnerability and multiple vulnerability for children attending separate (Francophone) schools versus those attending public schools significantly decreased (increased) by 14.9% (21.7%) and 19.5% (14.2%), respectively.

Secondly, aboriginal status had a significantly exacerbating effect on both child’s vulnerability and multiple vulnerability. Furthermore, for the first outcome, its effect was exacerbated as child’s neighborhood income inequality increased. Specifically, there were 31.1% (33.0%) increase in OR values of vulnerability for Aboriginal children versus non-Aboriginal children of moving from neighborhoods with low to medium (high) income inequality. The results for the multiple vulnerability outcomes were inconclusive. Figure 4.11 shows Predicted probabilities (95% CI) of vulnerability by neighborhood income inequality and Aboriginal status. As observed, the increase in vulnerability probabilities in aboriginal children is slightly sharper than non-aboriginal children.

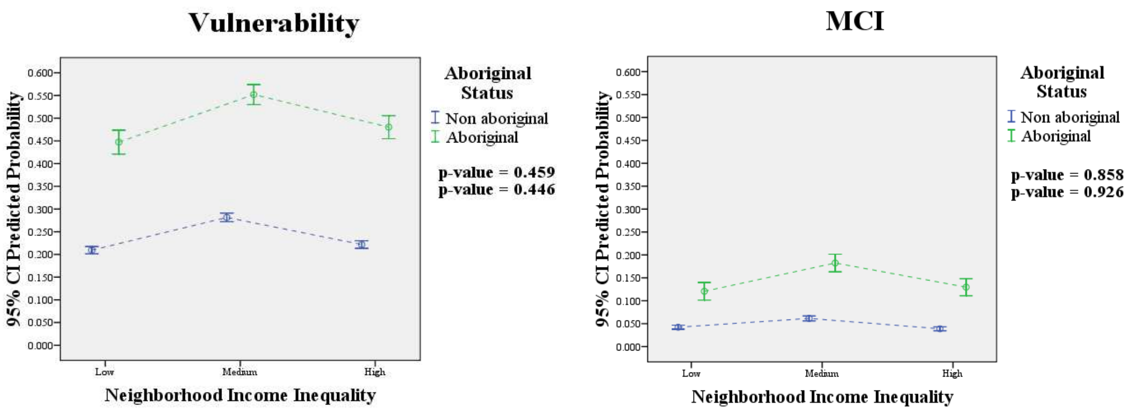

Figure 4.11. Predicted Probabilities (95% CI) of Vulnerability and MCI by Neighborhood Income Inequality and Aboriginal Status

Thirdly, non-parental care had an exacerbating weak effect on both child’s vulnerability and multiple vulnerability, and in addition, its exacerbating effect was mitigated and reversed very weakly according to child’s neighborhood income inequality (max 1% decrease in odds ratio of vulnerability and multiple vulnerability for children without parental care versus others among children living in neighborhoods with medium or high income inequality). Furthermore, child’s major city had very weak modifying effect on non-parental status (max 1% change in odds ratio of vulnerability and multiple vulnerability for children without parental care versus others among children living in Prince Albert, Regina or Saskatoon).

Fourthly, as observed above, Aboriginal status had a significantly exacerbating effect on both child's vulnerability and multiple vulnerability, and, in addition, child's major city can exacerbate or buffer such effect. Specifically, odds ratios of vulnerability and multiple vulnerability for Aboriginal children versus non-aboriginal children among those children living in Prince Albert were 29.7% and 13.4% higher than those children living in non-urban areas. However, odds ratios of vulnerability and multiple vulnerability for Aboriginal children versus non-Aboriginal children among those children living in Regina were 43.1% and 31.8% lower than those children living in non-urban areas. Similar to Regina's case, odds ratios of vulnerability and multiple vulnerability for aboriginal children versus non-Aboriginal children among those children living in Saskatoon were 26.9% and 60.8% lower than those children living in non-urban areas. Figure 4.12 presents predicted probabilities (95% CI) of vulnerability and multiple vulnerability outcomes by geography and Aboriginal status. As observed, Prince Albert children had the lowest predicted probabilities of vulnerability and multiple vulnerability in both Aboriginal and non-Aboriginal groups while Regina children had the highest predicted probabilities of vulnerability and multiple vulnerability for such groups.

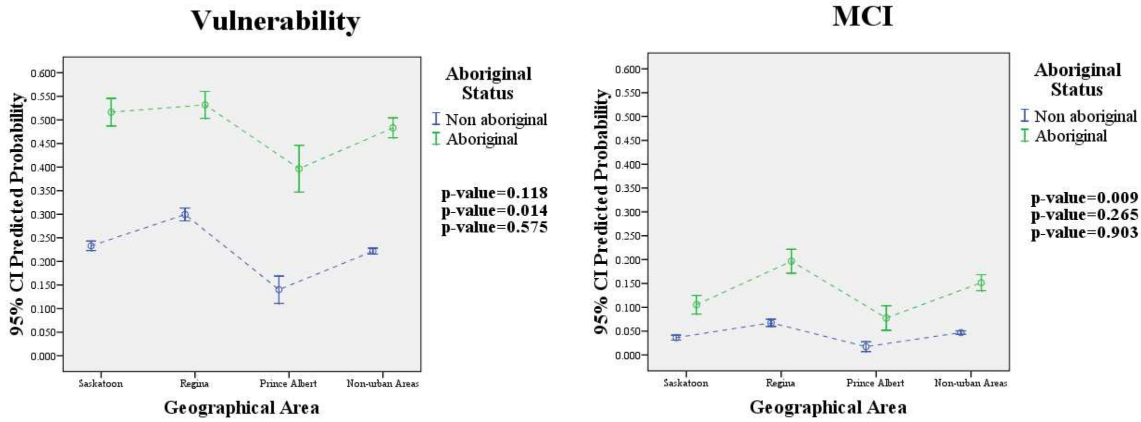

Figure 4.12. Predicted Probabilities (95% CI) of Vulnerability and MCI Outcomes by Geographical Area and Aboriginal Status

Fifthly, children living in Prince Albert had higher odds of vulnerability and multiple vulnerability, respectively, compared with non-urban areas children. Additionally, living in neighborhoods with medium or high income inequality exacerbated such effect. Firstly, in Prince Albert, odds ratios of vulnerability and multiple vulnerability for children living in neighborhoods with medium (high) income inequality versus those living in neighborhoods with low income inequality were 6.330 , 2.497 times (3.988, 0.399 times) higher than those of non-urban areas children, respectively. Next, children living in Regina had higher odds of vulnerability and multiple vulnerability compared with non-urban areas children, and, furthermore, among the Regina children, the odds ratio of vulnerability for children living in neighborhoods with medium income inequality versus those living in neighborhoods with low income inequality was 0.976 times higher than non-urban areas children. Similar to Regina's case, children living in Saskatoon had higher odds of vulnerability and multiple vulnerability and, in addition, among Saskatoon children odds ratios of vulnerability and multiple vulnerability for children living in neighborhoods with medium income inequality versus those living in neighborhoods with low income inequality were 0.234 times and 0.582 times higher than those of non-urban areas children, respectively. Figure 4.13 presents predicted probabilities (95% CI) of vulnerability and multiple vulnerability outcomes by geography and neighborhood income inequality. As observed, compared with non-urban areas children, those living in Prince Albert had the highest vulnerability outcome fluctuation as the neighborhood income inequality changed, while Saskatoon children had the lowest. Also, compared to non-urban areas children, Saskatoon children had the lowest multiple vulnerability outcome fluctuation as the neighborhood income inequality changed.

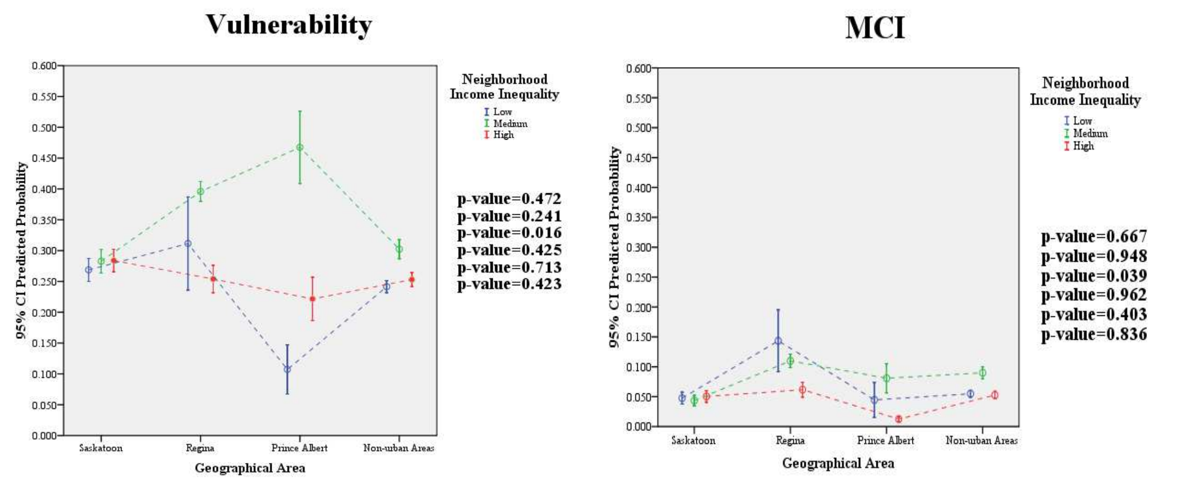

Figure 4.13. Predicted Probabilities (95% CI) of Vulnerability and MCI Outcomes by Geographical Area and Neighborhood Income Inequality

Finally, although children living in Prince Albert had higher odds of vulnerability and multiple vulnerability compared with non-urban areas children, respectively, odds ratios of vulnerability and multiple vulnerability for Prince Albert children versus non-urban areas children decreased by 55.2% and 63.5%, respectively for each $10,000 increase in child's residential neighborhood median income. Despite of the fact that children living in Regina had strongly higher odds of vulnerability and multiple vulnerability in comparison to non-urban areas children, for each $10,000 increase in child's neighborhood median income odds ratios of vulnerability and multiple vulnerability for Regina children versus non-urban areas children decreased by 39.0% and 42.6%, respectively. Similarly, for the city of Saskatoon, despite the fact that children living in this city had higher odds of vulnerability and multiple vulnerability compared with non-urban areas children, odds ratios of vulnerability and multiple vulnerability for Saskatoon children versus non-urban areas children decreased by 39.2% and 66.2%, respectively for each $10,000 increase in child's residential neighborhood median income.

4.5.4. Child-level, Neighborhood-level and Geographical Area-level Variances

Similar to the multilevel linear model, Table 4.11 presents variance, variance partition coefficient and intra-class coefficients at geography, neighborhood, and child levels for vulnerability and multiple vulnerability outcomes.

Table 4.11: Geographical Area Level, Neighborhood Level, Child Level Variance, VPC, and ICC Values of Multilevel Logistic Model

	Vulnerability Status	MCI		Vulnerability Status	MCI
$\sigma^2_{e(3)}$	0.004	3×10^{-7}	VPC (child)	0.881	0.806
$\sigma^2_{e(2)}$	0.439	0.794	ICC(geographical area)	0.001	7×10^{-8}
$\sigma^2_{e(1)}$	3.290	3.290	ICC(neighborhood)	0.119	0.194
VPC(geographical area)	0.001	7×10^{-8}			
VPC (neighborhood)	0.118	0.194			

First of all, similar to the multilevel linear case, as the hierarchy level increased, the level associated variance given by all variables in that level decreased for both vulnerability and multiple vulnerability outcomes. Partition of total variance by child, neighborhood and geography level is shown in Figure 4.14.

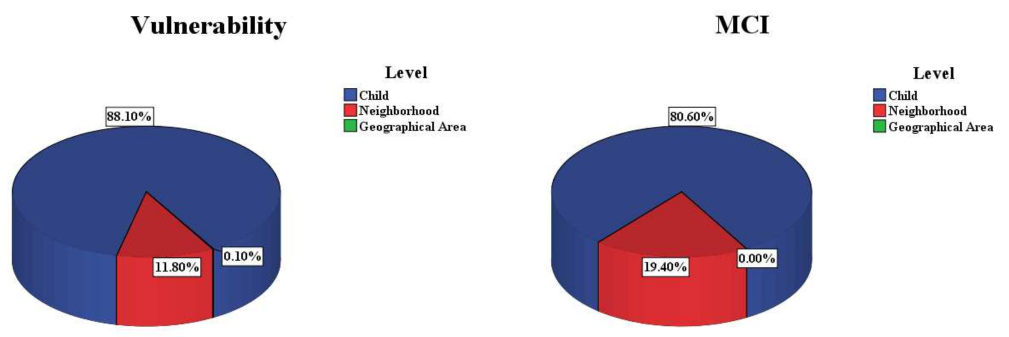

Figure 4.14: Variance Partition by Child, Neighborhood and Geographical Area Levels of Two Binary Outcomes

Secondly, geographical areas at the 97.5 th percentile of the geography distribution were estimated to have 24.8% higher log odds of vulnerability status than geographical areas at the 2.5 th percentile. For multiple vulnerability outcomes, such differences were negligible. Next, within geographical areas, neighborhoods at the 97.5 th percentile of neighborhood distribution were estimated to have 2.597 units higher log odds of vulnerability status than neighborhoods at 2.5 th percentile. For the multiple vulnerability outcome, the corresponding value is 3.493 units. Finally, within neighborhoods, children at the 97.5 th percentile of distribution were estimated to have 7.110 units higher log odds of vulnerability and multiple vulnerability than neighborhoods at the 2.5 th percentile.

Thirdly, looking at the VPC statistics, less than 1% of the variations in vulnerability status and MCI outcomes lied between geographical areas; 11.8% and 19.4% lied within geographical areas between different neighborhoods, while the remaining 88.1% and 80.6% lied within neighborhoods between children. More than three-quarters of vulnerability and multiple vulnerability outcomes variations were attributable to the children themselves.

Finally, observing ICC statistics, the correlations between two children living in the same geographical area but different neighborhoods were negligible for both vulnerability and multiple vulnerability outcomes, and the correlation between two children living in the same geographical area and same neighborhood were 0.119 and 0.194 for vulnerability and multiple vulnerability outcomes, respectively. Consequently, living in the same neighborhoods rather than adjacent neighborhoods played a key role in the similarity of children's vulnerability and multiple vulnerability outcomes.

4.6. Framework of significant variables

A framework for the groups of variables significantly associated with majority of 7 outcomes (5 EDI domains, vulnerability and multiple vulnerability), including main effects, within-level interactions and cross-level interactions is presented in Figure 4.15. For main predictors, 'Age', 'Days Absent from School', 'Number of Special Skills', 'Number of Special Problems', 'Aboriginal Status', 'ESL status', 'Francophone School', and 'Median Income' were significantly associated with most of the above 7 outcomes. In addition, among within-level interactions, only 'Aboriginal*Number of Skills' was significant in 6 of 7 outcomes. Finally, among cross-level interactions, 'Separate School*Days Absent', 'High Gini*Non-parental care', "Regina*Aboriginal', and 'Saskatoon*Median Income' were significant in at least 4 of 7 outcomes.

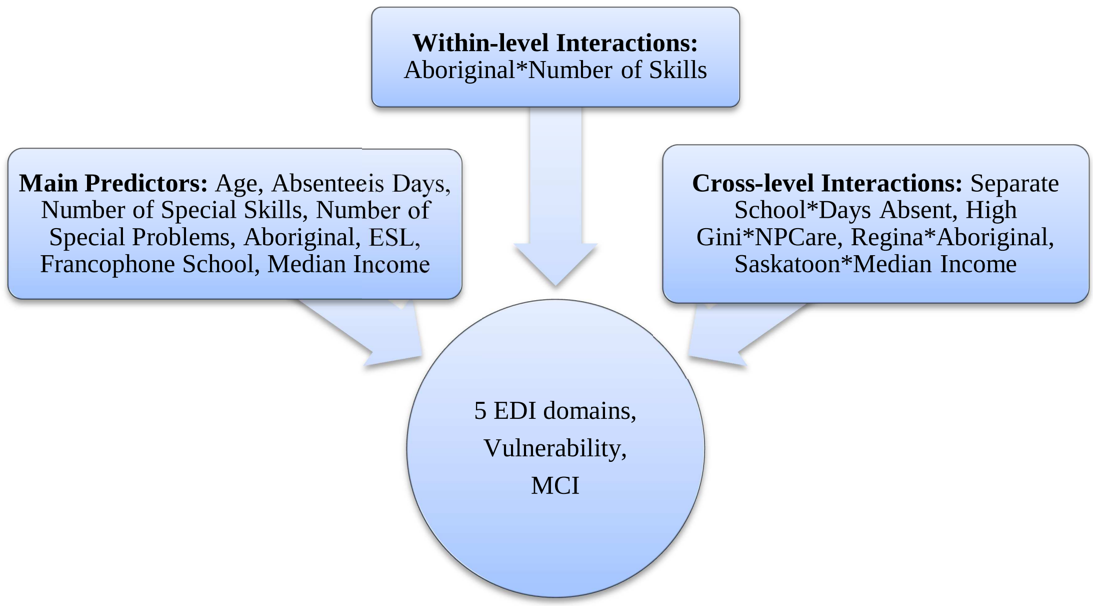

Figure 4.15. Significantly Associated Main Factors, Within-level Interactions, and Cross-level Interactions for 5 EDI Domains, Vulnerability and MCI Outcomes

4.7. Some Remarks on Model Fit and Hierarchy

This last section of the Results chapter deals with the multilevel model's goodness of fit and issues related to its hierarchy and number of levels.

First of all, Table 4.8c presents Pearson χ^2/df values for the 5 EDI domains. As observed, for the two domains of Physical well-being and Emotional maturity the model had good fit (1.0 < Pearson χ^2/df < 2.0), while for the two domains of Social competence and Language & cognitive development it had moderate fit (2.0 < Pearson χ^2/df < 4.0). For the domain of Communication & general knowledge, the model had weak fit (4.0 < Pearson χ^2/df).

Secondly, Table 4.10 presents Hosmer & Lemeshow Goodness of Fit test results for the two binary outcomes of vulnerability and MCI. As observed, for both outcomes the test p-values were greater than 0.05 implying that both models fit well.

Finally, Table 4.12 presents Likelihood Ratio (LR) test results comparing the three level model (geographical area-neighborhood-child) with the two level model (geographical area-child) and the other two level model (neighborhood-child). These tests are known as test of cluster effects and a test of super-cluster effects, respectively. In all 7 domains cluster effects were significant; however, only in one domain, Emotional maturity, super-cluster effects was significant. Consequently, only the 3-level linear model with Emotional maturity as its outcome variable was statistically significant and justifiable while the other six 3-level models were not statistically significant.

Table 4.12. LR Test Statistics (p-values) for Testing Cluster Effects and Super-cluster Effects

Model	Linear					Logistic	
	Physical Well-being	Social Competence	Emotional Maturity	Language& Cognitive Development	Communication & General Knowledge	Vulnerability Status	MCI
Cluster Effects	212.216 (<0.001)	231.588 (<0.001)	287.113 (<0.001)	129.986 (<0.001)	230.011 (<0.001)	44.063 (<0.001)	59.385 (<0.001)
Super -cluster Effects	1.808 (0.404)	0.785 (0.675)	6.293 (0.043)	1.676 (0.432)	0.847 (0.654)	3.141 (0.207)	0.695 (0.706)

CHAPTER 5: DISCUSSION

Old proverbs are the children of truth.
-Welsh Proverb

This chapter deals with a brief discussion of the study results. Firstly, a summary of descriptive and inferential statistics is presented, and then the three main epidemiologic questions are answered. Secondly, the study results in relationship to past publications are explored. It is shown that the results of child level variable of Aboriginal status, neighborhood level variable of income inequality and geographical area level variable of major city give more evidence of the past literature's results including income inequality hypothesis and urbanization impact on citizens. Finally, the chapter concludes by referring to the study strengths and limitations.

5.1. Summary of Major Findings

This study considered 9045 children age 4.5-8.0 years in 185 neighborhoods nested in 3 major cities (Saskatoon, Regina, and Prince Albert) and non-urban areas, recruited in the 2008-2009 school year.

The study began with the overall results in the EDI for Saskatchewan children. Firstly, Saskatchewan children's average scores in all 5 EDI domains were 1.2%-5.2% lower than those of national normative sample results. Moreover, 30% (6.4%) of children were rated as vulnerable in at least one (three) EDI domain(s). Secondly, on average a child was absent from his or her school for 3.8 days. Almost 50% of children were girls, almost 17% of children were Aboriginal status, almost 96% of children were native English speakers and almost 50% of them were beneficiary of parental care. Thirdly, on average each neighborhood had medium income inequality (Gini=0.127), a median income of $25,020, almost 57% of their residents holding high school diplomas, and almost 70% of children attending public school system. Finally, as the province's largest cities, Saskatoon, Regina and Prince Albert constituted almost 50% of the children in the study.

The study's primary inferential results described the differences in the 5 EDI and proportions of vulnerability and multiple vulnerability outcomes based on characteristics of child, neighborhood and geography. Firstly, for all 5 EDI domains, girls had significantly higher averages than boys (difference percentage 4.9% −12.4%) and their proportion of vulnerability and multiple vulnerability were significantly lower (difference percentage 41.2%−55.2%). Secondly, for all 5 EDI domains Aboriginal children had significantly lower averages than non-aboriginal children (difference percentage 8.8%−19.4%) and significantly higher proportion of vulnerability and multiple vulnerability (difference percentage 106.4%−211.0%). Thirdly, for 4 EDI domains native English speaking children had higher averages than children with ELS status (difference percentage 1.7%−35.9%) and significantly lower proportion of vulnerability (difference percentage 37.5%). Fourthly, for all 5 EDI domains, children who had parental care had higher averages than others (difference percentage 0.1%−6.9%) and significantly lower proportions of vulnerability and multiple vulnerability (difference percentage 19.0%−24.0%). Fifthly, as the neighborhood income inequality increased, children's EDI averages followed a quadratic (decreasing-increasing) trend in all 5 EDI domains, and in complementary results their proportions of vulnerability and multiple vulnerability followed a quadratic (increasing-decreasing) trend. Sixthly, for all 5 EDI domains children in neighborhoods with Francophone schools had higher averages than those in neighborhoods with public schools (difference percentage 0.3%−11.2%) and lower proportions of vulnerability and multiple vulnerability (difference percentage 40.7%−42.7%) as well. Finally, Saskatoon children had the highest averages in the province for 2 EDI domains, and Prince Albert children had the highest averages in the province in the other 3 EDI domains. In all 5 EDI domains Regina children had the lowest averages in the province. Moreover, Regina children had the highest proportions of vulnerability and multiple vulnerability and Prince Albert children had the lowest proportions of vulnerability and multiple vulnerability.

The study's main results dealt with the three main questions proposed in the objectives section 3.1. In the following, the answer for each question is presented.

Answering question 1, detailed in section 3.1, the significant determinants of developmental health at child, neighborhood and geography level were: (i) males had lower scores than females for all 5 EDI domains and for 2 EDI domains of Physical well-being and Language & cognitive development such differences were significant, in addition, the vulnerability and multiple vulnerability odds ratios for males versus females were in the range of 1.276-1.279 units; (ii) Aboriginal children had significantly lower scores than non-Aboriginal children for each of the 5 EDI domains. The vulnerability and multiple vulnerability odds ratios of Aboriginal children versus non-aboriginal children were significant and in range of $2.188 - 2.863$ units; (iii) attending French or English immersion school had a very weak effect on all 5 EDI domains, but vulnerability and multiple vulnerability odds ratios of non-attending (French/English immersion school) children versus attending children were within the range $1.039 - 2.512$ units; (iv) native English speaking children had higher average scores than ESL children for 4 EDI domains. Furthermore, odds ratios of vulnerability and multiple vulnerability in ESL children versus native English speaking children were significant and in the range of $1.886 - 2.643$ units; (v) not unexpectedly, child's age was significantly associated with all 5 EDI domains; for each additional year in age, odds ratio of vulnerability and multiple vulnerability decreased by

53.50% and 62.20%, respectively ; (vi) days absent from school had a significantly negative effect on all 5 EDI domains, and for each additional week absent from school, odds ratios of vulnerability and multiple vulnerability were significant and in the range of 1.040-1.047 units; (vii) higher neighborhood median income had negative effect on all 5 EDI domains and, in addition, it had significantly exacerbating weak impact on child's vulnerability and multiple vulnerability; (viii) compared to neighborhoods with public schools, neighborhoods with separate (Francophone) schools had significantly negative (positive) effect on 3 (4) EDI domains, and, in addition, they had the odds ratios of vulnerability and multiple vulnerability (vulnerability and multiple vulnerability) in the range of 1.135−1.332 units (the range of 0.052−0.138 units); and (ix) compared to non-urban areas, living in the major cities of the province, Saskatoon, Regina, and Prince Albert had negative effect on all 5 EDI domains and the associated vulnerability and multiple vulnerability odds ratios were in the range of 1.135-3.557 units and range of 3.165-14.267 units, respectively.

Answering question 2, detailed in section 3.1, effects that are modified by other determinants at the same hierarchical level (within-level effect modification), and effects that are modified by other determinants at a different hierarchical level (cross-level effect modification) were discussed in terms of significancy and strength as follows: (i) increased numbers of special skills significantly mitigated and even reversed the negative impact of Aboriginal status on all 5 EDI domains, and also on vulnerability and multiple -vulnerability. For each additional skill, EDI outcome differences between Aboriginal children versus non-Aboriginal children decreased in range of $0.046 - 0.223$ units and the related odds ratios of vulnerability and multiple vulnerability decreased by 9.1% and 16.2%, respectively; (ii) higher neighborhood income inequality and neighborhood median income had compounding negative effect on 4 EDI domains, and on vulnerability and multiple vulnerability. For each additional $10,000 increase in neighborhood median income, 4 EDI outcome differences between children living in neighborhoods with medium (high) income inequality and others increased in range of 0.155-0.309 units (range of 0.012-0.460 units). Additionally, related odds ratios of vulnerability and multiple vulnerability increase by at least 6.3% (9.3%); (iii) higher neighborhood income inequality exacerbated negative impact of days absent from school in 3 EDI domains and on vulnerability. For each additional week absent from school, 3 EDI outcome differences between children living in neighborhoods with medium (high) income inequality and others increased in a range of 0.091-0.182 units (0.084-0.119 units). The related vulnerability odds ratio increased by at least 12.6%; (iv) higher neighborhood income inequality exacerbated the negative effect of Aboriginal status on 4 EDI domains and on vulnerability. In fact, the 4 EDI outcome differences between Aboriginal children versus non-Aboriginal children for those living in neighborhoods with medium (high) income inequality increased in range of 0.008-0.293 units (0.050-0.137 units) and their associated vulnerability odds ratios increased by at least 31.1%; (v) child's major city can exacerbate (in Prince Albert's case) or mitigate (in Regina and Saskatoon's cases) the negative effect of Aboriginal status for all 5 EDI domains, on vulnerability and multiple vulnerability. For Prince Albert, 4 EDI outcome differences between Aboriginal children versus non-Aboriginal children increased in a range of 0.008−0.311 units, and their associated vulnerability and multiple vulnerability odds ratios increased by at least 13.4%. In contrast, for Regina and Saskatoon, 4 EDI outcome differences between Aboriginal children and non-Aboriginal children decreased in range of 0.276−0.476 and 0.167−0.480 units, respectively. Their associated vulnerability and multiple vulnerability odds ratios decreased by at least 31.8% and 26.9%, respectively; (vi) child's residential neighborhood higher income inequality can exacerbate the negative effect of living in the major city on all 5 EDI domains, on vulnerability and multiple vulnerabilities. For Prince Albert, Regina, and Saskatoon, respectively, 5 EDI, 3

EDI, and 2 EDI outcome differences between children living in neighborhoods with medium (high) income inequality and others increased in a range of 0.496−1.194 units (0.132−0.863 units), in a range of 0.569−0.745 units (0.138−0.348 units), and in range of 0.059−0.381 units (0.085 −0.160 units), respectively. In addition, the associated odds ratio of vulnerability increased at least 3.988, 0.976 and 0.234, for Prince Albert, Regina, and Saskatoon, respectively; and (vii) higher neighborhood median income can mitigate the negative effect of living in a major city on at least 4 EDI domains, on vulnerability and multiple vulnerability. For each additional $10,000 increase in neighborhood median income, 4 EDI, 5 EDI, and 5 EDI outcome differences of children living in Prince Albert, Regina and Saskatoon, respectively versus non-urban areas children decreased in a range of 0.324−0.855 units, 0.156−0.471 units and 0.191−0.678 units, respectively. In addition, for these three major cities, the associated odds ratio of vulnerability (and odds ratio of multiple vulnerability) decreased by at least 55.2%, 39.0% and 39.2% (63.5%, 42.6% and 66.2%), respectively.

Answering question 3, detailed in section 3.1, the relative contributions of main determinants at each level to the variance of the 5 EDI domains, probability of vulnerability and probability of multiple vulnerability were discussed as follows: (i) as one moves from child-level to neighborhood level and then to geography level, the variance of EDI outcomes explained decreased for all 5 EDI domains, and decreased also for vulnerability and multiple vulnerability; (ii) for fixed geographical area characteristics, the variance between neighborhoods constituted 20.6%−25.0% of the variation in the 5 EDI outcomes, while it constituted 11.8%−19.4% of the variation in probability of vulnerability and multiple vulnerability; (iii) for fixed geographical area and neighborhood characteristics, the variation between children contributed 70.8%−75.0% of the variation in the 5 EDI domains whilst it contributed 80.6%−88.1% of the variation in probability of vulnerability and probability of multiple vulnerability. Consequently, individual child characteristics had the most contribution to the variation of child's developmental health outcomes with at least a 70% contribution to the variance of each of the 7 continuous and binary outcomes; and (iv) children living in the same neighborhoods rather than adjacent neighborhoods had much more similarity in 5 EDI domains, vulnerability and multiple vulnerability as they had correlation in the range of 0.250−0.292 in all 5 EDI domains and 0.119−0.194 for vulnerability and multiple vulnerability outcomes.

5.2. Interpretation and Relation of Findings to Similar Studies

The current study explored the individual, neighborhood contextual, and geographical factors associated with children's school readiness and vulnerability.

The study's main individual factors associated with children's school readiness and vulnerability in terms of its significance and strength included gender, Aboriginal status, ESL status, number of special skills and days absent from school. Firstly, males had lower scores than females in all 5 EDI domains and had higher odds of vulnerability and multiple vulnerability, consistent with past research.[7,48,49] Secondly, compared with non-Aboriginal children, Aboriginal children had significantly lower scores in all 5 EDI domains and higher odds of vulnerability, concordant with past publications.[28,50] Also, increased number of skills significantly mitigated and reversed the adverse effect of Aboriginal status in all 5 EDI domains and lowered odds of vulnerability in large provincial scale consistent with previous smaller city scale result.[51] Thus, by advocating new policies and programs that introduce and work on increasing skills in Aboriginal children one may increase their functioning and improve their developmental health at preschool and elementary school years and reduce their associated school readiness gap with non-Aboriginal children. Thirdly, compared to ESL children, native English speaking children had higher scores on 4 EDI domains and lower odds of vulnerability and multiple vulnerability, consistent with past publications.[31,50,52] Finally, days absent from school had adverse effect on all 5 EDI outcomes, and exacerbating odds of vulnerability and multiple vulnerability confirming previous study's results,[51] on a much larger scale. This finding gives significant evidence that kindergarten and elementary school educational programs provide critical elements for children's developmental health.

The study's main neighborhood contextual factors associated with children's school readiness and vulnerability in terms of significance and strength included neighborhood school type, income inequality and median income. Firstly, among children without days absent from school, those studying at separate schools had lower scores than those studying at public school in all 5 EDI domains and had higher odds of vulnerability and multiple vulnerability; however, children studying at public schools had lower scores than those studying at Francophone school for 4 EDI domains and, in addition, they had significantly higher odds of vulnerability and multiple vulnerability. Thus, in terms of child developmental health, Francophone schools had better impact than public schools and the later had better effect than separate schools. The later conclusion challenges the common belief regarding the superiority of separate schools to public schools in terms of children's education and performance and is consistent with previous research in children's mathematics performance.[53] However, this conclusion is in disagreement with other previous studies regarding children's reading performance.[54] Secondly, higher neighborhood income inequality and days absent from school had compound negative effects on 3 EDI domains and invulnerability. Also, higher neighborhood income inequality exacerbated the adverse impact of Aboriginal status in 4 EDI domains and on invulnerability status. Furthermore, higher neighborhood income inequality and neighborhood median income had compounding negative effect in 4 EDI domains, exacerbating odds of vulnerability and multiple vulnerability. The above results give further evidence of the income inequality hypothesis,[55,56] and, furthermore, in companion with minimum neighborhood contribution of 11.8% to variance of 5 EDI outcomes, vulnerability and multiple vulnerability, they show that neighborhood contextual characteristics affect children's developmental health in terms of their significance and strength consistent with past publications.[7,8,30,31,57]

The study's main geographical factors associated with children's school readiness and vulnerability included living in the major provincial geographical areas of Prince Albert, Regina, Saskatoon, and non-urban areas. In addition, the modifying effects of Aboriginal status, neighborhood income inequality and neighborhood median income on children's geographical area were explored. Children living in the above major cities had lower scores than non-urban areas' children in at least 4 EDI domains and also had higher odds of vulnerability and multiple vulnerability. Firstly, compared to non-urban areas' Aboriginal children, Prince Albert Aboriginal children had even lower scores in 4 EDI domains, and higher odds of vulnerability and multiple vulnerability; however, Regina and Saskatoon Aboriginal children had higher scores in 4 EDI domains and had lower odds of vulnerability and multiple vulnerability. Consequently, in terms of child public health policy, Prince Albert should promote more special care and support programs for its Aboriginal children to narrow their developmental health gap with Aboriginal children living in non-urban areas. Secondly, compared to non-urban areas, living in major cities with higher neighborhood income inequality lowers children's scores in at least 2 EDI domains whereas living in these major cities with higher neighborhood median income increases children's scores in all 5 EDI domains. These results give further evidence of the income inequality hypothesis,[55] in particular the developmental health benefits accruing to children in urban setting where neighborhood income levels are generally higher. These results are consistent with those reported in previous studies as well.[58,59]

5.3. Strengths and Limitations

The current study added knowledge to the research literature on Saskatchewan children's school readiness and vulnerability and had the following strengths. Firstly, the current study was among a few child developmental health studies in Canada that was conducted on a provincial scale taking into account data available at multiple levels of social hierarchy. Previous studies that had utilized province-wide data had been conducted taking into account either two levels of hierarchy or on only one level of hierarchy. Secondly, the current study considered both linear multilevel model and logistic multilevel model in order to explore child developmental health by domains and also vulnerability. This approach present comprehensive approach including both continuous 5 EDI domains and binary outcomes of vulnerability and multiple vulnerability. This approach is in contrast to most former Canadian studies that had been conducted via either a multilevel linear model or multilevel logistic model. Thirdly, as the study sample covered almost all 4.5-8 years old children attending schools around the province (a population sample) it is unlikely that this study would have suffered from potential selection biases. Fourthly, the EDI and Census data have been shown to have high reliability and acceptable validity, and, therefore subject to little information bias. Finally, the study consisted of large sample size and included of a variety of predictors and their interactions yielding a relatively broad spectrum of information regarding individual, neighborhood and geographical areas impacts on children's developmental health and vulnerability.

The current study, however, did have some limitations. Firstly, there was a two year time difference between the collection of EDI data in 2008 and the census data in 2006 and this time difference could have affected the findings. Secondly, fitting the same set of predicting variables and their interactions for all 5 EDI domains, vulnerability and multiple vulnerability caused a loss of model fit for the Communication & general knowledge domain, vulnerability and multiple vulnerability. However, if in case of considering models with good fitness for all seven continuous and binary outcomes including the three mentioned outcomes, there would be a potential loss of the homogeneity of expressing results across all seven outcomes based on their associated predictors. Thirdly, the concept of 'neighborhood' for major cities (Saskatoon, Regina, and Prince Albert) and non-urban areas were in essence different as neighborhoods in major cities as in the first it was defined by each municipality whereas in the second it was defined for the purpose of this research. Fourthly, for continuous 5 EDI domains as outcome variables, there was no possibility to use non-identity transforms to make their distributions as possible as normal due to the fact any transformation other than identity function would create nuisance factor in differences of outcomes preventing plausible interpretations. Finally, the super cluster effect was significant for only the Emotional maturity domain and was insignificant with large p-values for the other 4 EDI domains, vulnerability and multiple vulnerability. This issue is accompanied with inflated neighborhood level VPC and underestimated geographical area level VPC. The source of this problem was defining city as a town with a minimum population of 35,000 (not 10,000) inhabitants, causing the exclusion of other provincial cities such as Estevan, Moose Jaw, North Battleford, Swift Current, Weyburn and Yorktown,[60] from geographical area level variables and their inclusion as neighborhoods of the non-urban areas. In this study, a three-level model of geographical area-neighborhood-child hierarchical data was considered. A geographical area-neighborhood-school-child four-level model was considered, but after investigating the data structure it was determined that school-level and neighborhood-level data were mutually nested, as children from different neighborhoods could attend the same school and several schools could be in the same neighborhood. Regarding the importance of neighborhood level socioeconomic variables in this study, school level was removed from the potential four-level model in favour of neighborhood level.

CHAPTER 6: CONCLUSION

Children are the bridges to heaven.
-Persian Proverb

The current study used a combination of 2008-2009 EDI data of Saskatchewan children and 2006 Census data and applied multilevel linear and logistic models in order to provide some insight regarding how individual, neighborhood contextual and geographical factors and their within-level and cross-level effects determine children's developmental health and vulnerability.

Individual characteristics of Aboriginal status, ESL status, male status and school absenteeism were associated with lower EDI averages and higher odds of vulnerability in terms of significance and or strength. Also, neighborhood contextual characteristics contributed to at least 11% of the child developmental health outcome variations. Neighborhood income inequality was associated with lower EDI averages and higher odds of vulnerability, giving further evidence for the income inequality hypothesis. However, neighborhood median income had inverse effects. Furthermore, children living in Regina had the lowest average EDI outcomes and the highest proportions of vulnerability, while Aboriginal children living in Prince Albert had lower average EDI outcomes and higher odds of vulnerability compared to non-urban areas.

Compound effects of Aboriginal status─number of skills, major city─Aboriginal status and major city─neighborhood median income were positive on EDI outcomes and mitigating vulnerability. In details, though Aboriginal children had lower EDI scores and higher odds of vulnerability, having more special skills or living in big city mitigated the gaps. Furthermore, though children living in big cities had lower EDI scores and higher odds of vulnerability, living in neighborhoods with higher median income mitigated the gaps.

Compound effects of neighborhood median income─neighborhood income inequality, school absenteeism─neighborhood income inequality, Aboriginal status─neighborhood income inequality, and major city─neighborhood income inequality were negative on EDI outcomes and exacerbating vulnerability. In details, living in neighborhoods with higher income inequality lowered scores of EDI domains and exacerbated odds of vulnerability for children who were living in neighborhoods with higher median income, had higher absent days from school, were Aboriginal, or living in big cities.

In terms of child public health policy, stakeholders, school policy-makers, and administrators' initiatives should focus on children with Aboriginal status, ESL status, males, and those with more days absent from school and who are living in neighborhoods with high income inequality. A recommendation is that the stakeholders design and promote child health programs that increase Aboriginal children's skills and school policy makers and administrators consider policies that minimize days absent from school for children living in neighborhoods with high income inequality or a high Aboriginal population. Also, on a large scale, there is a need to promote more child developmental health supporting programs by authorized institutions in the cities of Regina and Prince Albert.

On the basis of these findings, future research should continue to examine and clarify the significance and the strength of association between the above predictors and their compound effects on child developmental health status by considering a longitudinal design and inclusion of more small cities in the hierarchy.

REFERENCES

1. Silver R. B, Measelle J. R, Armstrong J. M & Essex M. J. Trajectories of Classroom Externalizing Behaviors: Contributions of Child Characteristics, Family Characteristics, and the Teacher-Child Relationship during Transition to School. *Journal of School Psychology* 2005; 43, 39-60.

2. Tremblay, R.E. & Masselink, B. Early Disruptive Behavior, Poor School Achievement, Delinquent Behavior, and Delinquent Personality: Longitudinal Analysis. *Journal of Consulting and Clinical Psychology* 1992; 60, 64-72.

3. Bongers, I. L, Koot H. M, Vander Ende J, & Verhulst F. C. The Normative Development of Child and Adolescent Problem Behavior. *Journal of Abnormal Psychology* 2003; 112, 179-192.

4. Kagan, S.L. Readiness Past, Present, and Future: Shaping the Agenda. *Young Children* 1992; 48, 48-53.

5. Snow, K.L. Measuring School Readiness: Conceptual and Practical Considerations. *Early Education and Development* 2006; 17, 7-41.

6. Janus, M, Brinksman, S, Duku, E, Hertzman, C, Santos, R, Sayers, M, Schroeder, J, & Walsh, C. The Early Development Instrument: Population Based Measure for Communities. *Offord Center for Child Studies*, McMaster University, 2007.

7. Lapointe V. R., Ford, L & Zumbo, B. D. Examining the Relationship Between Neighborhood Environment and School Readiness for Kindergarten Children. *Early Education and Development* 2007; 18, 3, 473-495.

8. Cushion, J. A, Vu, L. T. H, Janzen, B. L, & Muhajarine, N. Neighborhood Poverty Impacts Children's Physical Health and Well Being over Time: Evidence From the Early Development Instrument, *Early Education & Development* 2011; 22:2,183-205.

9. Hertzman, C. Social Determinants of Children's Developmental Health: Research Proposal, *Canadian Institute of Health Research*, 2012.

10. Saskatchewan Regional Intersectional Committees, Government of Saskatchewan, Ministry of Education, Human Services, 2012. Available at: http://www.education.gov.sk.ca/ Regional-Integrated-Services/Human-Services/RICs (Access: October 10, 2012).

11. Mitch, B, Stewart- Brown, S, Waterston, T & Crowther, R. Child Public Health.2ed edition. *Oxford University Press* 2010; pp. 1-9, 103-130.

12. Winslow, C, & Edward A. The Untitled Fields of Public Health. *Science* 1920; 51:1306, 23 – 33.

13. Trawick, S. J. Early Childhood Development: A Multicultural Perspective. 4th edition. *Pearson Education Inc* 2006; 1-10.

14. Rebello Britto P, Engle P. L, & Super C.M. Handbook of Early Childhood Development Research and Its Impact on Global Policy. *Oxford University Press* 2013; 3-13.

15. Richter L. An Introduction to Family Centered Services for Children Affected by HIV and AIDS. *Journal of International AIDS Society* 2010; 13 (2), S1.

16. Shonkoff, J. & Phillips D. (Eds). From Neurons to Neighborhoods: The Science of Early Childhood Development. Committee on Integrating the Science of Early Childhood Development, National Research Council and Institute of Medicine. Washington, DC: *National Academy Press,* 2000.

17. Knudsen, E. Heckman, J. J, Cameron, J. L, & Shonkoff, J. P. Economic, Neurobiological and Behavioural Perspectives on Building America's Future Workforce. *Proceedings of the National Academy of Science* 2006; 103, 10155 – 10162.

18. Bronfenbrenner, U. "The Bioecological Prospective From a Life Course Perspective: Reflections of a Participant Observer" In p. Moen, G.H. Edler, & K. Luscher (Eds.). Examining Lives in Context. *Washington Dc. American Psychological Association* 1995; 599-618.

19. Robinson, J.L. Are There Implications for Prevention Research from Studies of Resilience? *Child Development* 2000; 71, 570 – 572.

20. Roosa, M.W. Some Thoughts about Resilience Versus Positive Development, Main Effects Versus Interactions, and the Value of Resilience, *Child Development* 2000; 71, 567 – 569.

21. Norrie MacCain, M, Mustard, & F, Shanker, S. Early Years Study 2: Putting Science into Action. *Council for Early Child Development* 2007; 136.

22. Miller, L. J. First STEP Screening Test for Evaluating Preschoolers. San Antonio: *The Psychological Corporation*, 1993.

23. Dunn, L. M. & Dunn, L. Peabody Picture Vocabulary Test Revised. Manual for Forms L and M. Circle Pines, *Minnesota: American Guidance Service, 1981.*

24. Zady, Madelon F. Correlation and Simple Least Squares Regression, 2000. Available at: *http://www.westgard.com/lesson42.htm.* (Retrieved November 1, 2013)

25. Dubois, N. F, Lemelin, J. P, Boivin, M, Dionne, G, Seguin, J. R, Vitaro, F & Tremblay, R. E. Predicting Early School Achievement with the EDI: A Longitudinal Population-Based Study, *Early Education and Development* 2007; 18:3, 405-426.

26. Chew, A. L. The Lollipop Test: A Diagnostic Screening Test of School Readiness. Atlanta, GA: Humanics Limited, 1989.

27. Janus, M., Walsh, C., & Duku, E. Early Development Instrument: Factor Structure, Subdomains and Multiple Challenge Index. Hamilton, Ontario: Offord Center for Child Studies, McMaster University, 2005.

28. Muhajarine, N, Puchala, C, & Magdalena, J. Does EDI Equivalently Measure Facets of School Readiness for Aboriginal and Non-aboriginal Children? *Soc Indic Res* 2011; 103: 299 – 314.

29. Janus, M, and Duku, E. The School Entry Gap: Socioeconomic, Family, and Health Factors Associated with Children's School Readiness to Learn. *Early Education & Development* 2007; 18 (3), 375 – 403.

30. Oliver, L. N, Dunn, J. R, Kohen, D. E & Hertzman, C. Do Neighborhoods Influence the Readiness to Learn of Kindergarten Children in Vancouver? A Multilevel Analysis of Neighborhood Effects. *Environment and Planning* 2007; 39: 848-868.

31. Puchala, C, Vu L, & Muhajarine, N. Neighborhood Ethnic Diversity Buffers School Readiness Impact in ESL Children. *Canadian Journal of Public Health* 2010; (101), S13-S18.

32. Leeuw, J.de, & Meijer, E (eds). Handbook of Multilevel Analysis. *Springer* 2008; 1-5.

33. Fidell, B.G, & Tabachnick, L.S. Using Multivariate Statistics (5 th ed). Boston; *Pearson/A&B*, 2007.

34. Diez Roux, A.V. A Glossary for Multilevel Analysis. *Journal of Epidemiological Community Health* 2002; 56, 588 – 594.

35. Cohen, J, Cohen, P, West, S. G & Alken, L. S. Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences (3ed). Mahwah, NJ: *Lawrence Erlbaum Associate*, 2002.

36. Leckie, G. Three Level Multilevel Models- Concepts. LEMMA VLE Module 11, 2013; 1-47.

37. Leeuw, I. and Kreft, J.d. Introducing Multilevel Modelling. London. *Sage Publications Ltd*, 1998.

38. Leyland, A.H. & Goldstein, H. Multilevel Modeling of Health Statistics. *John Wiley & Sons, LTD*, 2004; *1 – 12.*

39. Bryk, A.S. and Raudenbush, S. W. Hierarchical Linear Models: Applications and Data Analysis Models, *Sage Publications* 1992; 155-174.

40. Gebbie, K, Rosenstock, L., & Hernandez, L. M. Who Will Keep the Public Healthy? Educating Public Health Professionals for 21st Century. Washington, D. C.: *The National Academic Press*, 2003.

41. Maddala, G. S. Limited-Dependent and Qualitative Variables in Econometrics, *Cambridge University Press*, Cambridge, UK 1983; 68.

42. McCullagh, P. & Nelder, J. A. Generalized Linear Models, 2nd edition. *Chapman & Hall, London* 1989; 20,144.

43. Hosmer, D. W, Lemeshow, S & Sturdivent, R. X. Applied Logistic Regression, 3d edition. *Wiley* 2013; 91.

44. Bendel R. B. & Afifi A. A. Comparison of stopping rules in forward regression. *Journal of the American Statistical Association* 1977; 72, 6-53.

45. Mickey J. & Greenland, S. A study of the impact of confounder-selection criteria on effect estimation. *American Journal of Epidemiology* 1989; 129, 125-137.

46. Kleinbaum, David G., Kupper, L. L, Muller, K. E & Nizam, A. Applied Regression Analysis and Other Multivariable Models, 3d edition, *Duxbury Press* 1998; 400-401.

47. Janus, M. & Duku, E. Normative Data for the Early Development Instrument, 2004. Available at: *http: //www.offordcenter.com/readiness/files/ results. normative_data.pdf.* (Retrieved September 27, 2013)

48. Guhn, M, Gadermann, A. M, Hertzman, C & Zumbo, B. D. Children's Development in Kindergarten : A Multilevel, Population Based Analysis of ESL and Gender Effects on Socioeconomic Gradient, *Child Ind Res* 2010; 3: 183-203.

49. Guhn, M., Gadermann, A., & Zumbo, B.D. Does EDI Measure School Readiness in the Same Way Across Different Groups of Children? Early Education and Development 2007; 18, 453-472.

50. Carpiano, R. M., Lloyd , J.E.V, & Hertzman, C. Concentrated Affluence, Concentrated Disadvantage, and Children's Readiness for School : A Population-based, Multilevel Investigation, *Social Science & Medicine* 2009; 69, 420-432.

51. Muhajarine, N, Sari, N, Pahwa, P & Soltanifar, M. Social Inequalities and Health disparities among Kindergarten Students: Evidence from a Medium Size Canadian City, In Press, 2014.

52. Kershaw, P, Irwin, L, Trafford, K & Hertzman, C. The British Columbia Atlas of Child Development, 1st ed , Vancouver: Human Early Learning Partnership, *Western Geographical Press* 2005; 40.

53. Lubienski, S. T, Lubienski, C., & Crance, C, C. Achievement Differences and School Type: The Role of School Climate, Teacher Certification, and Instruction, *American Journal of Education* 2008; 115(1), 97-138.

54. Braun, H, Jenkins, F. & Grigg, W. Comparing Private Schools and Public Schools Using Hierarchical Linear Modelling (NCES 2006-461). U.S. Department of Education, National Center for Education Statistics, *Institute of Education Science*. Washington, DC: U.S. Government Printing Office, 2006.

55. Olson, M. E, Diekema, D, Elliot, B. A, & Renier, C. M. Impact of Income and Income Inequality on Infant Health Outcomes in the United States, *Pediatrics* 2010; 126 (116). DOI: 10.1542/Peds.2009-3378.

56. Subramanian, S.V., Kawachi, I & Kennedy, B.P. Does the State You Live in Make a Difference? Multilevel Analysis of Self-rated Health in the U.S., *Social Science & Medicine* 2001; 53, 9-19.

57. Pickett, K.E., & Pearl, M. Multilevel Analysis of Neighborhood Socioeconomic Context and Health Outcomes: A Critical Review, *Journal of Epidemiological and Community Health* 2001; 55, 111-122.

58. Dye, C. Health and Urban Living, *Science* 2008; 319, 766-768.

59. Gunn, J.B. & Duncan G.J. The Effects of Poverty on Children, *The Future of Children*, Princeton University Press 1997; 7(2), 55-71.

60. Population and Dwelling Counts for Canada, Provinces and Territories, and Census Subdivisions (Municipalities), 2011 and 2006 Census (Saskatchewan), Statistics Canada, 2013. Available at: http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/TableTableau.cfm?LANG=Eng&T=302&SR=1&S=51&O=A&RPP=9999&PR=48&CMA=0 (Retrieved November 9, 2013).

INDEX

M

N

O

P

Q

R

S

T

U

V

W