FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

CAMILA MATA MACHADO SOARES

EARLY CHILDHOOD DEVELOPMENT IN VULNERABLE NEIGHBORHOODS: the case of São Paulo

São Paulo
2021

CAMILA MATA MACHADO SOARES

EARLY CHILDHOOD DEVELOPMENT IN VULNERABLE NEIGHBORHOODS: the case of São Paulo

Tese apresentada à Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas como requisito parcial para a obtenção do título de Doutor em Administração Pública.

Orientador: Prof. Dr. Ciro Biderman

São Paulo

2021

Soares, Camila Mata Machado.

Early childhood development in vulnerable neighborhoods : the case of São Paulo / Camila Mata Machado Soares. - 2021.

172 f.

Orientador: Ciro Biderman.

Tese (doutorado CDAPG) – Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo.

1. Crianças - Desenvolvimento. 2. Educação de crianças. 3. Igualdade. 4. Bairros - São Paulo (SP). 5. Creches. I. Biderman, Ciro. II. Tese (doutorado CDAPG) – Escola de Administração de Empresas de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 373.22(816.11)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

CAMILA MATA MACHADO SOARES

EARLY CHILDHOOD DEVELOPMENT IN VULNERABLE NEIGHBORHOODS: the case of São Paulo

Tese apresentada à Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas como requisito parcial para a obtenção do título de Doutor em Administração Pública.

Data da aprovação: 14/09/2021

Banca examinadora:

Prof. Dr. Ciro Biderman
(Escola de Administração de Empresas de São Paulo - FGV)

Prof. Dr. Rudi Rocha
(Escola de Administração de Empresas de São Paulo - FGV)

Prof. Dr. Naércio Menezes
(Universidade de São Paulo)

Prof. Dr. Ricardo Paes de Barros
(Insper)

AGRADECIMENTOS

Agradecer, após todo o (longo) percurso desse doutorado, não é apenas mera formalidade, mas também necessidade. São muitas (e muitos) as que tornaram essa jornada menos extenuante, menos solitária e mais divertida.

Agradeço ao meu orientador, Prof. Ciro Biderman, pelo apoio constante durante os quatro anos e meio de doutorado. Sua confiança no meu trabalho e nas minhas ideias, combinada a sua orientação, muitas vezes pragmática, foram indispensáveis nesse percurso.

Aos Professores Naércio Menezes e Rudi Rocha pela participação na banca de defesa e por todos os excelentes comentários e sugestões. Ao Prof. Ricardo Paes de Barros pela participação na banca, mas também pelas inúmeras trocas e aprendizados ao longo dos últimos anos.

Aos professores, funcionários e amigos da EAESP-FGV, por todo o suporte ao longo dos últimos anos. Agradeço também à CAPES pelo financiamento desse trabalho.

Ao professor Mario Small e o Departamento de Sociologia da Universidade de Harvard, por me receberem como Visiting Fellow e proporcionarem uma rica experiência de troca acadêmica.

Ao FGV EESP Clear, e toda a equipe, pela parceria nesse último ano e meio de pandemia e de escrita da tese. Agradecimento especial à Lycia e à Lara, pela confiança no meu trabalho, pelo apoio constante, pelo carinho sempre presente e por serem inspiração.

Às muitas amizades, das mais antigas às mais recentes, do mundo acadêmico às da vida profissional, meu eterno obrigada por me aturarem e por me acolherem. Muito obrigada por me inspirarem das mais diversas formas possíveis e me apoiarem ao longo dos últimos anos. Cada cerveja, risada, café, ioga e conversa compartilhada contribuíram (muito) nesssa caminhada.

Agradecimento especial às mulheres maravilhosas que tornam o percurso menos duro e mais vivo. Aquelas que me acompanham desde o colégio, as que entraram na minha vida graças à FACE/UFMG, as que o Rio me trouxe, as que São Paulo me apresentou. Agradecimento ainda mais especial às que me escutam diariamente, seja ao vivo e a cores (muito obrigada, Aninha!), seja por (muitas!) mensagens e por áudios (muito obrigada, Pri, Bela, Yanna, Lari, Ju, Marina [e Rica]!).

Aos companheiros da estadia em Boston, Mariah, Eduardo e Lucas, muito obrigada pelas trocas, pelo acolhimento e pela jornada compartilhada. Muito obrigada também à Alice e ao Luiz por serem (sempre) família longe do Brasil.

Finalmente, agradeço aos meus pais e meus irmãos. Aos meus pais por terem me ensinado a questionar, por me apoiarem sempre e também por me inspirarem constantemente. Aos meus irmãos pelo carinho e pelo apoio constante das maneiras mais diversas possíveis, seja no café entregue semanalmente, seja na resolução de (quase todas) as dúvidas tecnológicas.

ABSTRACT

This thesis discusses two main themes and their relation to early childhood development; neighborhood effects and center-based care. The first chapter presents a systematic review of the literature analyzing neighborhood effects on early childhood development of children between zero and six years old. The second chapter describes São Paulo, the focus of this thesis. Spatial segregation in the city is described at length using the *Índice de Vulnerabilidade Social*and the city's early childhood education and care (ECEC) policy is also presented. The third chapter provides evidence of the existence of neighborhood effects on children's educational outcomes in São Paulo by estimating a hierarchical model. To do so, data from *Provinha São Paulo* from 2018 is used. This empirical result is in line with the international evidence systematized in the first chapter. The fourth chapter uses administrative data from the city's enrollment system to understand the profile of those enrolled in the ECEC system throughout the years of 2010 and 2018. The profile of those entering the system is changing but the evidence is mixed as to whether this change is moving towards the inclusion of a more diverse population. Finally, the fifth chapter presents an estimation of the impact of daycare enrollment on children's cognitive abilities. Both the administrative database as well as data from *Provinha São Paulo* are used to implement a regression discontinuity design. The estimation provides evidence of a positive, marginally significant effect on a child's proficiency in mathematics on being offered a spot in the first level of the ECEC system. Together, the results presented in this thesis provide new evidence on both neighborhood effects and the importance of center-based care in São Paulo.

Key-words: early childhood, neighborhood effects, urban inequality, center-based care.

RESUMO

Essa tese discute a relação entre dois temas - efeitos de vizinhança e a matrícula em creches - e sua relação com o desenvolvimento infantil. O primeiro capítulo apresenta uma revisão sistemática dos efeitos de vizinhança no desenvolvimento infantil de crianças de zero a seis anos. O segundo capítulo descreve São Paulo, foco dessa tese. Nele, a segregação espacial na cidade é amplamente descrita, utilizando o *Índice de Vulnerabilidade Social*; também é apresentada a política de educação e atenção à primeira infância do município, com foco especial nas creches. O terceiro capítulo fornece evidências sobre a existência de efeitos de vizinhança em São Paulo através da estimação de um modelo hierárquico que analisa tais efeitos sobre os resultados educacionais das crianças. Para tanto, são utilizados os dados da *Provinha São Paulo* de 2018. Esse resultado empírico está em linha com as evidências internacionais sistematizadas no primeiro capítulo. O quarto capítulo utiliza dados administrativos do sistema de matrícula do município para compreender o perfil dos que tentaram se matricular nas creches municipais entre 2010 e 2018. O perfil dos ingressantes mudou; mas as evidências não permitem afirmar que essa mudança foi no sentido de uma inclusão maior de uma população mais diversa. Finalmente, o quinto capítulo apresenta uma estimativa do impacto da matrícula em creches sobre os resultados educacionais das crianças. Tanto o banco de dados administrativo, quanto os dados da *Provinha São Paulo* são usados para implementar uma regressão descontínua. A estimativa fornece evidências de um efeito positivo e marginalmente significativo da oferta de vaga no berçario I sobre a proficiência em Matemática. Juntos, os resultados apresentados nesta tese fornecem novas evidências sobre os efeitos de vizinhança e a importância das creches in São Paulo.

Palavras-chave: primeira infância, desigualdade, bairros, desenvolvimento infantil, creche.

List of Figures

List of Tables

Contents

Introduction

Early childhood is a crucial and sensitive developmental phase that shapes the foundation of a person's life course (Center on the Developing Child at Harvard University, ; HECKMAN, 2008). Healthy and integral early childhood development can be understood as the "ordered progression of perceptual, motor, cognitive, language, socio-emotional, and self- regulation skills"(BLACK et al., 2017). It provides a strong foundation for later well-being (Center on the Developing Child at Harvard University, ; HECKMAN, 2008; BRITTO et al., 2018) and the lack of adequate early childhood development can have lasting consequences not only on the individual's life, but also on their capability to participate fully in society (Center on the Developing Child at Harvard University,).

The importance of early childhood development has been acknowledged by society as a whole and by governments worldwide. A clear example of this is its inclusion in the Sustainable Development Goals (BRITTO et al., 2018). In Brazil, the approval of the *Marco Legal da Primeira Infância*[1] showcases the importance of this agenda in the country.

To have healthy developmental trajectories, young children need warm, responsive and safe environments in which they have the opportunity to explore, to play and to engage with others (SHONKOFF; PHILLIPS; (U.S.), 2000; Center on the Developing Child at Harvard University,). To reach their full developmental potential, children need an enabling environment that allows them to receive nurturing care from their caregivers, their family and their community (BLACK et al., 2017). The environment a child experiences matters whether it be it their home environment or their school environment (SHONKOFF; PHILLIPS; (U.S.), 2000; BRITTO et al., 2018; BLACK et al., 2017).

Following a recent literature trend, this thesis focuses on one specific environment: the neighborhood, understood as the community surrounding a child's home. From a theoretical perspective, the importance of the neighborhood is stated in the socio-ecological model of Bronfenbrenner e Morris (1998), but recent empirical evidence corroborates the connection between children's development and their neighborhood. In chapter 1, a systematic review of recent evidence is presented and it indicates that there is robust evidence of an association between a healthy child development and (i) the socioeconomic conditions of the neighborhood and (ii) the existence of a solid social network among its residents.

Understanding the association between neighborhood characteristics and a healthy child development can be especially relevant in an unequal country like Brazil. As dis-

[1] Law nº 13.257/2016.

cussed in chapter 2, the country and its largest city, São Paulo, the focus of this thesis, are marked by inequality, including spatial inequality. Focusing in São Paulo, chapter 2 presents a descriptive analysis of the status of the spatial segregation. Even though the pattern of center-periphery is questioned, segregation is still present and there are many neighborhoods with inadequate conditions.

What are the consequences of such a scenario for the children? To advance the understanding of the role of neighborhoods in a young child's development, chapter 3 empirically tests the existence of neighborhood effects on children's development. To do so, this chapter presents an analysis of how a neighborhood's social vulnerability is associated with children's cognitive skills in second grade, using a database from *Provinha São Paulo* - the city's standardized test for second-grade students. The results corroborate the international literature and indicate that there is a positive and robust association between children's achievement and their neighborhood's social vulnerability.

In the Brazilian context, one of the main public policy for young children is the provision of free center-based care which is the second main theme of this thesis. In chapter 2, a brief descriptive analysis of this policy, in the city of São Paulo, is presented. The access to this service is extremely connected to the neighborhoods families reside in as childcare centers give priority to families living nearby. This service can, then, be understood to be an important neighborhood characteristic, an institutional resource - using Sampson, Morenoff e Gannon-Rowley (2002) nomenclature.

The literature indicates that center-based care, and particularly those that are publicly provided, are central to a healthy and integral child development: especially for children from vulnerable backgrounds where high quality center-based care can have profound positive impacts on their development (SHONKOFF; PHILLIPS; (U.S.), 2000; BROOKS-GUNN, 2003). Additionally, childcare centers can also play an important role as brokers for vulnerable families in helping them reach other services (SMALL; JACOBS; MASSENGILL, 2008; ROOS; WALL-WIELER; LEE, 2019). The provision of center-based care becomes even more crucial when combined with the fact that women are more likely to be working outside of the house and that more traditional care arrangements such as that of a grandmother caring for her grandchildren, are becoming less common in our modern society.

To understand how the two main themes, neighborhood and center-based care, are connected, first, in chapter 3, the possible role of center-based care as a moderator of the association between neighborhood and children's cognitive outcomes is analyzed. No evidence of such role is found. Then, in chapter 4, the access to center-based care in São Paulo is analyzed at length. The importance of different family and child characteristics is estimated, along with an analysis of the distribution of spots throughout the city. The findings from this chapter indicate that, although the profile of those beneficiaries

of center-based care has changed with the expansion of number of spots in the system, there is no inclusion of the more vulnerable population in it. The evidence indicates that less vulnerable parents are more likely to enroll their children in center-based care before age three and children living in less vulnerable neighborhoods are also more likely to benefit from such a policy. There is, therefore, evidence that access to center-based care is connected to neighborhood characteristics.

In chapter 5, a regression discontinuity design is employed to estimate the impact of center-based care enrollment on children's cognitive skills as measured by the Provinha São Paulo. There is evidence of a positive, marginally significant effect on the child proficiency in mathematics on being offered a spot in the first level of the early childhood education and care system.

This thesis contributes to the existing literature in several ways. First, it provides an important contribution to the urban studies literature: whereas most of the neighborhood effects studies were conducted in developed countries, this study presents evidence to the literature on neighborhood effects in a new setting: a large, urban center in the Global South. Moreover, it contributes to the early childhood literature as it estimates the impact of a large scale early childhood development program: the provision of free, center-based care for children between zero and three years old, also in an under-researched context. These contributions are discussed in chapter 6 where policy implications are presented at length.

1 The State of Art of the Literature on Neighborhood Effects on Child Development

1.1 Introduction

From the work of sociologists like William Wilson and developmental psychologists such as Urie Bronfenbrenner, many have theorized about the importance of the neighborhood for integral early childhood development. The former, in describing the spatial concentration of poverty, addressed possible consequences of this phenomenon and its impacts on early childhood development were a clear continuation of these studies. On the other hand, Urie Bronfenbrenner's seminal work defined the ecological model of child development which indicated the importance of studying child development in the context in which it happens. Bronfenbrenner's work shed light on the centrality of studying the different contexts and environments that a child interacts with during their development, including the neighborhood.

These works led to the development of a field of study that is dedicated to understanding, both theoretically and empirically, how and by what characteristics neighborhoods affect integral child development. Even though the concept that a child grows up in a neighborhood and that this influences their development is not an innovation, there is still little understanding of the many ways that such influence happens. Researchers from many fields have contributed to it; from sociologists and developmental psychologists to economists and public health researchers and many have been dedicated to this question. Reviews from Leventhal e Brooks-Gunn (2000) and Sampson, Morenoff e Gannon-Rowley (2002) to more recent reviews from Minh et al. (2017) and Christian et al. (2015) have consolidated many of these findings, reinforcing the existence of neighborhood effects and the need for even more research to help understand specific neighborhood characteristics that affect children's development.

This chapter contributes to this field by conducting a thorough literature review of the evidence of neighborhood effects on early childhood development. This review indicates that there is robust evidence of an association between positive child development and two specific neighborhood characteristics; the level of socioeconomic vulnerability and the existence of a cohesive social network among residents. No such robust conclusion can be reached for other characteristics. Therefore, our review indicates the need for more research focussing on neighborhood characteristics outside the social and socioeconomic domains.

The work presented here also attempted to include more than just quantitative

evidence. Unfortunately, the effort to include more diverse, methodologically speaking, evidence was not very fruitful.

In addition to this introduction, this chapter has four other sections. In the next section the objective of this chapter is presented. In the third section the methodology used to review and systematize the evidence is discussed in detail. The fourth section presents the results by domain. The fifth analyzes the findings and limitations of the present study.

1.2 The current study

This study presents a systematic review of the existing evidence of neighborhood effects on child development. To do so, we systematically searched ten academic databases to select articles published in peer-reviewed journals between 1999 and 2019.

This study extends the literature by focusing on the effect the neighborhood has on the development of children between the ages of zero and six years old. This study also innovates as it attempts to include qualitative and mixed methods evidence by not excluding articles based on method.

To structure the analysis of the evidence collected, we use the theoretical model of Goldfeld et al. (2015), that analyzes five domains of the neighborhood that may influence a child's development: physical, social, service, socioeconomic and governance. This is another contribution of this study to the literature.

The main research question of the study is:

- • Does the literature present robust evidence that neighborhood characteristics as organized in Goldfeld et al. (2015) five domains, affect children's development?

1.3 Methodology

To understand what it is known so far about how the neighborhood the child grows up in affects their early childhood development, a systematic review was conducted of articles published in peer-reviewed journals from 1999 to 2019. The following databases were searched: Web of Science, PubMed, Embase, PsycINFO, GenderWatch, ERIC, Urban Studies, Academic Search Premier, Social Science Premium Collection, EconLit, LILACS and Scielo. These databases were chosen as they represent different fields of study that look at this particularly interdisciplinary question; Medicine/Health, Psychology, Social Science in general, Women's Studies, Urban Studies and Education are represented, as well as two specific databases for Latin America.

To conduct the search, two keywords were used: neighborhood and early childhood[1]. In each database, the search procedure was modified so to consider only papers that had both terms[2] in the title, abstract or in its keywords[3]. The search focused on papers published between 1999 and 2019[4] and covers only articles published in peer-reviewed journals[5].

In the present study, the neighborhood is understood as the small area near and around the child's home. There is no single way to operationalize this concept and there is no consensus in the literature on the best way to do it. In this review, there are studies that define neighborhood in different ways: (i) by census tract or block group; (ii) by report of the survey respondent[6]; and (iii) by authority definition[7]. We excluded articles that considered counties or school districts as neighborhoods, as we were not interested in the larger community around one's home. We also excluded articles that focused solely on rural neighborhoods, as we were specifically interested in the urban neighborhoods. Finally, articles that focused on the neighborhood around the school, or daycare center, attended by the child, were also excluded.

It is also important to mention that we are interested in the impacts of neighborhood characteristics on broadly defined early child development, including physical and motor development, as well as socio-emotional and cognitive development. We are not discussing here the evidence of neighborhood effects on diseases - such as asthma or developmental disorders - nor the impact on neglect or maltreatment. Articles focusing on these aspects were not considered in this review.

The inclusion criteria for the studies were the following: (i) empirical studies; (ii) general population of children between zero and six years old; (iii) focus on early child development outcomes[8]; (iv) studies published from January of 1999 to September of 2019; (v) studies published in English, Spanish and Portuguese; and (vi) studies published in peer-reviewed journals. There was no specification regarding the location of the study. There was also no specification on the types of study design employed or the methodology.

[1] Research was conducted in both Portuguese and English. In Portuguese, the terms were *bairro* or *território* or *vizinhança* and *criança*. Other terms were tested such as "early childhood development", but analysis suggested that they excluded too many articles.

[2] The AND operator was used. It is also important to note that early childhood was searched as a unique term, by using "".

[3] This was an option in most databases. The exception was the LILACS and the Scielo databases.

[4] Leventhal e Brooks-Gunn (2000) focus on the period from 1990 to 1998 and so, we are covering the academic production not covered by this seminal meta-analysis.

[5] Grey literature was not covered in this work.

[6] In the articles in this review, most of the survey's respondents are the child's caregivers. In this case, each parent may have a unique concept of what their neighborhood would be like.

[7] That is, by using administrative divisions of cities or counties.

[8] Defined in a broader sense, including physical, socio-emotional, language, social and cognitive development. This definition does not include diseases - either mental or physical ones - as we are interested in development and not in health. However, it does include behaviours, both internalizing and externalizing.

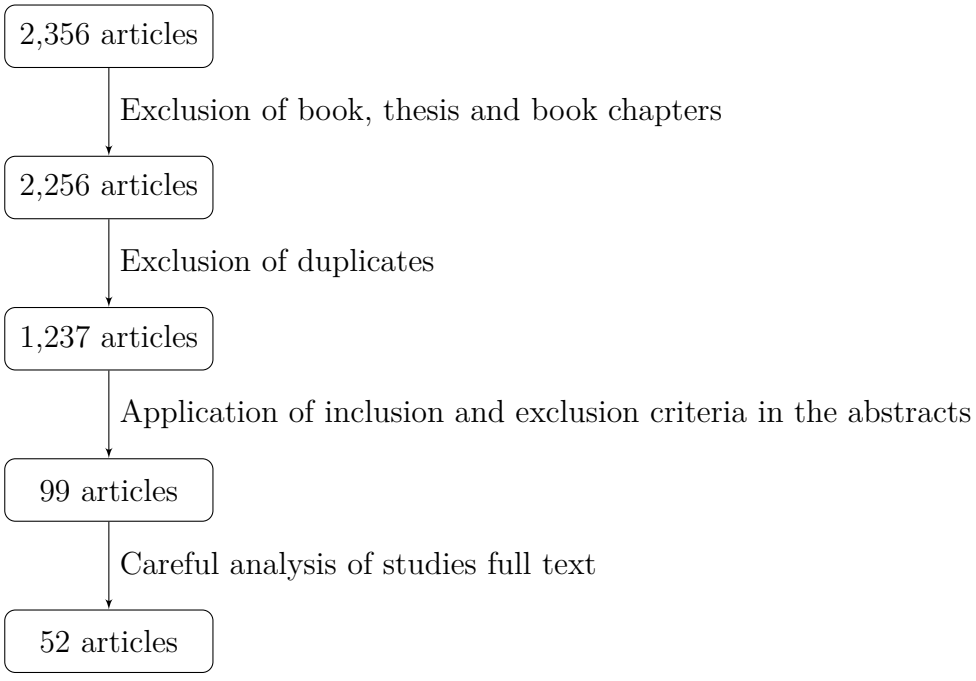
As discussed previously, this was kept open to allow for the inclusion of evidence produced by diverse methodologies.

The exclusion criteria were the following: (i) reviews; (ii) meta-analysis; (iii) editorials; (iv) opinion pieces; (v) book reviews; (vi) empirical studies that focused in a specific population of children (e.g.: immigrant children, aboriginal children, autistic children[9]; and (vii) studies that used the neighborhood as the context of study but did not analyze its characteristics.

The initial search in the twelve selected databases using the previously defined keywords, yielded 2,356 articles. These articles were organized using the software Zotero. Duplicated articles, books and book chapters[10]were excluded resulting in 1,237 articles remaining. These articles were then uploaded to the software Rayyan where the inclusion and exclusion criteria were applied to the remaining articles' abstracts. After this process, the full text of 99 articles was read. 42 studies were then excluded due to exclusion and inclusion criteria. The final sample included 52 articles.

The flowchart of the entire article selection process is presented in Figure 1.1.

Figure 1.1 – Search Process



Source: Elaborated with the 52 selected articles.

After the 52 articles were selected, we use a conceptual model, defined by Goldfeld et al. (2015), to organize the evidence and facilitate the analysis. Goldfeld et al. (2015) define five domains that affect early childhood development. These domains —

[9] Studies that focused in high risk populations, defined through socioeconomic criteria, were included.
[10] Some databases did not allow to select only papers published in peer-reviewed journals.

socioeconomic, service, social, physical, and governance, encompass the characteristics of neighborhoods that influence child development and should be understood as interconnected, composing a "a dynamic, eco-epidemiological approach to understanding children's developmental outcomes with a focus on constructs that may be amenable to change"(GOLDFELD et al., 2015). The five dimensions underscore factors that are not generally analyzed by the neighborhood effect literature such as the governance dimension and by doing so, shed light on processes that were previously ignored but which could be of importance for early childhood development.

Each one of the 52 articles were classified according to the model's domains. Articles could be classified in more than one domain if they analyzed different neighborhood characteristics. In the next section, our results are presented by domain.

Before we proceed, it is important to point out that no articles that investigated the governance domain were selected. This domain includes citizen participation in public decisions and also characteristics of the governance environment of initiatives taking place in the neighborhood. As already mentioned by Goldfeld et al. (2015), there is little empirical evidence related to child development for this domain.

1.4 Analysis of Evidence

Table 1.1 presents the 52 articles and a systematization of their main information. In this section, the evidence of all the selected articles is analyzed by neighborhood domain. For each domain, we present its definition, the number of articles that were analyzed in it and a systematization of the findings.

The socioeconomic domain This encompasses the socioeconomic status of the population residing in the neighborhood (GOLDFELD et al., 2015). Our search found twenty-six studies that analyzed this domain and its relationship with child development.

In 21 studies the socioeconomic domain is either characterized by a single measure; either neighborhood median income or the neighborhood poverty level, or by a composite index[11]. Taken together, these studies provide evidence that neighborhood socioeconomic status is negatively related to child development as measured by broad measures of child development (MILBRATH; GUHN, 2019; BROWNELL et al., 2016; LOVASI et al., 2011), cognitive ability (COULTON et al., 2016; MORRISSEY; VINOPAL, 2018b; WOLF; MAGNUSON; KIMBRO, 2017; BENSON; BORMAN, 2010; KOHEN et al., 2002; HANSON et al., 2011; CAUGHY; O'CAMPO, 2006; VADEN-KIERNAN et al.,

[11] Different variables were used in the construction of such index; some examples include the number of unemployed, the education level of neighborhood residents, the proportion of households headed by single women, the proportion of households who receive social benefits and the proportion of immigrants and/or minorities.

2010; ROOS; WALL-WIELER; LEE, 2019), receptive vocabulary (KOHEN; OLIVER; PIERRE, 2009; VADEN-KIERNAN et al., 2010) or social competence (KERSHAW et al., 2007). There is also evidence of a negative relationship between neighborhood socioeconomic status and different behavioral problems (WOLF; MAGNUSON; KIMBRO, 2017; SHAW et al., 2016; HEBERLE et al., 2014; ODGERS et al., 2012; BARRY et al., 2015; VADEN-KIERNAN et al., 2010; HART; ATKINS; MATSUBA, 2008; WINSLOW; SHAW, 2007).

Another group of five studies analyze the relationship between neighborhood affluence[12] and child development as measured by broad indicators of child development (KERSHAW et al., 2007; CARPIANO; LLOYD; HERTZMAN, 2009) or by cognitive skills (ANDERSON; LEVENTHAL; DUPéRé, 2014; ANDERSON; JOHNSTON; LEVENTHAL, 2019; VADEN-KIERNAN et al., 2010).

There is a smaller group of studies - four in total - that do not find evidence to corroborate the results presented up to this point (KOHEN et al., 2008; HURT; BETANCOURT, 2017; KINGSTON et al., 2013; PALAMAR et al., 2015)[13].

Besides the analysis of a direct relationship between neighborhood socioeconomic status and child development, these studies also provide some other interesting evidence. First, there is some evidence that the longer the child's exposure to neighborhood poverty, the stronger its effects are on child development (ROOS; WALL-WIELER; LEE, 2019). Second, there is evidence that neighborhood poverty is positively associated with an increase in behavioral problems over time (HART; ATKINS; MATSUBA, 2008; PALAMAR et al., 2015; ODGERS et al., 2012)[14].

The social domain This includes characteristics of neighborhood social dynamics, including the existence of support networks between neighbors and the level of social cohesion (GOLDFELD et al., 2015). This dimension also includes the residents' perception of security and violence-related indicators (GOLDFELD et al., 2015). 26 of the selected articles covered topics related to this domain.

Thirteen studies analyze neighborhood social disorder[15] and its relationship to children's development, providing evidence of a significant relationship between these variables. In these studies, children's development is either measured through children's

[12] Measured by indexes that included at least the percent of individuals that were employed in a professional occupation.

[13] A fourth study, Morrissey e Vinopal (2018b) finds significant results when cognitive skills are analyzed, but not when social and behavioral outcomes are analyzed.

[14] This association loses statistical significance when indicators of parental involvement and maternal warmth are controlled for (ODGERS et al., 2012).

[15] Measured in different ways that include the existence of abandoned houses, robberies and drug trafficking or even questions about the safety of the neighborhood for children to play, the presence of garbage , and empty houses. These information is generally collected through survey questions, answered by the child's caregiver.

externalizing behavior (COLEY; LYNCH; KULL, 2015; PEI et al., 2019)[16] or through cognitive skills (KOHEN et al., 2002; FROILAND, 2019; MCWAYNE et al., 2007; BARBARIN et al., 2006; CAUGHY; O'CAMPO, 2006)[17]. The relationship with children's externalizing behavior is still found when neighborhood social disorder is analyzed along with more structural aspects of neighborhood environment[18] and with neighborhood social cohesion (COLDER et al., 2006)[19]. Taken together, the evidence indicates that children who live in neighborhoods with more social disorder have worse indicators of child development.

These studies also provide two other important pieces of evidence. First, there is some indication that neighborhood social disorder moderates the impact of socioeconomic status on children's executive functions: when there is a high level of neighborhood social disorder[20], children's socioeconomic status is positively related to children's executive function (JOHN; TARULLO, 2019). Second, cumulative exposure to neighborhood social disorder might be especially important. One study finds a significant relationship only at age six but not before (COLEY; LYNCH; KULL, 2015).

The existence of a support network is another neighborhood characteristic that has a positive association with children's development. Two studies analyze this question (SHIN et al., 2019; BAYDAR; AKCINAR, 2015). Shin et al. (2019) measures children's development through a broad instrument of child development at age two and Baydar e Akcinar (2015) find similar evidence when they analyze children's prosocial behavior at age three[21].

In addition to analyzing only the existence of such neighborhood social networks, seven studies analyze the relationship between neighborhood collective efficacy[22] and child development. Three studies measure child development through indicators of internaliz-

[16] In this case, the relationship is positive, as children's externalizing behavior is not a positive developmental outcome. It is important to point out that Barbarin et al. (2006) does not find a significant relationship, but, unlike the other studies, it includes cognitive skills in this analysis.

[17] Fantuzzo et al. (2005) do not find a significant relationship between these two variables. This study is quite similar to McWayne et al. (2007) and the different results are striking; they might be a consequence of the fact the McWayne et al. (2007) does not control for family covariates. John e Tarullo (2019) also does not find a significant direct relationship between children's cognitive skills and neighborhood social disorder, but they also consider children's executive function. Finally, analyzing smaller samples of students, all from a vulnerable background, Kiernan et al. (2008) and Caughy e O'Campo (2006) do not find a significant relationship between neighborhood social disorder and children's outcomes.

[18] Like crime and adequate police protection.

[19] In this case, this result is valid for children that, in their infancy, were characterized by certain temperaments (high positive affect/low fear, low positive affect/high fear).

[20] A measure defined by the parents perception of it in the neighborhood and including vandalism, sense of security and perception of activities related to drug trafficking and use.

[21] No such relationship is found for other two outcomes: receptive vocabulary knowledge and externalizing behavior.

[22] Measured through questions regarding the existence of informal social control in the neighborhood, social cohesion and trust among neighbors.

ing and externalizing behaviors and these studies find evidence of a significant, and negative, relationship between the analyzed variables (MA; KLEIN, 2018; MA; GROGAN-KAYLOR, 2017; MA, 2016). Another two studies measures child development through cognitive skills and find evidence of a significant and positive relationship with neighborhood collective efficacy (KOHEN et al., 2002; CAUGHY; O'CAMPO, 2006). A fifth study finds evidence of such a relationship when it analyzes the risk for developmental problems and its relationship with neighborhood collective efficacy (BLAIR; FORD, 2019). The sixth study analyzes not only neighborhood collective efficacy, but also crime and deteriorated structural conditions; no significant relationship is found with children's externalizing behavior in this case (MA; GROGAN-KAYLOR; LEE, 2018)[23].

Two other studies analyze constructs that are similar - but less embracing - than collective efficacy. The first one analyzes neighborhood belongingness and finds no evidence of a significant association with children's internalizing and externalizing behaviors (CALLAHAN et al., 2011). The other analyzes neighborhood trust and finds a positive and significant association with child development (NEVES et al., 2016).

Taken together, these sixteen studies indicate the importance of the relationships between neighbors when one is analyzing the relationship between neighborhood and child development.

A final group of studies analyze the relationship between neighborhood violence and child development. Three studies find a significant and negative relationship, as expected, between violence[24] and child development as measured through different indicators of behavior problems (CALLAHAN et al., 2011; MA; GROGAN-KAYLOR; LEE, 2018; BRIGGS-GOWAN; CARTER; FORD, 2012; VILSAINT et al., 2013) and through children's social competence (BRIGGS-GOWAN; CARTER; FORD, 2012). A fourth study compares violent and non-violent communities in the US, finding that children from non-violent communities had greater behavioral problems than children from violent communities(FEDOR; BENDER; CARLSON, 2010). When safety, and not violence, is analyzed, no significant relationship with child development is found. FFour studies use this analysis and each one uses a different measure of child development: mathematics achievement (BAKER, 2015), reading achievement (AIKENS; BARBARIN, 2008), the risk of developing developmental disorders (BLAIR; FORD, 2019) and internalizing and externalizing behaviors (CALLAHAN et al., 2011). These studies provide mixed evidence to the study of the relationship between child development and neighborhood violence, but they indicate the importance of such studies, and specially, of understanding why results are different when safety, and not violence, is analyzed.

[23] Blair e Ford (2019) find a similar result. When they analyze collective efficacy together with neighborhood safety and physical disorder, only neighborhood physical disorder remained statistically associated with moderate risk for developmental disability.

[24] Measured through the perception of the survey respondent.

The physical domain This includes the neighborhood's housing conditions, both physical[25] and related to ownership[26]. The neighborhood's physical characteristics[27] are also considered in this domain as well as the presence of vandalism. This domain also includes characteristics of the public transport available in the neighborhood[28] (GOLDFELD et al., 2015). The systematic search found nine studies that provide evidence on this domain and its association with children's development.

Five studies analyze general housing conditions of the neighborhood and their relationship with child development. Analyzing Australian data, two studies find relationships between housing conditions and children's development (VILLANUEVA et al., 2019; KERSHAW et al., 2007). The most recent study uses a mixed method approach and finds evidence that more affordable housing, less renters, and a lower density of high-rise, high-density public housing are characteristics of neighborhoods related to better development in young children (VILLANUEVA et al., 2019). Kershaw et al. (2007) find evidence of a negative association between the rate of housing ownership in a neighborhood and children's vulnerability in their communication skills. Another Australian study does not find evidence of a relationship between neighborhood residential density and children's developmental vulnerability (CHRISTIAN et al., 2017). Analyzing the US context, two studies find mixed evidence. McWayne et al. (2017) finds an association between the neighborhood housing types[29] - and children's achievement: five-year-old children living in neighborhoods with a larger proportion of semi-detached or detached homes had higher achievement in math and language, than children who lived in neighborhoods with primarily row homes. Another US study finds no relationship between general housing conditions in the neighborhood[30] and the child development (LOVASI et al., 2011).

Two studies analyze the presence of green areas in the neighborhood and their relationship with children's development, but neither study finds evidence of a significant relationship (FLOURI; MIDOUHAS; JOSHI, 2014; CHRISTIAN et al., 2017). The British study, however, finds evidence that the family's socioeconomic status moderates such relationship: for children from poor families, living in a greener neighborhood is associated with better emotional development than their counterparts who live in neighborhoods with less green area (FLOURI; MIDOUHAS; JOSHI, 2014). This same study also concludes that the use of green areas by children is significantly associated with children's

[25] Including measures of housing quality, of housing density and of the existence of high-rise buildings versus single-dwelling houses. Also includes the presence of condemned properties.

[26] Including the percentage of homes that are occupied by renters and the percentage of public housing.

[27] This includes the presence of green areas: the existence of parks, squares and green areas, as well as measures of their quality, as well as indicators of structural hazards, such as the incidence of domestic fires.

[28] Mainly, its proximity and accessibility.

[29] That is, row homes, detached or semi-detached houses, apartment buildings, government owned properties.

[30] Measured by the proportion of unoccupied houses and the proportion of houses with adequate sanitation.

emotional development and behavior problems (FLOURI; MIDOUHAS; JOSHI, 2014). This might explain the non-significant result found in these two studies that having parks near children's home is not enough; children need to use them frequently to benefit from them.

Another five studies focus on general physical characteristics of the neighborhood and their relationship with child development. Analyzing two samples of kindergarten students in large, urban centers in the US, two studies did not find robust evidence for the relationship between children's cognitive skills and the existence of structural hazards[31] in the neighborhood (MCWAYNE et al., 2007; FANTUZZO et al., 2005)[32]. Another study for the US context found evidence of a significant relationship between neighborhood physical disorder[33] and a moderate risk of developmental disabilities, even when other neighborhood characteristics, safety and isolation, are controlled for (BLAIR; FORD, 2019). A Brazilian study analyzes general neighborhood infrastructure and finds evidence of a significant relationship between the neighborhood infrastructure and the language of children between 24 and 36 months old (NEVES et al., 2016). In the Australian context, the presence of more low-traffic roads is related to lower odds of developmental vulnerability in the social competence domain (CHRISTIAN et al., 2017), although this same study does not find evidence that the number of public transport stops is related to any child development indicators.

Taken together, these nine studies indicate that more research is needed to reach robust conclusions on the relationship between the neighborhood physical domain and children's development. The evidence presented gives some indication that neighborhood housing conditions matter and that it is not sufficient to consider the presence of green areas when studying their relationship with children's development.

The domain of service provision This includes not only the existence of and easy access to public services and facilities, but also their quality (GOLDFELD et al., 2015). Very little evidence was found for this domain as only three studies were identified in our search and all were correlational studies.

A study with a nationally representative sample of three-year-old Turkish children did not find a significant association between the presence of public services and facilities - reported by parents - and children's pro-social and externalizing behavior, or their receptive vocabulary (BAYDAR; AKCINAR, 2015). Another study, analyzing a sample

[31] Defined as a higher density of condemned properties, a high incidence of domestic fires, and a higher proportion of children with higher lead levels.

[32] McWayne et al. (2007) finds evidence of a negative relationship, but the study does not control for family-level covariates. Fantuzzo et al. (2005) employs a more robust estimation procedure and does not find a significant association. This last study also analyzes other types of skills - motor skills, social knowledge, work habits - and also finds no significant relationship.

[33] This indicator considered the existence of sidewalk rubbish, vandalism, and poorly maintained housing.

of Brazilian children between 24 to 36 months old, also did not find a significant association between the presence and quality of public services and facilities - also reported by parents - and children's cognitive development and expressive language (NEVES et al., 2016). Both studies employed controls for family characteristics including parental socioeconomic status and education, children's characteristics, their home environment and neighborhood social support (BAYDAR; AKCINAR, 2015; NEVES et al., 2016). In both cases, the services and facilities were defined broadly and included those that directly served children such as daycare centers as well as others that served the community as a whole such as phamarcies and police stations (BAYDAR; AKCINAR, 2015; NEVES et al., 2016).

The third study uses the Australian Early Development Census to analyze the relationship between the distance to child-relative destinations and the children's development (CHRISTIAN et al., 2017). The study finds a very counterintuitive result: a negative association between children's developmental vulnerability and the distance to daycare centers and family support services. The same study found no relationship between child development and proximity to health posts or preschools, but it finds evidence that an increase in the access to different utilitarian and recreational destinations is associated with an increase in the odds of children's vulnerability in the social competence domain. It is important to point out that unlike Baydar e Akcinar (2015) and Neves et al. (2016), Christian et al. (2017) does not collect the service provision information from parents but rather uses a secondary database; nor do they consider the quality of the services. These differences might explain the difference in results.

Taken together, these studies do not allow us to reach any conclusions as to the relationship between service provision and child development. They do give some indication that it might be important to consider service and facility quality - and not only their distance from the child's home.

1.5 Discussion

The empirical evidence presented in this chapter is very recent. This is a booming area of study and one in which more research is needed. Specifically, there is a clear need for more research focussing on domains beyond the socioeconomic and social as defined by Goldfeld et al. (2015). It is important to understand which factors, in addition to those related to socioeconomic vulnerability and the social networks, present in the neighborhoods, are responsible for neighborhood effects. More research might also point out to other neighborhood characteristics - not identified in the model proposed by Goldfeld et al. (2015) - that affects child development.

The systematization of the empirical literature also sheds light on the lack of

research that was done outside developed countries. Only two studies out of the 52 selected analyze the context of developing countries[34]. Research from more diverse contexts would allow us to understand if the vulnerability of the context as a whole plays a role in which characteristics of the neighborhoods matter more or less.

Finally, the analysis of the selected literature indicated not only a lack of studies, especially for some domains and those outside developed, English-speaking countries, but also that the existing evidence does not allow for a clear systematization of its conclusions. On more than one occasion, evidence points in different directions which in turn does not allow us to clearly define which neighborhood characteristics actually affect child development. More research is needed to solidify the conclusions indicated in this chapter and allow researchers to reach similar conclusions for other characteristics.

This literature has a clear methodological caveat; families can choose where they live and therefore, it is very difficult to isolate the effects of the neighborhood from those of family characteristics. After all, families might move to specific neighborhoods due to factors that the studies neither observe nor manage to control for and which in turn, may be associated with the child's level of development. In other words, unobservable factors can significantly impair the statistical analysis. Much of the work presented here despite being methodologically robust, does not resolve this issue.

Even though this study tried keep the systematization methodology as open as possible to allow for quantitative and qualitative methodologies, success was limited. Only one article had only qualitative methodology and another used a mixed methods approach. This might be a consequence of the databases searched even though their diversity was a criteria in choosing them and it might also be a consequence of the terms employed.

Despite these limitations, the systematization of the literature indicates that there are neighborhood characteristics that matter to child development. Socioeconomic conditions of the neighborhood affect child development as does the existence and characteristics of neighborhood social networks which are two groups of such characteristics. These results are robust to controls and methodologies that try to separate the neighborhood effect from the family effect. No clear conclusions can be reached about the possible effects of the neighborhood's violence level, its physical space or the provision of public services in the community.

This evidence is correlational and no claim of causality should be made from it. However, this evidence allow us to understand which neighborhood characteristics are associated with child development. Moreover, it also leads to the conclusion that children from vulnerable neighborhoods, however vulnerable neighborhoods are defined, tend to have poorer development. Taken together, these conclusions can influence the

[34] The choice to only examine articles in Portuguese and English has a role in this.

formulation and the implementation of public policy as they indicate the importance of focusing not only on children from vulnerable families, but also on children from vulnerable neighborhoods.

The theme of neighborhood effects and child development is further developed in the next chapters of this thesis. But, its starting point is the conclusion of this review: neighborhoods matter for child development; more specifically, neighborhood socioeconomic vulnerability affect children's cognitive development.

1.6 Appendix A

Table 1.1 – Selected articles

Study population (setting, sample size)	Developmental outcome	Neighborhood-level covariates (indicators)	Neighborhood definition	Controls	Results	Citation
AIKENS, N. L.; BARBARIN, O. Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. Journal of Educational Psychology, v. 100, n. 2, p. 235-251, 2008.	Kindergarten students who were part of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (US, n = 17401).	Reading achievement.	Neighborhood Safety; neighborhood problems (garbage or litter in the street, individuals selling or using drugs in the street, burglary or robbery in the area, violent crime in the area, or vacant homes in the area).	Defined by survey respondents.	Socioeconomic family covariates, parent well-being and involvement in school, individual child characteristics (including early child care experiences), home environment. School characteristics and literacy related activities.	Neighborhood safety and problems are not associated with initial reading achievements. But neighborhood covariates are associated with the rate of children's monthly reading growth as children aged.
ANDERSON, S.; JOHNSTON, W.; LEVENTHAL, T. When neighborhoods matter: Developmental timing and youth reading achievement and problem behaviors. Social Science Research, v. 81, p. 1-11, jul. 2019.	6, 9, and 12 year-olds who were part of the Project on Human Development in Chicago Neighborhoods (PHDCN) (US, n=2628)..	Reading achievement, externalizing and internalizing problems.	Neighborhood advantage (the percent of residents with a B.A. and employed in managerial or professional positions); Neighborhood disadvantage (the poverty rate, percent of residents receiving of public assistance, of female-headed families, and the unemployment rate).	Two to three contiguous and relatively homogenous census tracts (about 8,000 residents).	Socioeconomic family covariates, maternal depression, child individual characteristics.	Neighborhood advantage is associated with children's reading achievement and this relationship is stronger the younger the child is. No significant association was found for behavior problems nor when neighborhood disadvantage is analyzed.

ANDERSON, S.; LEVENTHAL, T.; DUPR, V. Exposure to Neighborhood Affluence and Poverty in Childhood and Adolescence and Academic Achievement and Behavior. Applied Developmental Science, v. 18, n. 3, p. 123-138, jul. 2014.	Children who were part of the NICHD Study of Early Child Care and Youth Development (US, n=1364).	Reading and math achievement, externalizing and internalizing problems.	Neighborhood affluence (the percentage of adults with at least a B.A. degree and the percentage of adults in managerial/professional jobs). Neighborhood poverty (the percentage of single mothers, the percentage of households under the U.S. poverty level, and the percentage of adults unemployed).	Census block-group.	Socioeconomic family covariates (including maternal verbal ability), mothers' beliefs about parenting, maternal personality, maternal depression, residential instability and individual child characteristics.	Neighborhood affluence is associated with children's reading and math achievement; the same cannot be said of neighborhood poverty. No significant associations emerged for neighborhood affluence or poverty and children's externalizing behaviors. Internalizing behaviors are associated with neighborhood affluence but not neighborhood poverty.
BAKER, C. E. Does Parent Involvement and Neighborhood Quality Matter for African American Boys Kindergarten Mathematics Achievement? Early Education and Development, v. 26, n. 3, p. 342-355, abr. 2015.	5 year old African American boys who were part of Early Childhood Longitudinal Study (ECLS-K) Kindergarten Class of 1998-1999 (US, n=1202).	Mathematics achievement.	Neighborhood safety.	Defined by survey respondents.	Socioeconomic family covariates, child's age and home environment.	Neighborhood safety is not associated with initial mathematics achievement.
BARBARIN, O. et al. Children Enrolled in Public Pre-K: The Relation of Family Life, Neighborhood Quality, and Socioeconomic Resources to Early Competence. American Journal of Orthopsychiatry, v. 76, n. 2, p. 265-276, 2006.	4 year-old children who were part of the National Center for Early Development and Learnings (NCEDL) Multi-State Study of Pre-Kindergarten (US, n = 501)	Pre-academic competences (receptive language, letter identificiation, mathematics) and behavior problems.	Neighborhood quality (physical safety, litter, crime, abandoned buildings, and drugs).	Defined by survey respondents.	Socioeconomic family covariates, parent and child's health, partners relationship, child's individual characteristics.	Neighborhood quality is associated with children's language development, but not with the other outcomes.

BARRY, S. J. E. et al. Mapping area variability in social and behavioural difficulties among Glasgow preschoolers: linkage of a survey of pre-school staff with routine monitoring data. Child Care, Health and Development, v. 41, n. 6, p. 853-864, 2015. ISSN 03051862.	Pre-school students in Glasgow City between 2010-2012 (Scotland, n=10409).	Behavior problems.	Neighborhood deprivation status (income, employment, health, education, housing, geographic access to services and crime data).	Datazones (key small-area statistical geography in Scotland).	Child individual characteristics.	Neighborhood deprivation is associated with behavior problems.
BAYDAR, N.; AKCINAR, B. Ramifications of socioeconomic differences for three year old children and their families in Turkey. Early Childhood Research Quarterly, v. 33, p. 33-48, 2015.	3 year-old children from a nationally representative sample (Turkey, n=902)	Receptive vocabulary knowledge; externalizing behavior; prosocial behavior.	- Availability and adequatability of playgrounds, sports fields, health centers, schools, preschools, community education centers, job training centers, and police stations. '- Social support from neighbors.'	Defined by survey respondents, regardless of the size of that area or its official administrative status.	Socioeconomic family covariates (maternal education and economic well-being), individual child characteristics, parenting behaviors and home environment.	Physical resources were not related to any of the outcomes analyzed. Social support was significantly related to prosocial behavior, but not to the other outcomes.
BENSON, J.; BORMAN, G. D. Family, Neighborhood, and School Settings Across Seasons: When Do Socioeconomic Context and Racial Composition Matter for the Reading Achievement Growth of Young Children? Teachers College Record, v. 112, n. 5, p. 1338-1390, 2010.	Children who were part of the Early Childhood Longitudinal Study- Kindergarten Cohort (US, n=4180).	Reading achievement.	Neighborhood social context (educational level and income)	Census zip code tabulation areas	Socioeconomic family covariates, individual child characteristics, school covariates.	Neighborhood social context is associated with children's reading achievement levels at school entry and for their reading achievement growth during the summer.

BLAIR, L. M.; FORD, J. L. Neighborhood Context and the Risk for Developmental Disabilities in Early Childhood. Maternal and Child Health Journal, v. 23, n. 9, p. 1213-1219, set. 2019.	0-5 year-old children who were part of the National Survey of Children's Health (NSCH) 2011/2012 (US, n = 29.997)	Risk for developmental problems	Neighborhood safety; neighborhood physical disorder; neighborhood isolation	Defined by survey respondents.	Socioeconomic family covariates, parental mental and physical health and individual child characteristics.	All three neighborhood-level covariates were statistically associated with moderate (but not severe) risk for developmental disability. When all three covariates were analyzed together, only neighborhood physical disorder remained statistically associated with moderate risk for developmental disability.
BRIGGS-GOWAN, M. J.; CARTER, A. S.; FORD, J. D. Parsing the Effects Violence Exposure in Early Childhood: Modeling Developmental Pathways. Journal of Pediatric Psychology, v. 37, n. 1, p. 11-22, jan. 2012.	5-6 year-old children who were part of subsample of a longitudinal representative birth cohort (US, n=437)	Externalizing and internalizing behavior; social competence.	Children's exposure to neighborhood violence	Defined by survey respondents.	Sociodemographic risk, children's lifetime violence exposure, individual child's characteristics.	Neighborhood violence is associated with worse behavior and lower social competence in early elementary school.
BROWNELL, M. D. et al. A population-based analysis of factors that predict early language and cognitive development. Early Childhood Research Quarterly, v. 35, p. 6-18, 2016.	Children who took a grade 3 reading assessment in the 2009/10 school year (Canada, n = 8983)	Language and cognitive development	Neighborhood socioeconomic status (unemployment, household income, and high school completion)	Dissemination area level, which includes about 400-700 people.	Maternal health behaviors during the prenatal period, child's health at birth, family risk.	Neighborhood socioeconomic status is negatively associated with children's language and cognitive development in kindergarten.

CALLAHAN, K. L. et al. Neighborhood disadvantage as a moderator of the association between harsh parenting and toddler-aged childrenÕs internalizing and externalizing problems. Journal of Family Psychology, v. 25, n. 1, p. 68-76, 2011.	2 year-old children who were enrolled in Head Start (US, n= 55)	Externalizing and internalizing behavior.	Neighborhood danger; neighborhood belongingness.	Defined by survey respondents.	Harsh parenting,	Neighborhood danger is significantly associated with higher levels of behavior problems.
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, v. 69, n. 3, p. 420-432, ago. 2009.	Kindergarten students who were part of the Human Early Learning Partnership (HELP) (Canada, n=37798)	Child development.	Index of Concentration at the Extremes (ICE); educational heterogeneity (% Non-official language as a mother tongue); cultural heterogeneity (% Reporting Aboriginal identity); and residential instability (% Moved in last year)	Defined by stakeholders in each school district.	Individual child characteristics, median income of the child's residential postal code.	The ICE is positively associated with child development. Educational heterogeneity is negatively associated with child development . No other associations are found.
CAUGHY, M. O.; OÕCAMPO, P. J. Neighborhood Poverty, Social Capital, and the Cognitive Development of African American Preschoolers. American Journal of Community Psychology, v. 37, n. 1-2, p. 141, 2006.	3 and 4 year-old children and families with children whose primary caregiver self-identified as African American (US, n=200)	Cognitive competence	Neighborhood structure (economic impoverishment and population instability); Neighborhood social capital; Neighborhood physical and social disorder	Census block group	Family demographic characteristics, positive parent involvement, cultural context of the home.	Children's problem-solving skills varied by neighborhood impoverishment levels. Neighborhood physical and social disorder is not asssociated with problem-solving skills. And neighborhood social capital is associated with such skills.

CHRISTIAN, H. et al. Relationship between the neighbourhood built environment and early child development. Health & Place, v. 48, p. 90-101, nov. 2017.	5 year old children in Perth, who were part of the Australian Early Development Census (Australia, n = 23.395 children)	Developmental maturity in five domains: Physical Health and Wellbeing, Social Competence, Emotional Maturity, Language and Cognitive Skills, and Communication Skills and General Knowledge.	Walkability(Street connectivity, Land use mix, Residential density, Low traffic exposure, Public transport stops), Green Spaces (Distance to nearest park, Distance to nearest attractive park, Distance to nearest pocket park, Distance to nearest nature/conversation area, Distance to nearest school grounds). Distance to Child-Relevant Destinations (kindergarten, child-center-based-care, family support service, child health clinic, playgroup venue).	Local Communities, localities with an average population size of 10,000 persons	Socioeconomic family covariates and child individual characteristics.	The odds of developmental vulnerability was negatively associated with each 1 km increase in the distance to the nearest child-centre-based-care (Social Competence domain) and family support service (Social Competence domain). The odds of developmental vulnerability was negatively associated with the increase in proportion of low traffic roads in the neighborhood (Social Competence domain) and with the increase in the distance to the nearest school grounds (Physical Health and Wellbeing and Emotional Maturity domains) and attractive park (Social Competence domain). An increase in land use mix was associated with an increase in the odds of developmental vulnerability (Social Competence domain). No relationship was found for developmental vulnerability and distance to the nearest kindergarten, child health clinic or play group, pocket park, nature/conservation area, number of public transport stops or residential density.

COLDER, C. R. et al. Temperament in context: Infant temperament moderates the relationship between perceived neighborhood quality and behavior problems. Journal of Applied Developmental Psychology, v. 27, n. 5, p. 456-467, set. 2006.	Children who were born to mothers who were part of National Longitudinal Survey of Youth (NLSY) and who were part of the 1986 NLSY (US, n = 316).	Externalizing and internalizing behavior (at age 6 and 12).	Neighborhood quality (both structural - crime, unemployment, abandoned buildings, adequate police protection - and related to social cohesion - residence respect rules, supervise children, care about what goes on). Ê	Defined by survey respondents.	Child temperament before age of 1, child's sex, family income.	Low neighborhood quality was positively associated with externalizing behavior for children at age 6, who were, in infancy, caracterized by certain temperaments (high positive affect/low fear, low positive affec/high fear). This association was also present when the change in externalizing behavior between ages 6 and 12 was analyzed. For internalizing behavior, a positive association was only found when the change in such behavior was analyzed
COLEY, R. L.; LYNCH, A. D.; KULL, M. Early exposure to environmental chaos and childrenÕs physical and mental health. Early Childhood Research Quarterly, v. 32, p. 94-104, 2015.	Children who were part of the Three-City Study and had less than 2 years old at the first wave of data collection, in 1999 (US, n = 495).	Developmental delays, externalizing and internalizing behavior (at age 6).	Neighborhood disorder (abandoned houses and burglaries, assaults, and drug dealing)	Defined by survey respondents.	Child's health and individual characteristics, socioeconomic family covariates, housing disorder, family instability, parental functioning,	Neighborhood disorder is associated with behavior problems for children at age 6. Such association is not found for developmental delays, measured when the children were younger than 2.
COULTON, C. J. et al. Temporal effects of distressed housing on early childhood risk factors and kindergarten readiness. Children and Youth Services Review, v. 68, p. 59-72, set. 2016.	Kindergarten students in the Cleveland Metropolitan School District during the 2007-2010 academic years (US, n= 13.762).	Literacy readiness.	Neighborhood disadvantage (welfare receipt, poverty, unemployment, female-headed households, percentage African American, and density of children).	Census tract.	Socioeconomic family covariates, child individual characteristics (including low birth weight), family housing conditions, occurance of child maltreatment investigations, residential instability, elevated blood lead level.	Neighborhood disadvantage has a negative association with literacy readiness.

DA ROCHA NEVES, K. et al. Growth and development and their environmental and biological determinants. Jornal de Pediatria, v. 92, n. 3, p. 241-250, maio 2016.	24-36 month-old children with no congenital or acquired disabilities who were enrolled in the municipal early childhood education network (Brasil, n = 92)	Child development in two domains: cognitive and expressive language.	Accessibility and quality of services, both public and private; perceived neighborhood security, social turmoil, and interaction and trust among neighbors; neighborhood infrastructure	Defined by survey respondents.	Socioeconomic family covariates (parental education and socioeconomic status), individual child characteristics and home environment.	Neighborhood infrastructure and interaction and trust between neighbors were positively related to children's expressive language development. No other significant relationships were found.
FANTUZZO, J. W. et al. Early childhood experiences and kindergarten success: A population-based study of a large urban setting. School Psychology Review, v. 34, n. 4, p. 571-588, 2005.	Representative sample of kindergarten students in an urban school district (n=3.969).	Language art skills, mathematics skills, motor skills, social knowledge, work habits.	Neighborhood social problems (rates of truancy, child poverty, teen births, delinquent and dependent out of home placements, and substantiated abuse and neglect cases); neighborhood structural danger (density of dangerous properties, incidences of residence fires, and lead levels).	Census block-group	Socioeconomic family covariates, individual child characteristics (including early care experiences).	No relationship was found between neither neighborhood social problems nor neighborhood structural danger with any of the children's skills.
FEDOR, M. C.; BENDER, S. L.; CARLSON, J. S. Examining Risk and Protective Factors in Head Start Populations Located in High- and Low-Violence Communities: Infants & Young Children, v. 23, n. 3, p. 209-217, jul. 2010.	Children enrolled in 5 Head Start programs in high-violence communities and in 5 Head Start programs in low-violence communities (US, n = 388).	Externalizing and internalizing behavior, self-control and initiative.	Community violence (number of violent crimes reported)	Defined as the neighborhood were the Head Start center the child is enrolled in is located.	None.	Children who lived in less violent neighborhoods had more behavior problems than those who lived in high-violence ones.

FLOURI, E.; MIDOUHAS, E.; JOSHI, H. The role of urban neighbourhood green space in children's emotional and behavioural resilience. Journal of Environmental Psychology, v. 40, p. 179-186, dez. 2014.	3-7 year-old children who lived in urban English neighborhoods and were part of the Millennium Cohort Study (England, n= 6348).	Emotional and behavioural problems	Neighborhood green space (% of space within LSOA that was green); Neighborhood disadvantage (LSOA's 2004 Index of Multiple Deprivation)	Lower layer Super Output Areas (LSOAs), groups of Census Output Areas, that have, on average, 1500 residents.	Socioeconomic family covariates, individual child characteristics, maternal and child general health, maternal and child physical activity and maternal psychological distress	Neighborhood green space was unrelated to children's behavioral problems. A significant effect was found for the interaction between green space and family socioeconomic status when emotional problems were analyzed: for poor children, more green space in their neighborhood was associated with fewer emotional problems from age 3 to 5, relative to their counterparts in less green neighbourhood
FROILAND, J. M. Examining the Effects of Location, Neighborhood Social Organization, and Home Literacy on Early Cognitive Skills in the United States. International Journal of Psychology: A Biopsychosocial Approach, v. 9, p.29-42, 2011.	Kindergarten students who were part of the Early Childhood Longitudinal Study, Kindergarten Class of 1998 (US, n = 21409).	Cognitive skills (a composite of individually administered early reading, math and general knowledge scales),	Neighborhood social organization (how safe it is for children to play; presence of litter; sale of drugs; violent crime; burglary; and vacant homes).	Defined by survey respondents.	Socioeconomic family covariates, access to cultural resources, home literacy, individual child characteristics and type of community (small town/rural and suburban/large town)	Neighborhood social organization is positively associated with children's cognitive skills.

HANSON, M. J. et al. Neighborhood Community Risk Influences on Preschool Children's Development and School Readiness. Infants & Young Children, v. 24, n. 1, p. 87-100, 2011.	Children enrolled in a preschool program that either lived in poverty, had identified disabilities, or whose families spoke a primary home language other than English (US, n=1006)	Receptive vocabulary, early literacy, mathematics achievement, social skills.	Neighborhood economic hardship (family poverty, female-only head of household, male educational attainment, male unemployment); neighborhood composition (the percent of English speaking households in the community).	Census block	Maternal education and individual child characteristics.	Neighborhood community hardship is negatively associated with both children's math achievement and early literacy. No such relationship was found for receptive vocabulary, nor for social skills. Neighborhood composition is not associated with children's academic outcomes, but it is to social participation (a dimension of social skills).
HART, D.; ATKINS, R.; MATSUBA, M. K. The association of neighborhood poverty with personality change in childhood. Journal of Personality and Social Psychology, v. 94, n. 6, p. 1048-1061, 2008.	Children who were part of the 1979 National Longitudinal Survey of Youth (US, two sub-samples: n= 1692, n=1550)	Behavior problems, personality (overcontrol, resiliency and undercontrol).	Neighborhood poverty (the percentage of all individuals in the census tract living in households with incomes below the federal poverty level), Neighborhood distrust.	Census tract.	Socioeconomic family covariates, home environment, individual child characteristics (including Head Start participation), maternal depression.	Children's personality change is negatively associated with neighborhood disadvantage. This association is also found for behavior problems.
HEBERLE, A. E. et al. The Impact of Neighborhood, Family, and Individual Risk Factors on Toddlers' Disruptive Behavior. Child Development, v. 85, n. 5, abr. 2014.	12-42 months old children who were born in the State of Connecticut between July 1995 and September 1997 (US, n=1204)	Externalizing behavior.	Neighborhood Social Disadvantage (men not attached to the labor force, high school drop-outs, families, with children, headed by women, and households dependent on public assistance).	Census tract.	Socioeconomic family covariates, violence or conflict exposure, parent depressive symptoms, parenting behavior, individual child characteristics	Children's externalizing behavior is positively associated with neighborhood social disadvantage.

HURT, H.; BETANCOURT, L. M. Turning 1 Year of Age in a Low Socioeconomic Environment: A Portrait of Disadvantage. Behavioral Pediatrics, v. 38, n. 7, p. 8, 2017.	Infants whose mothers American-born African-American (US, n=60).	Cognitive, language, and motor development; language skills.	Neighborhood Disadvantage (percentage of individuals below the poverty line, unemployed, receiving public assistance, African-Americans, children younger than 18 years, and female-headed households).	Census tract.	Socioeconomic family covariates, maternal cognitive function, parental stress, maternal depression, home environment, individual child characteristics	
JOHN, A. M.; TARULLO, A. R. Neighbourhood chaos moderates the association of socioeconomic status and child executive functioning. Infant and Child Development, 2019.	4.5-5.5 year-old children(US, n = 121 children).	Executive function (working memory accuracy, accuracy in go/no-go trials), receptive language, nonverbal IQ.	Neighborhood quality (percentage of adults >25 years old with less than a high school education; percent- age of unemployed males; percentage of households with an income below the poverty line; percentage of households receiving public assistance; percentage of households with children that are headed by a female; and median household income); neighborhood chaos (vandalism, feeling safe, and drug activity).	Census tract and defined by survey respondents.	Socioeconomic family covariates (including parent's education and ocupation), parenting stress, parents' depression, household chaos and individual child characteristics.	Neighborhood chaos is not directly associated with children's executive function, but it is a moderator of the relationship between family's socioeconomic status and children's executive function. Neither relationship is found for neighborhood quality and children's executive function.

KERSHAW, P. et al. Toward a Social Care Program of Research: A Population-Level Study of Neighborhood Effects on Child Development. Early Education & Development, v. 18, n. 3, p. 535-560, out. 2007.	Kindergarten students (Canada).	Five dimensions of school readiness: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge	Neighborhood structural characteristics (residents' mode of transportation to paid work, and characteristics of local dwellings). Residential mobility. Neighborhood demographic and socioeconomic characteristics (including % that were of a certain ethnicity). Neighborhood occupational characteristics (including educational level of the neighbors).	Defined by stakeholders in each school district so that each is home to between 35 and 200 kindergarten children.	Individual child characteristics.	The neighborhood median family income was negatively correlated with children's social competence. The rate of home ownership rate in the neighborhood was also negatively correlated with children's communication skills.
KIERNAN, G. et al. the school readiness of children living in a disadvantaged area in Ireland. Journal of Early Childhood Research, v. 6, n. 2, p. 119-144, jun. 2008.	Children in their first year of formal schooling, attending four schools designated as disadvantaged and located in a socio-economically deprived community (Ireland, n = 89).	Cognitive abilities, emotional well-being and social behaviour.	Neighborhood quality (crime, antisocial behaviour and general environment).	Defined by survey respondents.	Individual child characteristics (including early care experiences) and health, home context (including parenting behavior, parent's mental and physical health, parent's relationship status), community support.	No association was found between neighborhood quality and the children's outcomes.

KINGSTON, S. et al. Parent Involvement in Education as a Moderator of Family and Neighborhood Socioeconomic Context on School Readiness Among Young Children. Journal of Community Psychology, v. 41, n. 3, p. 265-276, 2013.	Pre-k students at the eight study schools during the study period (US, n=554).	Cognitive skills (motor, language and conceptual skills), prosocial and externalizing behavior.	Concentrated affluence (the percentage of households with annual incomes above $75,000, the percent of individuals that were employed in a professional occupation, childcare burden).	Census tract.	Socioeconomic family covariates, parent involment in school, individual child characteristics	Neighborhood childcare burden is associated with lower prosocial skills. No other significant direct associations were found. The percent of households with annual income above $75,000 was positively associated with prosocial skills when interacted with parent involvement in education. The same association was found for childcare burden.

KOHEN, D. E. et al. Neighborhood income and physical and social disorder in Canada: Associations with young childrenÕs competencies. Child Development, v. 73, n. 6, p. 1844-1860, 2002.	4- and 5-year-old children who were part of the National Longitudinal Survey of Children and Youth of 1994 (Canada, n = 3497).	Receptive vocabulary, Externalizing and internalizing behavior.	Neighborhood socioeconomic characteristics (income and family structure), neighborhood physical and social disorder (traffic, garbage, litter, or broken glass, lottering, threatening behavior, drunken or otherwise intoxicated persons, general condition of buildings), neighborhood social cohesion (if there is a problem, neighbors get together to deal with it; there are adults in the neighborhood that children can look up to; people are willing to help their neighbors; you can count on adults in the neighborhood to watch that children are safe and out of trouble; and when I am away, I know that my neighbors will keep their eyes open for possible trouble).	Enumeration area (census units of approximately 300 households).	Socioeconomic family covariates, maternal emotional and social characteristics, and individual child characteristics.	Neighborhood affluence is positively associated with children's receptive vocabulary and negatively associated with behavior problems; neighborhood poverty is negatively associated with children's receptive vocabulary. Neighborhood social cohesion is negatively associated with children's behavior problems, but no association is found for children's receptive vocabulary. Neighborhood physical and social disorder is negatively associated with children's receptive vocabulary, but no association is found for children's behavior problems.

KOHEN, D. E., LEVENTHAL, T., DAHINTEN, V. S., MCINTOSH, C. N. Neighborhood Disadvantage: Pathways of Effects for Young Children. Child Development, v. 79, n. 1, p 156-169, 2008.	4- and 5-year-old children who were part of the National Longitudinal Survey of Children and Youth of 1994 in its Cycle 3 (Canada, n = 3528).	Receptive vocabulary, externalizing and internalizing problems.	Neighborhood disadvantage (Neighborhood income; Neighborhood education; Neighborhood family structure; and Neighborhood unemployment). Neighborhood cohesion.	Enumeration area (census units of approximately 300 households).	Socioeconomic family covariates, maternal depression, home literacy, parenting characteristics, family functioning, home literacy, and individual child characteristics.	There is no direct effect of neighborhood disadvantage on either receptive vocabulary or behavior problems; it does have a significant indirect effect on behavior problems through both neighborhood cohesion and family-level covariates.
KOHEN, D., OLIVER, L., PIERRE, F. Examining the effects of schools and neighbourhoods on the outcomes of Kindergarten children in Canada. International Journal of Speech-Language Pathology, v. 11, n. 5, p. 404-418, 2009.	The sample consisted of a total of 2743 children attending 181 Kindergarten schools in seven Canadian cities (South Eastman, MB; Hampton, NB; Abottsford, BC; Mississauga, ON; Niagara, ON; Saskatoon, SK; and Montreal, PQ) and living in 272 neighbourhoods (Canada, n=2743)	Receptive vocabulary, early literacy and numeracy skills, inattention to task, behavior problems (Hyperactivity-Inattention, Emotional Disorder-Anxiety, Physical Aggression-Conduct Disorder), child development (Social Knowledge and Competence, Emotional Maturity, and Language and Cognitive Development)	Proportion of low income residents in the neighborhood, Youth Unemployment Rate in the neighborhood, proportion of residents with less than high school graduation in the neighborhood, proportion of Aboriginal Peoples in the neighborhood, proportion of recent immigrants in the neighborhood, residential mobility.	Census tract.	Socioeconomic family covariates, individual child characteristics, school covariates.	The proportion of low income residents is positively associated with children's behavior problems (Hyperactivity-Inattention, Physical Aggression-Conduct Disorder). The YUR is also positively associated with children's behavior problem (Hyperactivity-Inattention). The proportion of residents with less than high school graduation was negatively associated with children's receptive language. The proportion of recent immigrants is negatively associated with children's receptive language and positively, with early literacy and numeracy skills and behavior problems (Emotional Disorder-Anxiety, Physical Aggression-Conduct Disorder).

LOVASI, G. S. et al. Chlorpyrifos Exposure and Urban Residential Environment Characteristics as Determinants of Early Childhood Neurodevelopment. American Journal of Public Health, v. 101, n. 1, p. 63-70, jan. 2011.	3 year-old children who were part of a birth cohort established by the Columbia Center for Children's Environmental Health. (US, n=327).	Cognitive and psychomotor development	Neighborhood socioeconomic deprivation, composition, linguistic isolation, crowding, and physical infrastructure	Defined as a 1-kilometer network buffer from the child's home.	Socioeconomic family covariates (including maternal intelligence), individual child characteristics , home environment and chlorpyrifos exposure.	Neighborhood socioeconomic deprivation is significantly associated with children's development - both cognitive and psychomotor. The associations for the other neighborhood characteristics was not robust.
MA, J. Neighborhood and parenting both matter: The role of neighborhood collective efficacy and maternal spanking in early behavior problems. Children and Youth Services Review, v. 70, p. 250-260, nov. 2016.	5 year-old children who were part of the Fragile Families and Child Wellbeing Study (FFCWS) (US, n= 2472)	Externalizing and internalizing behavior.	Neighborhood collective efficacy (informal social control, social cohesion and trust); Neighborhood income.	Census tracts and defined by survey respondent.	Socioeconomic family covariates, maternal spanking, maternal warmth, maternal depression, and individual child characteristics.	Neighborhood collective efficacy is negatively associated with children's behavior problems. No association is found for neighborhood income and children's behavior problems.
MA, J.; GROGAN-KAYLOR, A. Longitudinal associations of neighborhood collective efficacy and maternal corporal punishment with behavior problems in early childhood. Developmental Psychology, v. 53, n. 6, p. 1027-1041, jun. 2017.	Children who were part of the Fragile Families and Child Wellbeing Study (FFCWS) waves 2, 3 and 4 (US, n=3705)	Externalizing and internalizing behavior.	Neighborhood collective efficacy (informal social control, social cohesion and trust); Neighborhood income.	Census tracts and defined by survey respondent.	Socioeconomic family covariates, maternal spanking, maternal warmth, maternal depression, individual characteristics child, child's temperament (at age 1).	Neighborhood collective efficacy is negatively associated with children's behavior problems. This association is stronger for younger children. No association is found for neighborhood income and children's behavior problems.

MA, J.; GROGAN-KAYLOR, A.; LEE, S. J. Associations of neighborhood disorganization and maternal spanking with childrenÕs aggression: A fixed-effects regression analysis. Child Abuse & Neglect, v. 76, p. 106Ð116, fev. 2018. ISSN 01452134.	Children who were part of the Fragile Families and Child Wellbeing Study (FFCWS) waves 3 and 4 (US, n=2472)	Externalizing behavior.	Neighborhood disorganization (neighborhood collective efficacy, crime and violence, and deteriorated structural conditions in neighborhood); neighborhood income.	Census tracts and defined by survey respondent.	Socio-economic family covariates, maternal spanking, maternal warmth, maternal depression, and child individual characteristics.	Only crime and violence in the neighborhood is associated with children's externalizing behavior. No such association is found for neighborhood collective efficacy nor deteriorared structural conditions nor neighborhood income.
MA, J.; KLEIN, S. Does Race/Ethnicity Moderate the Associations between Neighborhood and Parenting Processes on Early Behavior Problems? Journal of Child and Family Studies, v. 27, n. 11, p. 3717-3729, nov. 2018.	Children who were part of the Fragile Families and Child Wellbeing Study (FFCWS) waves 3 and 4 and whose caregiver was either non-Hispanic White, non-Hispanic Black, or Hispanic (US, n=2388)	Externalizing and internalizing behavior.	Neighborhood collective efficacy (informal social control, social cohesion and trust); Neighborhood income.	Census tracts and defined by survey respondent.	Socioeconomic family covariates, maternal spanking, maternal warmth, maternal depression, and individual child characteristics.	Neighborhood collective efficacy is negatively associated with children's behavior problems. This association is found when the neighborhood collective efficacy is interacted with the respondents' race only for Hispanics. No association is found for neighborhood income and children's behavior problems.

MCWAYNE, C. M. et al. Employing community data to investigate social and structural dimensions of urban neighborhoods: An early childhood education example. American Journal of Community Psychology, v. 39, n. 1-2, p. 47-60, 2007.	Public school kindergarten students in a large urban school district (US, n=5.026).	Language and Mathematics achievement.	Neighborhood social stress dimension (rates of truancy, child poverty, teen births, delinquent and dependent out of home placements, and substantiated abuse and neglect cases); neighborhood structural danger (dangerous property, fires on property, high lead). Neighborhood racial composition.	Census block group	Individual child characteristics.	Neighborhood structural danger is associated with the child's language arts achievement. Neighborhood social stress is associated with mathematics achievement. Neighbohood racial composition did not have any significant associations with children's achievement in the final model.
MILBRATH, C.; GUHN, M. Neighbourhood culture and immigrant children's developmental outcomes at kindergarten. Early Childhood Research Quarterly, v. 48, p. 198-214, 2019.	Children born between 1995 and 2005 who attended selected selected school districts (Canada, n=45290)	Child development.	Neighborhood cultural composition (ethnic identity, home language, mother tongue); neighborhood socioeconomic status (median family income, percentage of income from government transfers, unemployment rate, and per- centage of residence with less than a high school education).	Defined by researchers in consultation with local community members .	Family cultural background, immigrant status, family socioeconomic status, individual child characteristics.	Neighborhood poverty is negatively associated with child development. No significant association was found for neighborhood cultural composition and integral child development. There is a significant, negative, association between neighborhood cultural composition and the communication skills.

MORRISSEY, T. W.; VINOPAL, K. M. Neighborhood Poverty and Children's Academic Skills and Behavior in Early Elementary School. Journal of Marriage and Family, v. 80, n. 1, p. 182-197, 2018.	Children who were part of the 2010-2011 ECLS-K (US, n=13350-16600)	Children's reading and math scores, self-control, approaches to learning, attention and focus, inhibitory control, interpersonal skills, externalizing behaviors, and internalizing behaviors.	Neighborhood poverty (average value of the percent of residents living below the federal poverty threshold from 2008 to 2012).	Census tract.	Soci-economic family covariates, maternal depression, individual child characteristics (including early care experiences), disciplinary strategies and spanking, and the availability of books in the home.	Neighborhood poverty is negatively associated with children's achievement between Kindergarten and Second Grade. No significant relationship is found for social-emotional and behavioral outcomes,
ODGERS, C. L. et al. Supportive parenting mediates neighborhood socioeconomic disparities in children's antisocial behavior from ages 5 to 12. Development and Psychopathology, v. 24, n. 3, p. 705-721, ago. 2012.	Children who were part of the Environmental Risk (E-Risk) Longitudinal Twin Study (UK, n = 2232).	Antisocial behavior	Neighborhood socioeconomic status (Classification of Residential Neighborhoods index was built using over 400 variables from the 2001 census).	Enumeration district level (around 150 households)	Socioeconomic family covariates, parent's history of antisocial behavior, family history of mental health problems, physical child maltreatment, domestic violence, maternal warmth, parental monitoring, child's individual characteristics.	Neighborhood SES and children's antisocial behavior are negatively associated at age 5, 7, 10 and 12. The association became stronger as children aged.
PEI, F. et al. The influences of neighborhood disorder on early childhood externalizing problems: The roles of parental stress and child physical maltreatment. Journal of Community Psychology, v. 47, n. 5, p. 1105-1117, 2019.	3 year-old children who are part of the Fragile Families and Child Wellbeing Study (FFCWS) (US, n=3036).	Externalizing behaviors.	Neighborhood disorder (perceptions of disorganized neighborhood environment with descriptions such as misbehaving groups of young children in the neighborhood, unemployed adults loitering in the neighborhood, and drug dealers or users hanging around in the neighborhood.")	Defined by survey respondents.	Parental stress, physical child maltreatment, maternal depression, socioeconomic family covariates and individual child characteristics, including child's temperament (at age 1).	Children's externalizing behavior is positively associated with neighborhood disorder.

ROOS, L. L.; WALL-WIELER, E.; LEE, J. B. Poverty and Early Childhood Outcomes. Pediatrics, v. 143, n. 6 , jun. 2019.	Kindergarten students between 2005 and 2015 (Canada, n= 46 589)	School readiness in five domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge.	Neighborhood poverty (living in a neighborhood with median income in the lowest quintile)	Census dissemination areas (around 400 individuals).	Socioeconomic family covariates, maternal characteristics during pregnancy, social isolation and lone parent status, individual child characteristics (including if low birth weight).	Neighborhood poverty is negatively associated with school readiness.
SHAW, D. S. et al. Transactional effects among maternal depression, neighborhood deprivation, and child conduct problems from early childhood through adolescence: A tale of two low-income samples. Development and Psychopathology, v. 28, n. 3, p. 819-836, 2016.	Study 1: Boys from vulnerable backgrounds who were part of the Pitt Mother & Child Project (US, n = 310); Study 2: Children from vulnerable backgrounds who were part of the Early Steps Multisite Study (US, n = 560);	Behavior problems.	Neighborhood disadvantage factor (median family income, % families below poverty, % households on public assistance, % unemployed, % single-mother households, and % bachelor degree or higher).	Census block group	Socioeconomic family covariates, individual child characteristics, maternal depression, parent behavior (study 1), parent-child coercion (study 2).	Neighborhood disadvantage and behavior problems are positively associated beggining at age 3.5 in study 1 and 5 in study 2.
SHIN, E. K. et al. Association of Maternal Social Relationships With Cognitive Development in Early Childhood. JAMA Network Open, v. 2, n. 1, p. e186963, jan. 2019.	2 year-old children whose mothers were recruited for the project Conditions Affecting Neurocognitive Development and Learning in Early Childhood (US, n= 1082)	Child development.	Neighborhood embeddedness (amount of people mothers knew in the neighborhood).	Defined by survey respondents.	Socioeconomic family covariates (including father's educational level and cohabitation), mother's social network, maternal inteligence and individual child characteristics (includin birth weight).	Children's child development is not associated with neighborhood embeddedness.

VADEN-KIERNAN, M. et al. Neighborhoods as a Developmental Context: A Multilevel Analysis of Neighborhood Effects on Head Start Families and Children. American Journal of Community Psychology, v. 45, n. 1, p. 49-67, 2010.	Children who attended Head Start and were part of the Family and Child Experiences Survey (FACES) 1997 or 2000 (US, n= 6225)	Receptive vocabulary, early literacy and math skills, externalizing and internalizing behaviors.	Neighborhood Low Socioeconomic Status (poor, largely African-American, female-headed families, with high levels of unemployment, particularly for males); Neighborhood with high proportion of residents with English as a Second Language (ESL); Neighborhood High Socioeconomic Status (families with higher levels of income, educational attainment, and professional-managerial occupational status); Neighborhoods with Mobile, Young, and Diverse population (ethnic diversity, recent immigrant population that is mobile and made up of young adults); Neighborhood Family Density (high numbers of families with children and families with multiple children, particularly young children often living in more crowded housing and among fewer older adult).	Census tract.	Socioeconomic family covariates, child's individual characteristics, residential mobility, social risk measure, parents' locus of control, parents' social support, family and child exposure to violence and criminality, family involvement in child's education,	Neighborhood Low Socioeconomic Status is negatively associated with children's receptive vocabulary, early math skills and behavior problems (total, aggressive and withdrawn). Neighborhood with high proportion of ESL residents is negatively associated with behavior problems (total, aggressive and withdrawn) and positively associated with early writing skills. Neighborhood High Socioeconomic Status is negatively associated with behavior problems (total, aggressive and hyperactive) and positively associated with early literacy skills. Neighborhood's family density is negatively associated with children's receptive vocabulary, and early math skills. No other direct associations were significant. When the gains in different outcomes are analyzed, some results change; neighborhood low socioeconomic status is no longer associated with gains in early math skills and the associations for both neighborhood High SES and neighborhood proportion of ESL residents with cognitive skills disappear. For behavior outcomes, all associations became insignificant when gains are analyzed.

VILSAINT, C. L. et al. The Ecology of Early Childhood Risk: A Canonical Correlation Analysis of Children's Adjustment, Family, and Community Context in a High-Risk Sample. The Journal of Primary Prevention, v. 34, n. 4, p. 261-277, ago. 2013. ISSN 0278-095X, 1573-6547.	2 year-old children recruited at Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) sites (US, n=364).	Externalizing and internalizing behavior.	Neighborhood danger.	Defined by survey respondents.	Socioeconomic family covariates, parents' depression, alcohol and drug use, parenting daily hassles, home environment, conflict with child, parental discrimination, individual child characteristics.	Neighborhood danger is significantly associated with higher levels of behavior problems.
WINSLOW, E. B.; SHAW, D. S. Impact of neighborhood disadvantage on overt behavior problems during early childhood. Aggressive Behavior, v. 33, n. 3, p. 207-219, 2007. ISSN 1098-2337.	Boys living in a large metropolitan area in the Mid-Atlantic region of the US whose mothers received Women, Infant, and Children Nutritional Supplement Program (WIC) (US, n=218).	Externalizing and internalizing behavior.	Neighborhood disadvantage factor (median family income, % families below poverty, % households on public assistance, % unemployed, % single-mother households, and % bachelor degree or higher).	Census block group.	Socioeconomic family covariates, individual child characteristics, parent criminality, maternal depression, residential instability.	Neighborhood disadvantage has a negative association with behavior problems at age 6 (but not before). The relationship between both variables was not linear.
WOLF, S.; MAGNUSON, K. A.; KIMBRO, R. T. Family poverty and neighborhood poverty: Links with children's school readiness before and after the Great Recession. Children and Youth Services Review, v. 79, p. 368-384, ago. 2017.	Kindergarten students who were part of the Early Childhood Longitudinal Study-Kindergarten Cohorts (ECLS-K) from 1998 and 2010 (US, n = 19200/15700)	Math and reading skills, externalizing behavior, self-control and approaches-to-learning.	Neighborhood poverty (the percent of persons living below the federal poverty threshold).	Census tract.	Socioeconomic family covariates, individual child characteristics and location.	Neighborhood poverty is associated with lower academic skills; there is also a significant, but smaller, association with behavioral outcomes.

PALAMAR, J. J. et al. Family- and Neighborhood-Level Factors as Predictors of Conduct Problems in School among Young, Urban, Minority Children. Behavioral Medicine, v. 41, n. 4, p. 177-185, out. 2015. ISSN 0896-4289, 1940-4026.	Students who attended selected five schools (US, n = 489)	Behavior problems.	Percentage of poor residents.	Census tract.	Socioeconomic family covariates, individual child characteristics.	Neighborhood poverty is not associated with children's behavior problems at pre-K level. When time is taken into account, a significant and positive relationship emerges betwen neighborhood poverty and the annual increase in behavior problems.

Source: Elaborated with the 52 selected articles.

2 The case of São Paulo: spatial segregation and access to public services

2.1 Introduction

Brazil is a country marked by substantial vulnerability and high inequality: 25.3% of its population lives in poverty[1] and the Gini Index of distribution of household income distribution per capita is trending upwards, having reached its highest value since 2012 in 2018 at 0.545 (Instituto Brasileiro de Geografia e Estatística, Coordenação de População; e Indicadores Sociais., 2019) Brazil is also marked by spatial segregation between the rich and the poor. Such segregation can be found both when the country as a whole is examined, as well as when neighborhoods, within cities, are compared. There is a clear and long-lasting spatial segregation pattern between macro-regions where the South and the Southeast are the richest regions and with the Northeast being home to almost half of the Brazilian poor (Instituto Brasileiro de Geografia e Estatística, Coordenação de População; e Indicadores Sociais., 2019). There is also a pattern of spatial segregation within the Brazilian urban areas. The most vulnerable live in neighborhoods characterized by greater negative externalities - a consequence of the incomplete and/or low quality urban infrastructure, low quality and irregular housing, the lack of urban sanitation, with scarce public transport and with a concentration of violence (Instituto Brasileiro de Geografia e Estatística, Coordenação de Geografia, 2017).

Villaça (2011) defines spatial segregation in Brazilian cities as a spatial manifestation of the inequality that prevails in the Brazilian society. The combination of socioeconomic inequality and vulnerability with spatial segregation has a cumulative effect that creates a substantial burden for the poorest as it reduces social mobility and traps families in poverty (TORRES, 2004). Poor families, who live in segregated spaces, are exposed to worse living conditions and economic opportunities than the poor who live in non-segregated spaces (TORRES et al., 2003). Space segregation also affects delivery and access to public policies and services, isolating the most vulnerable who need such support and services even more (VILLAçA, 2011). Moreover, those who live in these neighborhoods tend to not have access to the diversity that cities can provide to those who live in less vulnerable neighborhoods (KUHNEN; SILVEIRA, 2008).

This segregation may have substantial consequences for the children. As discussed in chapter 1, children are greatly affected by neighborhood conditions and characteristics.

[1] This was calculated using the World Bank's poverty line of US$ 5.50 PPP (Purchasing Power Parity) (Instituto Brasileiro de Geografia e Estatística, Coordenação de População; e Indicadores Sociais., 2019).

These effects can be long-lasting and have consequences for different domains of life.

In this thesis, neighborhood effects are examined in one specific context, the city of São Paulo. To understand the following analysis and discussion, it is important to first understand spatial inequality within São Paulo, including the inequality in access to different public goods and services. In section 2.1 and section 2.3, a portrait of the spatial inequality in São Paulo as well as a discussion of the inequality in access, is presented. Of all public goods and services available, this thesis analyzes in-depth the role of one specifically, center-based care for young children. Due to this, it is important to analyze how this service is distributed in the city. This is done in the fourth section of this chapter. The final section presents some final remarks and connects to the rest of the thesis.

2.2 Spatial segregation

Since the 1970s, urban researchers have described the spatial segregation of metropolitan regions of Brazil using a pattern of center-periphery (VILLAçA, 2011; TORRES et al., 2003; MARQUES; BICHIR, 2001). This pattern is defined by a duality between the central area of the metropolitan region and those areas further away, near the borders of this region. The center is described as home to not only the more affluent population but also to most of the job opportunities, public goods and services. By contrast the periphery is described as home to more vulnerable populations who have worse access to the labor market. These the peripheral neighborhoods are also characterized by a lack or non-existence of public goods and services and as areas "forgotten" by the government (VILLAçA, 2011; TORRES et al., 2003; MARQUES; BICHIR, 2001).

In the past twenty years, this pattern of duality has been questioned. First, it became clear that the periphery was not an area without public investments, that peripheral neighborhoods were not totally devoid of public goods and services and that there was a significant expansion of these in the more recent decades (TORRES et al., 2003; MARQUES; BICHIR, 2001). Moreover, other transformations took place in the urban space where the rich have left the city center and the poor have occupied areas outside of the periphery(TORRES et al., 2003; MARQUES; BICHIR, 2001).

Reflecting upon these transformations, many scholars have started to re-think the dichotomy between center and periphery (NERY; SOUZA; ADORNO, 2019; TORRES et al., 2003; MARQUES; BICHIR, 2001). For them, the periphery should not be understood as a homogeneous poor area. Within peripheral neighborhoods there is plenty of heterogeneity when well-being indicators or public investment are analyzed (TORRES et al., 2003; MARQUES; BICHIR, 2001; BUGNI; JACOB, 2017). Moreover, within these neighborhoods there are "critical areas" with a high concentration of negative indicators with a very vulnerable population which is exposed to severe physical risks, such as flooding

or landslides for example and are without access to public goods and services (TORRES et al., 2003).

There is, however, a consensus that Brazilian cities are spatially segregated. Generally, the poor live in areas with worse urban infrastructure and living conditions and the rich, in areas with better conditions (Instituto Brasileiro de Geografia e Estatística, Coordenação de Geografia, 2017). This is a consequence of different processes that shaped the urban space throughout the years - processes operated by the government, by society and by the economic elites (Instituto Brasileiro de Geografia e Estatística, Coordenação de Geografia, 2017). It is not within the scope of this thesis to discuss these processes, but it is important to acknowledge them.

2.2.1 São Paulo

As the largest metropolitan region in Brazil and Latin America, São Paulo has been the focus of much of the research on spatial segregation in Brazil. Many studies have described the city's spatial segregation as having the previously discussed pattern of center-periphery (MARQUES; BICHIR, 2001). Evidence from the 2000s, however, found much more heterogeneity in the spatial distribution of the population than previously described and expected by the literature (TORRES et al., 2003). Despite the highest levels of vulnerability being found in districts further way from the city center, in the north, east and south of the Metropolitan Region, and the lowest levels being located in the central area of the city, there was high variability of vulnerability, particularly in the periphery (TORRES et al., 2003). The center-periphery model seems therefore, to be a simplification of the occupation of the urban space (TORRES et al., 2003). Evidence from São Paulo triggered much of the reflection discussed in the previous section. More recent evidence corroborates this conclusion for São Paulo (BUGNI; JACOB, 2017; NERY; SOUZA; ADORNO, 2019).

To describe and explore the spatial segregation within the city of São Paulo, we follow Bugni e Jacob (2017) and use a database from Ipea with information on the vulnerability of Brazilian metropolitan regions. The database measures vulnerability through a multi-dimensional index - the *índice de vulnerabilidade social* (IVS), an index of social vulnerability that provides information at the human development unit (HDU) level[2]. This multi-dimensional index is composed of sixteen indicators, divided in three dimensions: urban infrastructure, human capital and income and labor[3] (COSTA; MARGUTI,

[2] These geographical subdivisions of metropolitan regions in Brazil were constructed according to socioeconomic homogeneity by combining contiguous *setores censitários* - the smallest subdivision in the Brazilian Census (COSTA; MARGUTI, 2015). They were validated by local researchers and should be recognizable for the population as a whole, as a neighborhood or, in some rare cases, a small city (COSTA; MARGUTI, 2015) . The city of São Paulo has 1.593 HDUs.

[3] For the urban infrastructure dimension, there were three indicators: (i) percentage of people in households with inadequate water supply and sanitation; (ii) percentage of the population living in urban

Figure 2.1 – Spatial Distribution of IVS in 2000

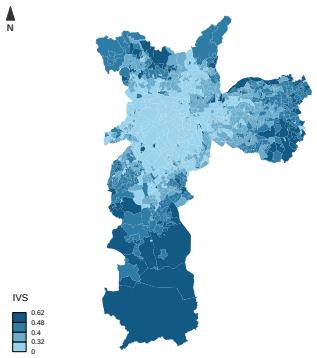
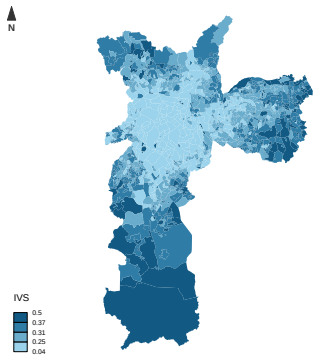


Figure 2.2 – Spatial Distribution of IVS in 2010



Source: Based on data from *IVS* from Ipea.

2015). It was developed by Ipea, in partnership with the United Nations Development Programme, with data from the Brazilian Census (COSTA; MARGUTI, 2015). IVS values range from zero, representing the lack of vulnerability in a neighborhood; to one, representing the maximum possible vulnerability (COSTA; MARGUTI, 2015).

Figure 2.1 and Figure 2.2 present the spatial distribution of social vulnerability index by human development unit in 2010 and in 2000. There is a clear pattern where the more vulnerable neighborhoods, those with higher IVS values, are located farther from the city center; in the southern and eastern city limits, and the less vulnerable, those with lower IVS values, nearer the city center; similar to that described by Torres et al. (2003).

Compared to 2000, the level of vulnerability in São Paulo decreased in the first decade of the century and the global index improved its value by 21%, moving from 0.368 in 2000 to 0.291 in 2010 (BUGNI; JACOB, 2017). All three dimensions saw improvements with the income and labor dimension showing the largest one at 39% and the infra-

households without the garbage collection service; (iii) percentage of people living in households with *per capita* income less than half the minimum wage and who spend more than one hour to work in the total number of employed, vulnerable people who return daily from work. For human capital dimension, there were eight indicators: (i) infant mortality; (ii) percentage of children aged 0 to 5 who do not attend school; (iii) percentage of children aged 6 to 14 who do not attend school; (iv) percentage of women aged 10 to 17 years who had children; (v) percentage of mothers who are heads of households, do not have complete elementary school education and have at least one child under 15 years of age in the total of mothers who are heads of households; (vi) illiteracy rate of the population aged 15 or over; (vii) percentage of children living in households where none of the residents have completed elementary school; (viii) percentage of people aged 15 to 24 who do not study, do not work and are vulnerable to poverty in the total population of this age group. For the income and labor dimension, there were five indicators: (i) proportion of people with *per capita* household income equal to or less than half the minimum wage in 2010; (ii) unemployment rate of the population aged 18 or over; (iii) percentage of people aged 18 or over with no complete elementary education and informally employed; (iv) percentage of people in households vulnerable to poverty and dependent on the elderly; (v) employment rate of children between 10 to 14 years old.

structure dimension the smallest at 0.2% (BUGNI; JACOB, 2017).

There was also a reduction in the number of neighborhoods, HDUs, with a very high level of vulnerability[4]. No neighborhood in the city of São Paulo had this classification in 2010 - a reduction of 229 neighborhoods (BUGNI; JACOB, 2017).

Even though the vulnerability in the city of São Paulo significantly reduced in the first decade of the 2000, inequality is still present: Bugni e Jacob (2017) find evidence that there is spatial segregation in the city of São Paulo using both a Moran Index of the IVS and local indicators of spatial associations (LISAs). Moreover, there is evidence that the segregation has risen in the city: when measured by the Moran Index, it has risen 13% between 2000 and 2010 (BUGNI; JACOB, 2017).

The complete picture drawn by Bugni e Jacob (2017) with IVS data is corroborated by other analyses of spatial inequality in São Paulo. One example is the *Mapa da Desigualdade*, a combination of maps that plot over fifty indicators in ten dimensions - from mobility to education - for all 96 districts[5] of São Paulo[6]. Another example is the study from Nery, Souza e Adorno (2019) that concludes that the dichotomy center-periphery cannot be used to describe the pattern of spatial inequality in São Paulo. UUsing census data combined with mobility data and information on physical geography characteristics, these authors find eight different categories of urban occupation; none of which are reflected in the administrative subdivisions of the city (NERY; SOUZA; ADORNO, 2019). Instituto Brasileiro de Geografia e Estatística, Coordenação de Geografia (2017), in their analysis of spatial differentiation in Brazilian cities, describe the pattern of São Paulo as radial, but underscore the existence of neighborhoods with worse living conditions interspersed with those with better living conditions.

2.3 Inequality in the access to public services

The inequality in the access to public services and goods is intrinsically associated with spatial segregation. For some authors, the access to public services is part of what characterizes the segregation itself; examples can be found in the inclusion of the infrastructure dimension in the IVS or of basic sanitation services in the analyses done by (NERY; SOUZA; ADORNO, 2019). For others, a consequence of segregation is more difficult access to public goods and services for vulnerable populations. In a segregated city, those services tend to be located farther away from where the more vulnerable live and it takes this population more time - and, potentially, more money - to reach them

[4] Neighborhoods were classified as having a very high level of vulnerability when their IVS was between 0.501 and 1 (BUGNI; JACOB, 2017).

[5] The districts are administrative sub-divisions of the city.

[6] The maps and the methodology used can be found at: $https://www.nossasaopaulo.org.br/wp-content/uploads/2020/10/Mapa-da-Desigualdade-2020-MAPAS-site-1.pdf$.

(TORRES; BICHIR, 2007; BICHIR, 2009).

Evidence for São Paulo corroborates this association. Nery, Souza e Adorno (2019) find that access to basic sanitation services is an important element when spatial segregation in the city is analyzed, even though this service is almost universal to the residents of São Paulo. Bichir (2009) finds that the region of residence is the main driver of inequality in the access to basic urban infrastructure even though there is high coverage of many of these goods and services.

A set of studies analyze the inequality in the access to health care services. Torres e Bichir (2007) analyze the waiting time for consultations in the public health system and find that those who live in the most central regions have shorter waiting times than those who live in the periphery. Moreover, the authors argue that the spatial dimension has larger relevance in explaining differences in access (TORRES; BICHIR, 2007). Analyzing the access to the public healthcare system in a broader way, Coelho, Szabzon e Dias (2014) present a description of the evolution of this access by city region[7]. In the first decade of this century, there was reduction in the inequality of access[8], but access was still unequal: there was a step-wise pattern when the distribution of services was analyzed by city region in conjunction with the level of human development index of each region. Coelho, Marcondes e Barbosa (2019) investigate the evolution of access for a longer time period, from 2001 to 2016, and corroborate the findings of the previous study: the distribution of health services has become more equitable in São Paulo when the city regions are compared by their human development index.

In several of these studies, the role of social movements in advocating for the expansion of public goods and services to the periphery is discussed (TORRES; BICHIR, 2007; BICHIR, 2009; COELHO; SZABZON; DIAS, 2014) The authors suggest that these efforts could be, in part, responsible for the reduction in inequality of access that has been found in the literature (TORRES et al., 2003).

2.4 Center-based Care

The access to affordable, center-based care became a worldwide need as women's participation in the labor market rose and the urbanization process deepened (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; DUNCAN,). In Brazil, the need for center-based care gained strength in the 1960s as a consequence of the incorporation of middle class women into the labor market, along with a rise in the proportion of manufacturing jobs that required mothers to be away from their children for extended periods of time (OLIVEIRA, 1988). The government's response to such needs was to finance charities that

[7] This study uses *subprefeituras*, administrative regions in the city of São Paulo that aggregate the city's districts.

[8] The authors use a Human Development Index calculated at the municipal level.

provided center-based care. During this period, providing center-based care was still understood as an assistentialist practice, targeted at low-income children and their families (OLIVEIRA, 1988). In the seventies, however, this conception changed and center-based care started to be understood as a worker's right and a duty of the State. Organized women's movements played an important role in this transformation (OLIVEIRA, 1988). There was an increase in government supported daycare centers either directly run by cities or state-level governments, or by charities or other private organizations but financially supported by government (OLIVEIRA, 1988).

The demands for free, public provision of center-based care influenced the discussion of the National Constitution of 1988 and its final version included the provision of early childhood care and education for children between zero and five years old, as a constitutional right in Article 208. Other legislation followed. The principal Brazilian educational law, *Lei das Diretrizes Básicas da Educação*, defined that a duty of the State was to provide free early childhood education and care for children up to five years of age[9]. The law also defined the division between the care and education of children younger than age four, which would be the purview of daycare centers, and the education of the those between ages of four and five which would occur in preschool[10]. The latter also became mandatory. Finally in 2006, constitutional amendment 53 guaranteed the provision of financial support for early childhood care by the federal government to the municipal governments which are responsible for the provision of this service.

In line with this legislation, the 2010 National Educational Plan established two goals related to the expansion of access to early childhood education. By 2016, all children between ages four and five should be in preschool, and by 2024, at least 50% of children under three should be enrolled in center-based care[11]. These goals, however, are far from being achieved and as of 2015, only 30.4% of children between zero and three were attending center-based care and 90.5% of children between four and five attended preschools[12].

While the provision of center-based care is the responsibility of municipal governments in Brazil, federal government provides monetary support[13] as well as the curriculum guidelines[14]. However, unlike provision of other educational services such as primary and

[9] This can be found at article 4 of *Lei das Diretrizes Básicas da Educação*.

[10] It is worth noting that, originally, preschools were for children between ages 4 and 6. But the law nº 12.796, from 2013, changed this age limit.

[11] The National Educational Plan is available at: $http://portal.mec.gov.br/index.php?option = com_ocmanview = downloadalias = 7116 - pl - pne - 2011 - 2020 Itemid = 30192$.

[12] Data from *Observatório do Plano Nacional de Educação*. Available at: $http://www.observatoriodopne.org.br/indicadores/metas/1 - educacao - infantil/indicadores$.

[13] Through the main financing scheme for Brazilian public education the *Fundo de Desenvolvimento da Educação Básica* (FUNDEB) and also through special programs like *Programa de Apoio a Novos Estabelecimentos de Educação Infantil*, *Programa de Apoio a Novas Turmas de Educação Infantil* and *ProInfancia*.

[14] The main document is the *Base Nacional Comum*, approved in 2017, and contains learning ob-

secondary schools, not all free and public center-based care needs to be administered directly by municipalities. Cities are allowed to establish partnerships with third parties to allow them to administer publicly funded daycare centers[15]. In 2019, 71.39% of early childhood enrollment was in municipal centers and schools, 8.20% in centers run by third-parties in partnership with the municipal government and 19.73% in private centers[16].

2.4.1 The history of center-based care in São Paulo

The provision of center-based care in São Paulo followed a similar timeline to that discussed previously. While though the first centers to care for young children date back to 1935[17], it was only after 1966 that the local government began a structured early childhood care policy (CHALITA; CIPRIANO, 2016; OLIVEIRA; FERREIRA, 1986). The first centers supported by the municipality were inaugurated in the following years: 9 in 1967 and 7 in 1968 and 1969 (OLIVEIRA; FERREIRA, 1986)[18]. The seventies saw an intensification of social movement fights for the public provision of center-based care[19] (OLIVEIRA; FERREIRA, 1986; FERNANDES; DOMINGUES, 2017) that culminated in the implementation of the first city-run daycare centers in 1980 (OLIVEIRA; FERREIRA, 1986; FERNANDES; DOMINGUES, 2017).

During the next two decades, there was a gradual expansion of the number of children in the municipal network of daycare centers both in directly run and privately run, government-supported centers. Figure 2.3 presents the growth in enrollment in these types of daycare centers from 1985 to 2015[20]. It is important to point out that this graph contains information on enrollment for children between zero and five years of age; that is, of children in daycare centers and preschools.

jectives for the different education levels, including early childhood. This is available at: *http : //basenacionalcomum.mec.gov.br/*.

[15] This is only valid for children up to age three. These daycare centers are called *creches conveniadas* and the municipal government receives funding to support those institutions through Fundeb. This was established by Decree Nº 6.253 from 2007.

[16] 0.7% were in centers administered by state or federal government. This data came from the from the Technical Report of the 2019 Educational Census.

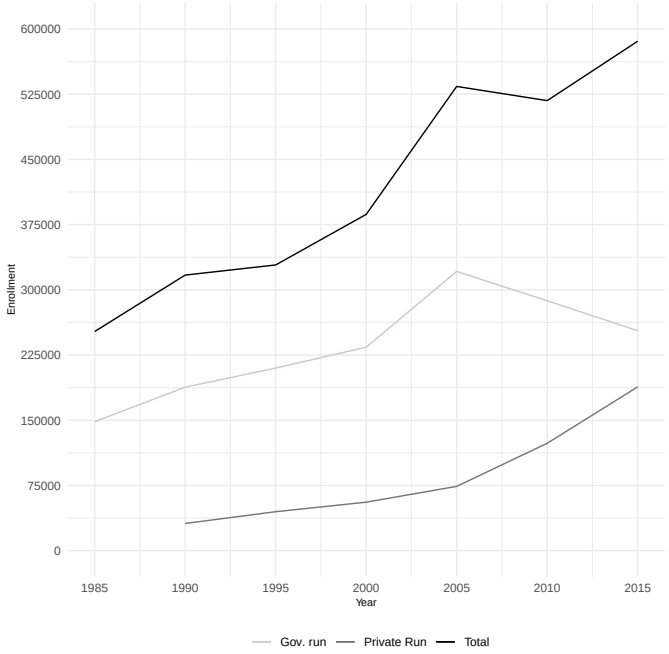
[17] Three centers, called *parques infantis municipais*, were created during Mario de Andrade's tenure at the city's, then, Department of Culture (CHALITA; CIPRIANO, 2016; OLIVEIRA; FERREIRA, 1986).

[18] All centers were were by third-parties with financial support from the city. After a dispute with one of these organizations, one of the centers became government-run in 1969 (OLIVEIRA; FERREIRA, 1986).

[19] Some groups were specially important in this fight: *Movimento de Luta por Creches*, *Pastoral do Menor*, *Sociedade Amigos de Bairro*.

[20] The data in this graph came from Chalita e Cipriano (2016). There was no information for the enrollment in private-run daycare centers in 1985.

Figure 2.3 – Expansion of Center-Based Enrollment, São Paulo, 1985-2015



Source: Based on data from Chalita e Cipriano (2016).

Even though the growth in total enrollment was substantial - from 252.070 children in 1985 to 585.942 children in 2015 -, it is clear that there was a change in the way such growth happened. If in the 90s the growth was mostly due to an increase in enrollment in government-run centers, between 2000 and 2015, this was substituted by a growth in enrollment in private-run, but government supported, centers. There was even a decrease in enrollment in government-run centers between 2005 and 2015.

Besides the continuing expansion of enrollment, the 2000s saw the approval of several important pieces of legislation that substantially changed early childhood care and education policy in the city of São Paulo. First, following federal legislation[21], the Municipal Department of Education became responsible for the daycare centers - both for their administrative and pedagogical structure[22]. There was also legislation that defined basic infrastructure conditions that all centers needed to adhere to[23] and that specified the criteria and the conditions for support of privately run centers[24]. More importantly, the enrollment process as discussed in more details in chapter 5 - was enhanced[25].

[21] More specifically, *Lei das Diretrizes Básicas da Educação* in its article 89 defines that daycare centers should be under the responsibility of municipal Department of Education.

[22] Before that, the centers were under the responsibility of the municipal department of social assistance. Several legislation discussed this matter. The law n. 13.326 of 2002 defined how this transition should happen.

[23] Instruction n. 3.479 of the municipal Department of Education.

[24] Instruction n. 3.477/2011 of the municipal Department of Education. This instruction defined the conditions the organizations should comply to apply for government support and also basic guidelines for the center operation. It also specified how the government should monitor such operation.

[25] Law 14.127 of 2006

Currently, São Paulo has 3,075 publicly funded early childhood centers[26]. In analyzing centers that offer care for children between zero and three years and eleven months in the municipal daycare system, 85% of the centers are run by third parties, in government owned buildings or not[27]. At the end of 2020, these centers served 374.631 children[28]. This subset of the early childhood education system of the city of São Paulo is the focus of this work.

All of daycare centers in the municipal system are open 10 hours a day[29]. Children are divided in four groups, according to age, and all groups have at least one teacher with a Bachelor's Degree in Education. The first group, *Berçário 1*, serves children from zero to one year and eleven months and this group has one teacher for seven students. The following group, *Berçário 2*, serves children from one year and two months to one year and ten months and this group has one teacher for up to nine children. Children from one year and eleven months to two years and ten months are in *Minigrupo 1*, with one teacher for up to twelve students and finally, *Minigrupo 2*, children from two years and eleven months to three years and ten months, are in groups of up to 25 children per teacher[30].

The city's curriculum guidelines are established in two documents: the *Currículo Integrador da Infância*, published in 2015, that includes guidelines for children between zero and twelve years of age, and the *Currrículo da Cidade – Educação Infantil*[31] published in 2019 building on the previous document and on experiences of the school system. Both directly run and privately run, government-supported centers develop their own curriculum based on these guidelines.

[26] There are six types of centers in the city serving children between zero and 5 years old: three of them are centers for children between zero and three years old - the *Centros de Educação Infantil Direto*, run directly by the municipal government, the *Centros de Educação Infantil Indireto*, centers located in city-owned buildings but administered by third parties, and the *Centros de Educação Infantil Parceiros*, the *Creches Conveniadas*, administered by a third party but recipients of public funding; there is also the *Centro Municipal de Educação Infantil* for children between zero and 5 years and 11 months of age, the *Escola Municipal de Educação Infantil* that serve children between 4 and five years of age and the *Centros de Convivência Infantil*. This information came from the city's Department of Education webpage: $https://educacao.sme.prefeitura.sp.gov.br/escolaaberta/conheca-a-rede.$

[27] Data available at: $https://educacao.sme.prefeitura.sp.gov.br/escolaaberta/conheca-a-rede.$

[28] Data from December, available at: $https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/$

[29] Daycare centers can independently choose the opening and closing times to serve the needs of their community.

[30] This information came from: $http://www.capital.sp.gov.br/noticia/readequacao-de-turmas-na-educacao-infantil-altera$

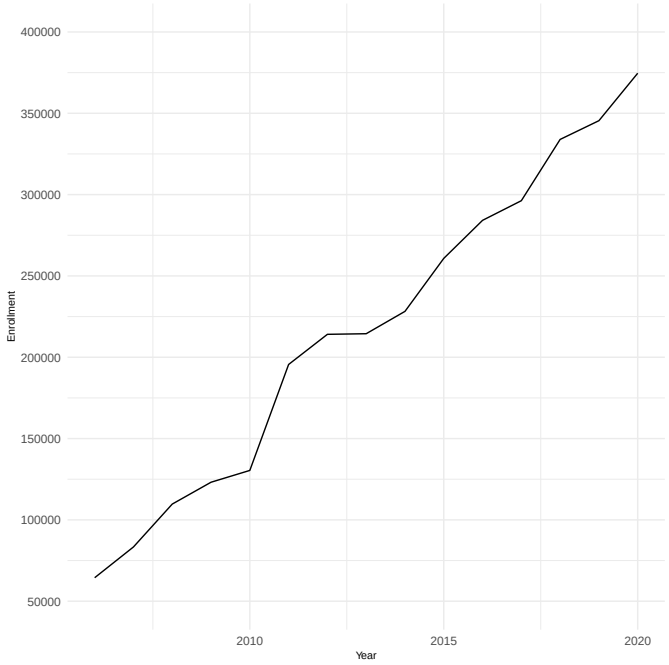
[31] In both guidelines, it is evident that there is a focus on not only caring for the young children – in the sense of bathing, feeding them, and taking care of all their physical needs – but also helping them develop as human beings in a rich social environment. There is also a concern with understanding that early childhood center should not be seen as places where young children are prepared for school; they are places were children play, interact with others and develop as a whole person.

2.4.2 The recent expansion of access

In the 2000s, the demand for daycare center spots became a point of contention between the city hall and the city's population. There was substantial pressure from civil society, through organized movements like *Movimento Creche para Todos*, and the judicial system through the Public Defenders office of the State of São Paulo, to expand the network of early education and care centers and to have a more efficient enrollment process (OLIVEIRA; SILVA; MARCHETTI, 2018).

As with the movement in the 70s which generated the early expansion of the municipal system, there was also a response from the city government with an expansion in the number of centers and enrollment. Figure 2.4 depicts the growth of enrollment for children in *creches*. There was a significant expansion in the enrollment of children between zero and three years and eleven months. In December of 2006, there were 64,436 children enrolled and in December of 2020, 374,631[32].

Figure 2.4 – Expansion of Center-Based Enrollment for Children 0-3, São Paulo, 2006-2020



Source: Based on data from the Department of Education at: https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/.

This growth has been mostly stable during the years analyzed - with a median growth of 13% per year, except for the period between December of 2010 and December of 2011, when the enrollment grew 50%.

[32] Data from the Department of Education at: $https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/$.

To determine if the growth in enrollment was well distributed throughout the city, we analyzed the enrollment by city district[33] in 2006 and 2020[34]. Figure 2.5 and Figure 2.6 present this information.

Figure 2.5 – Enrollment in Center-based care by district, São Paulo, 2006

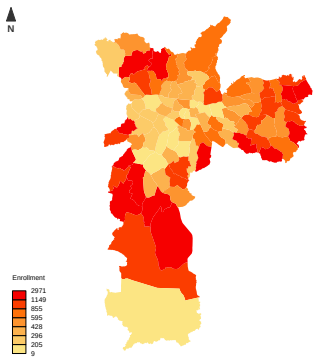
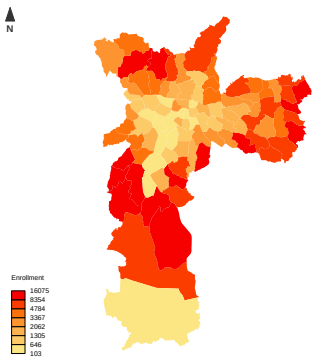


Figure 2.6 – Enrollment in Center-based care by district, São Paulo, 2020



Source: Based on data from the Department of Education at: https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/.

In both years, most of the enrollment of children between zero and three years and eleven months in center-based care is located in the peripheral neighborhoods. In 2020, this was especially true: the center of São Paulo shows lighter colors indicating lower enrollment while the peripheral neighborhoods show mostly dark colors indicating higher enrollment. Two factors may explain this pattern. First, most of the target groups for municipal daycare centers live in the periphery; children whose mothers work and need affordable care. Second, the periphery also has a larger number of children. These areas have experienced higher demographic growth in the past several decades (MARQUES; REQUENA, 2013). As a result, it is not sufficient to analyze only the absolute size of enrollment in center-based care to reach any conclusions on distribution of enrollment in the city.

Figure 2.7 and Figure 2.8 present two maps that allow us to better understand how enrollment grew by city district. First, Figure 2.7 presents a map of the spatial distribution of the absolute growth in enrollment in center-based care for children between zero and three years and eleven months. To calculate this, for each district, the enrollment in 2006

[33] Districts are administrative subdivisions of São Paulo. The city has 96 districts. The HDUs discussed in the previous section can be grouped into districts and districts can be grouped into administrative regions.

[34] It is important to point out that, although the information indicates the location of the center the child is enrolled in, and not where the child lives, it is a reasonable hypothesis that children live within the same district as the center they attend is located, as enrollment is by proximity. More about the enrollment process is discussed in the next section.

was subtracted from the enrollment in 2020. Relative growth in percentage points was then calculated by dividing this difference by the size of enrollment in 2006. The spatial distribution of this growth is in Figure 2.8.

Figure 2.7 – Absolute growth in enrollment in center-based care by district, São Paulo

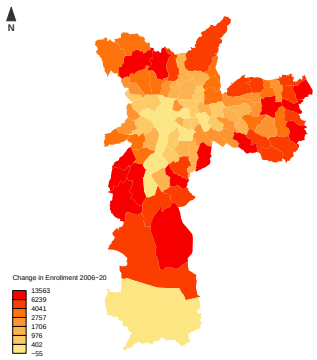
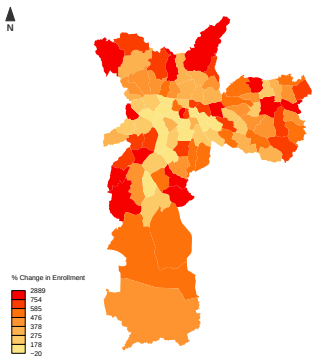


Figure 2.8 – Relative growth in enrollment in center-based care by district, São Paulo



Source: Based on data from the Department of Education at: https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/.

The first map indicates that the larger increases in center-based care enrollment occurred in peripheral neighborhoods. For example, the largest increase was in the district of Grajau in the south of the city where in 2020 there were 16,075 children between zero and three years and eleven months enrolled in centers in that district, an increase of 13,563 children in 14 years.

In addition to analyzing absolute growth, it is important to understand whether growth happened in neighborhoods that already had large enrollments or not. The second map allows us to do this as it presents the spatial distribution of relative growth. It indicates that the pattern is less clear when the relative growth is analyzed as even though there is large relative growth in several neighborhoods outside the city center, there are neighborhoods near the center with significant relative growth too. There was not a clear pattern of expansion of the service in the periphery as the maps in Figure 2.7 and Figure 2.8 can imply. Analysis of the size of the waiting lists, covered in the next section, allow us to have a better understanding this movement.

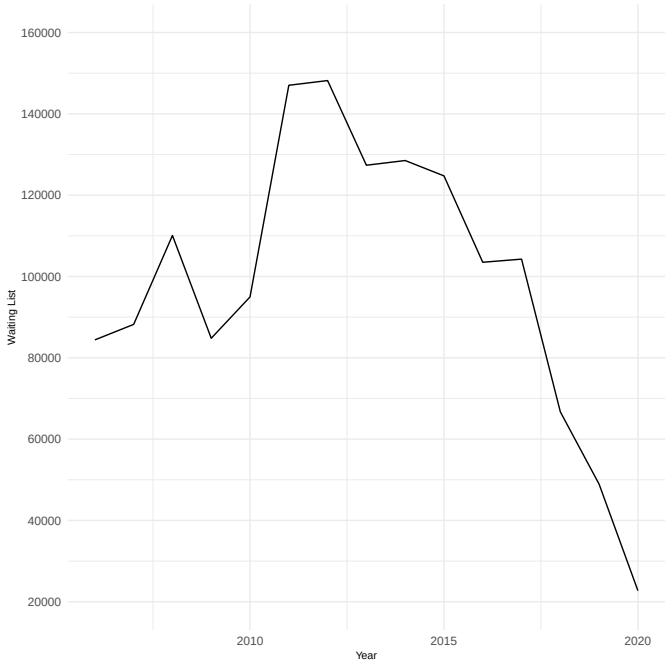
2.4.3 Waiting Lists

Even with a significant expansion of enrollment in the city as depicted in the previous section, there was a significant excess of demand for spots in the municipal daycare system throughout the 2000s and 2010s. As a result, over the past fifteen years,

waiting lists for a spot in the municipal daycare system were significant. There was a significant number of children waiting for a chance to be enrolled in center- based care for the entire period analyzed. This became a topic of dispute and an important political platform.

Figure 2.9 presents the evolution of the number of children on waiting lists between 2006 and 2020[35].

Figure 2.9 – Number of Children in waiting lists for Center-Based Enrollment for Children 0-3, São Paulo, June of 2006-2020



Source: Based on data from the Department of Education at: https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/.

There was a tendency of increase in the size of the waiting list between 2006 and 2012 and then, a drastic reduction after 2012. This reduction was not only in the magnitude of the waiting lists, but also in their relative size compared to enrollment[36]. In 2006, there were 84,408 children on waiting lists and 61,729 children enrolled in center-based care and the number of children on waitlists were 136% of the number of children receiving care. In 2012, the number of children on waiting lists reached its highest number, 148,185, but the number of children enrolled had also grown considerably, reaching 207,605 in June of that year. With that, in 2012, the size of the waiting lists represented 71.4% of enrollment. In 2020, there were 22,732, the lowest number of children on

[35] To calculate this, we used the waiting list in June of each year. We have data on the size of the waiting list on March, June, September and December of each year.

[36] We opted to compare the enrollment in June, and because of that, this data does not correspond perfectly with information depicted in 2.5.

To understand how the size of the waiting lists were distributed spatially in São Paulo, Figure 2.10 and Figure 2.11 present maps with the relative size of the waiting lists for center-based care by district, for both 2006 and 2020. To calculate this relative size, the number of children on waiting lists in June of 2006 was divided by the enrollment in June of that year; the same was done for 2020.

Figure 2.10 – Relative size of waiting lists in daycare centers by district, São Paulo, 2006

Figure 2.11 – Relative size of waiting Lists in daycare centers by district, São Paulo, 2020

Source: Based on data from the Department of Education at: https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/.

In comparing the maps, we can see a change in the pattern between 2006 and 2020. In 2006, the relative size of the waiting lists was higher in districts in the north of the city and in the south, close to the city center. In 2020, the relative size of the waiting lists was larger in the districts on the southern boundary of the city

The increase in the size of the waiting list in the beginning of the period depicted in Figure 2.9 is possibly a consequence of a change in the enrollment system. In 2006, the enrollment process became a centralized informational procedure that took into account the child's address and the place he/she was on waiting lists for a specific age-group[37] in centers close to their home. In chapter 5, this process is explained in detail.

With the implementation of this procedure, the city's Department of Education began to keep track of its waiting lists and later, began to publicize them[38]. Before the centralized procedure was adopted, parents had to register their child on the waiting

[37] Defined according to the child's date of birth and the ordinances from the city's Department of Education.

[38] This system started in 2006, but the waiting lists were not made available to the public until 2008. For more information, see Oliveira, Silva e Marchetti (2018). Currently, one can find the waiting lists from 2008 to 2020 in: $https://educacao.sme.prefeitura.sp.gov.br/acesso-a-informacao/demanda-escolar/$.

list for a place in each specific daycare they were interested in (OLIVEIRA; SILVA; MARCHETTI, 2018). With the centralized system, the cost for parents to register their children reduced which in turn can explain the increase in waiting list size. The increase in the waiting lists can also be a consequence of new centers opening in neighborhoods that did not previously have them which also reduces the cost of enrolling a child.

The database Throughout this thesis, we use a database from the enrollment process of the ECEC system of São Paulo[39]. This database includes all children who entered the system between 2010 and 2018 and/or whose registration was updated in any manner during this period[40]. It provides information on the child's trajectory within the enrollment process of the municipal ECEC system[41], including: (i) the date the child's parent requested a spot in the system; (ii) the date of the child's placement in a certain school and age group; (iii) whether the child accepted the placement; (iv) the date of the child's enrollment in a certain school and age group; and (v) whether the child had any kind of priority in placement[42]. It also contained data on child characteristics such as gender, race, month and year of birth, as well as the parents official ID, their *Cadastro de Pessoa Física* number and the family's residential postal code at the time of request.

This database was anonymized by the municipal Department of Education and ID codes were created[43]. These ID codes allowed us to match this information with information on child proficiency as measured by the *Provinha São Paulo*, and discussed at length in chapter 3. It is important to point out that there could be multiple requests per child[44], but, in all cases, the ID code stayed the same.

2.5 Discussion

Analysis of the social vulnerability index data at the human development unit level for the city of São Paulo corroborated the literature on spatial segregation. It indicated that, in the city, the more vulnerable neighborhoods, where most of the vulnerable families live, are clustered together, generally, outside the city center.

This pattern is associated with the access to public services and goods. The literature suggested that there should be an association between neighborhood vulnerability

[39] This database was shared with us using an special permission.

[40] This might include children whose parents registered them for a spot, children who were offered a spot, those who denied a spot, those who gave up their place or those who enrolled in between those years.

[41] The process was also used for placement in municipal preschools. This thesis does not analyze this part of the ECEC system.

[42] In chapter 5, the types of priorities are discussed in details.

[43] These codes do not match the official ID code of the child in the city's educational system.

[44] Children could be taken off of the waiting lists and put back on a couple of months later. Or parents could deny a spot and enter the list again.

and access to public services. For the city of São Paulo, evidence for the access to health services corroborated this theoretical hypothesis as districts with higher vulnerability had worse access to basic health services.

This thesis is concerned with one particular public service: the provision of free center-based care for young children. The expansion of this service in the past twenty years occurred across the city. Even though the neighborhoods outside the city center, the most vulnerable, suffer with large wait-lists for spots in the daycare centers, it was in those neighborhoods that enrollment increased the most in the past fifteen years. One limitation for this analysis should be pointed out. It would be important to understand demographic growth in these neighborhoods as this could be related to the large waitlists. In addition, it would also be interesting to compare waitlists with the vulnerability of neighborhoods. However, as data for waitlists were aggregated by district and vulnerability is at the HDU level, this analysis requires further assumptions and this is not done here.

The focus on the provision of center-based care for younger children, which will be discussed in detail in chapter 4 and chapter 5, can be justified because of the evidence that center-based care can have profound impacts in the development of young children, specially those from more vulnerable contexts (CURRIE, 2001; CAMILLI et al., 2010; ENGLE et al., 2011; YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; BRITTO et al., 2017; MCCOY; WALDMAN; FINK, 2018). Considering that São Paulo is marked by spatial segregation and that neighborhood conditions matter to child development - as discussed in chapter 1 and tested in chapter 2 -, children in the city can have very distinct development trajectories. The access to center-based care can, then, be an important tool in the reduction of inequalities if its impacts on child development are true for the city's context - as tested in chapter 5.

This chapter provided descriptive information on spatial segregation in São Paulo as well as the context and the access to center-based care in the city. This information is important to set the stage for the empirical analyses carried out in the next chapters. It is important for interpreting and better understanding the results of the different analyses.

3 Testing Neighborhood Effects in São Paulo

3.1 Introduction

Several fields of study, from psychology and public health to education and criminology, have studied the effects of "place" in peoples' lives. There is robust evidence that the context in which people live, be it the neighborhood, the city, or the region, influences well-being, earnings, and life opportunities. A specific body of research focuses on the effect of the immediate geographical space around one's home, the neighborhood[1] and analyzes the impact of this immediate context - known as the neighborhood effect - for different outcomes and different populations.

The evidence for neighborhood effects is particularly robust for children and adolescents. Analyzing the Moving to Opportunity Program, Chetty, Hendren e Katz (2016) find evidence that the neighborhood one moves to before age 13 has impacts on their life outcomes in their young adulthood. In chapter 1, the existing evidence of neighborhood effects on children's development was also discussed at length. Taken together, this evidence suggests that to better understand and explain persistent social inequalities, it is important to study the neighborhood as the characteristics of the neighborhood, especially its socioeconomic vulnerability, are related to outcomes in children and adolescents; particularly in academic performance (CROWDER; SOUTH, 2011; CHETTY; HENDREN; KATZ, 2016; MORRISSEY; VINOPAL, 2018b; WODTKE; HARDING; ELWERT, 2011; DUPERE et al., 2010; LEVENTHAL; BROOKS-GUNN, 2000).

To contribute to this discussion by bringing evidence of São Paulo neighborhood effects in academic outcomes, a novelty in the literature, this study uses a sample of second grade students from São Paulo's municipal education system to test the hypothesis of whether there is a relationship between the vulnerability of the residential neighborhood that a child grows up in and their achievement in early elementary school. Following the literature on neighborhood effects, our main hypothesis is that the vulnerability of the neighborhood could negatively impact children's outcomes.

The remainder of this chapter is organized as follows. In the next section, evidence of neighborhood effects in education is reviewed - with special focus on the early grades of elementary school. Research questions are presented in the third section. The data and the methodology are introduced in the fourth section, while the fifth section presents the results. The sixth section presents a critical discussion of the results, followed by a

[1] In this literature, the neighborhood can correspond to areas of different sizes, and generally, is defined by by administrative barriers. There are studies that define the neighborhood as a subdivision of a city - like a district -, as a zip-code, or even a school district or a policy precinct.

discussion of the limitations of this study and potential future studies.

3.2 Neighborhood Effects in Educational Outcomes

Evidence indicates that the neighborhood a child experiences while growing up has a significant effect on their short-term development and also on their long-term results (CHETTY; HENDREN; KATZ, 2016; MINH et al., 2017; LEVENTHAL; BROOKS-GUNN, 2000). The negative effect of neighborhood's socioeconomic level on both short and long-term outcomes is specially well-documented: growing up in a poor neighborhood affects children's outcomes above and beyond their family's socioeconomic condition - chapter 1 provides evidence for short-term outcomes and Chetty, Hendren e Katz (2016) presents some long-term evidence.

The literature that studies neighborhood effects on children's development was discussed at length in chapter 1. The evidence presented indicates that the neighborhood a child grows up in matters for their development as measured by a myriad of indicators, from those that focus on cognitive abilities to behavioral ones. Because of such neighborhood effects, we can hypothesize that children who grow up in more disadvantaged neighborhoods, defined according to their socioeconomic level or other characteristics, begin school behind their peers who grew up in better-off neighborhoods. The existence of this gap is specially concerning as evidence indicates that early school skills predict a child's later educational outcomes (DUNCAN et al., 2007). Another concern is that children who grow up in more disadvantaged neighborhoods tend to be from more vulnerable families and therefore growing up in a disadvantaged neighborhood compounds the established negative consequences of family poverty (MORRISSEY; VINOPAL, 2018c).

The research on the role of the neighborhood context, specifically its socioeconomic level, on predicting gaps at school entry is still limited but confirms this hypothesis (WOLF; MAGNUSON; KIMBRO, 2017). Analyzing children enrolled in Head Start centers in the USA, different studies find evidence that low neighborhood socioeconomic status predicted some dimensions of children's school readiness; math skills for example, but not others like teacher-reported behavior problems or approaches to learning (HANSON et al., 2011; VADEN-KIERNAN et al., 2010; MCCOY et al., 2015). Generally, cognitive results were more affected by neighborhood socioeconomic status than non-cognitive or behavioral results (HANSON et al., 2011; VADEN-KIERNAN et al., 2010; MCCOY et al., 2015).

These findings are also supported by analysis of children's cognitive skills in kindergarten. An emerging body of studies use the Early Childhood Longitudinal Study-Kindergarten Cohorts (ECLS-K), a nationally representative database of USA children, to study this matter. Using the ECLS-K from 1998 and 2010, Wolf, Magnuson e Kimbro

(2017) find that both family poverty and neighborhood poverty are associated with children's lower levels of academic skills in the autumn of their kindergarten year. The authors find that at the start of the school year, differences across neighborhoods with differing poverty levels is substantial as children who grow up in the highest poverty neighborhoods are almost a full year of learning behind their peers who grew up in the lowest poverty neighborhoods (WOLF; MAGNUSON; KIMBRO, 2017). Morrissey e Vinopal (2018a), Morrissey e Vinopal (2018b), Vinopal e Morrissey (2020) use more robust methodologies and find similar results. During the fall of kindergarten, there is a substantial difference in terms of math and reading skills between those children who live in poorer neighborhoods and those children who live in low-poverty neighborhoods even when family and child characteristics are controlled for Vinopal e Morrissey (2020) estimate a gap in math scores between children from low poverty neighborhoods and those from high poverty neighborhoods of over half a year of learning. These authors also analyze behavior and socio-emotional outcomes and find no significant neighborhood effects (MORRISSEY; VINOPAL, 2018a; MORRISSEY; VINOPAL, 2018b). Moreover, the authors find very little indication that the effect of neighborhood poverty varied according to the child's own family poverty (MORRISSEY; VINOPAL, 2018a; MORRISSEY; VINOPAL, 2018b).

Evidence from the ECLS-K suggests that the school system does not succeed in reducing this gap, at least in the early elementary years. Morrissey e Vinopal (2018a), Morrissey e Vinopal (2018b) find a persistent negative association between neighborhood poverty and children's achievement - in math and reading - from Kindergarten to spring of second grade. These associations hold when a myriad of covariates are included in the model (MORRISSEY; VINOPAL, 2018a; MORRISSEY; VINOPAL, 2018b). Analyzing child growth trajectories in math and reading skills from kindergarten to third grade, Vinopal e Morrissey (2020) find some evidence of a catch-up during kindergarten. Children from poorer neighborhoods learn more relative to their peers from low-poverty neighborhoods but this does not hold true for other school years, particularly for reading skills (VINOPAL; MORRISSEY, 2020). These patterns did not vary in a statistically significant way by children's sex, household poverty or urbanicity however they did vary by children's early childhood education experience, race, Hispanic ethnicity and parents immigrant status (VINOPAL; MORRISSEY, 2020).

When neighborhood advantage - and not vulnerability - is analyzed, the results corroborate previous findings. Using a database from National Institute of Child Health and Human Development (NICHD), Dupere et al. (2010) analyze the relationship between a neighborhood advantage index and children's vocabulary, reading and math skills and the authors find a positive significant association in first grade. However, when the annual linear changes in child achievement, between four and a half years old and grade five (for reading) or 15 years old (for vocabulary and math), were analyzed, no statistical association is found with neighborhood poverty (DUPERE et al., 2010)

Evidence from another American nationally representative database, the Panel Study of Income Dynamics, goes in another direction and finds that children who live in poorer neighborhoods experience smaller growth in their math achievement than do children in less poor neighborhoods (PEARMAN, 2019). This effect is robust to controlling for prior achievement, parental cognitive ability, and school level characteristics (PEARMAN, 2019).

Taken together, all of these results can be interpreted as evidence that neighborhood socioeconomic vulnerability sets children on a lower achievement course at the beginning of their schooling and that the school system is not capable of changing this, at least not in the child's first years of schooling. This hypothesis is in line with findings from the literature that analyzes the existence of long-term neighborhood effects on individual outcomes. There is evidence that the neighborhood where one grows up impacts long-term individual results from educational attainment to the probability of marriage and teenage pregnancy (CROWDER; SOUTH, 2011; WODTKE; HARDING; ELWERT, 2011; WODTKE, 2013; CHETTY; HENDREN, 2018; HARDING et al., 2010).

Crowder e South (2011) analyze the Panel Study of Income Dynamics (PSID) and find evidence that the socioeconomic level of the neighborhood as measured by a multi-item, standardized index, is positively and significantly correlated with the probability of graduating high school. Using the same database, but estimating a marginal structural model and analyzing neighborhood disadvantage, Wodtke, Harding e Elwert (2011) find evidence that growing up in more disadvantaged neighborhoods has negative consequences on the probability of graduating high school among Blacks. These authors correct their results using an inverse probability of treatment weighting to adjust for confounding by time- varying covariates. Their results are larger than those from conventional regression models and indicate that prior results, as reported in the literature which indicate a not so substantial influence of neighborhood poverty, might be a consequence of estimation issues. Also using the Panel Study of Income Dynamics (PSID), combined with Census Data, Wodtke (2013) analyzes a cohort of children from age 4 to 19, collecting yearly information on neighborhood context and other characteristics, to estimate the effects of neighborhood poverty on adolescent parenthood. The results indicate that growing up in a poorer neighborhood has effects positive effects on adolescent parenthood. Chetty e Hendren (2018) use tax record data from 1996 to 2012 to identify a sample of children born between 1980 and 1988 whose parents moved once between 1996 and 2012. This sample is used to analyze the relationship between neighborhood conditions and different children's outcomes - their future income, their employment status, their college attendance, the probability of teenage birth and of marriage. The authors conclude that moving to a better neighborhood during childhood has positive impacts on the long-term outcomes analyzed. Additionally, Chetty e Hendren (2018) found robust linear relations between each additional year in a better or worse "quality" neighborhood and a collection

of individual outcomes in adult life, providing evidence that the amount of time a child is exposed to vulnerable neighborhoods matters.

Experimental evidence supports and reinforces this conclusion. Combining data from the Moving to Opportunity (MTO) study with tax record data, Chetty, Hendren e Katz (2016) analyze the impact of the MTO voucher treatment on long-term outcomes. Their findings indicate that moving to a better neighborhood before age 13 has positive impacts on children's college attendance, on the quality of the college these children attend, on earnings in young adulthood and also on marriage and fertility behavior. These impacts were not found for children who moved after age thirteen. Previous studies of the MTO experiment had not found any significant impact on short-term test scores of treated children (SANBONMATSU et al., 2006).

Even though most empirical research on neighborhood effects has not explored the mechanisms through which neighborhoods' vulnerability impact individual outcomes, classical theories, from the social-ecological model from Bronfenbrenner to more contemporary sociologists, point to two main pathways: a structural and a relational one (MCCOY et al., 2015; HARDING et al., 2010). The first pathway is related to the access to quality institutional resources, from schools and child care centers, to groceries, pharmacies and parks (DUPERE et al., 2010; MCCOY et al., 2015; SMALL; NEWMAN, 2001). This pathway is also related to the quality of the physical environment; the presence of toxins and of pollution is an indicator of this quality (DUPERE et al., 2010; MCCOY et al., 2015; WOLF; MAGNUSON; KIMBRO, 2017).

The second pathway is related to social aspects of the neighborhood and of its network; the social isolation of the residents of more vulnerable neighborhoods - from networks that provide job opportunities and that connect them to the city as a whole - is part of it (WODTKE, 2013). The level of social cohesion and the existence of collective norms of the neighborhood are also part of this pathway; these are related to both the level of support the parents receive from the community as well as the level of safety in the neighborhood (DUPERE et al., 2010; MCCOY et al., 2015; SAMPSON; MORENOFF; GANNON-ROWLEY, 2002) Whether the structural or the relational pathway are true, the theoretical explanations of neighborhood effects rely on the spatial interdependence between residents: it is the fact that people, and institutions, are located in proximity that generates such effects (ZANGGER, 2019).

There is an emerging body of research that focuses on the role of childcare centers in shaping neighborhood effects, analyzing one component of the first pathway described above. This is of special interest to this thesis.

Although there is evidence that childcare experiences are related to neighborhood characteristics, there is no clear evidence that child care centers and children's experiences in it, are responsible for neighborhood effects (DUPERE et al., 2010; MCCOY et al., 2015;

VINOPAL; MORRISSEY, 2020). Dupere et al. (2010) find evidence that the quality of child care environment is a mediator between neighborhood advantage and children's vocabulary scores. But, McCoy et al. (2015), analyzing a sample of Head Start students and focusing on neighborhood disadvantage, did not find any statistically significant evidence of the childcare center's role as a mediator of neighborhood effects and different child development indicators. Vinopal e Morrissey (2020) also do not find evidence that childcare experience affects the association between neighborhood poverty and children's achievement, at least during their kindergarten year. More detailed exploration of the potential role of childcare centers as a mediator of neighborhood effects is needed and this thesis contributes to this in other chapters.

3.3 Modelling neighborhood effects

The methodological difficulties of estimating neighborhood effects have been discussed at length in the literature (e.g. Harding et al. (2010) and Blume et al. ()). A good starting point for understanding such difficulties is Manski's discussion of identification problems in neighborhood effects research and his proposed model, called a linear-in-means model (MANSKI, 1993).

Let Y denote be the outcome of interest, X, individual independent characteristics, and Z, attributes that characterize a predefined neighborhood. Neighborhood effects can be modelled as:

$$Y_{ig} = \beta_0 + \beta_1 X_i + \beta_2 Z_g + \beta_3 m_{ig}^e + e_i \tag{3.1}$$

where i identifies the individual, g, the neighborhood and m_{ij}^e denotes the average outcome in the group g; e_i the error term. Following Manski's original formulation, β_1 represents direct influence of individual attributes on the outcome, β_2, the exogenous neighborhood effect - the effect of the characteristics of neighborhood on individual outcomes -, β_3, the endogenous neighborhood effect - the effect of neighbors' attributes on the individual outcome.

Most empirical research estimates a reduced form version of equation 1.1 in that the Y_{ij} is determined entirely by observables and the individual-specific error (Blume et al., 2010). This reduced form can be described by:

$$Y_{ig} = \theta_0 + \theta_1 X_i + \theta_2 Z_g + e_i \tag{3.2}$$

where θ_2 represents the parameter of interest. Generally, Z is modelled as an aggregated measure of a neighborhood's characteristic, such as its socioeconomic vulnerability or some sort of physical characteristic. If the main objective of such research is to isolate

neighborhood effects - the total effect of the exposure to one neighborhood instead of another - then, the differences between endogenous and exogenous effects do not matter (HARDING et al., 2010). But, even in this case, it is still necessary to deal with the selection bias to find a causal estimate.

Selection bias exists because, as pointed out by Leventhal e Brooks-Gunn (2003), "It is difficult to separate the effect of neighbourhood on children's developmental outcomes from the effect of the family, as families often choose where they live". Even if families do not completely chose where they live, the reasons behind such choice can be directly related to children's development. This causes selection bias. Therefore, it is questionable whether differences in outcomes are due to neighborhood effects or reflect the differences between residents of different neighborhoods. To overcome such problem, it is necessary to find an exogenous source of variation in neighborhood contexts - the Moving to Opportunity experiment, for example, is a study design that managed to do just that.

However, even if the selection bias problem is resolved, there is still an issue about how to distinguish between endogenous and exogenous neighborhood effects. This is a consequence of the "reflection problem" - using Manski's terminology -, the difficulty of identifying the direction of causality between individual and group characteristics. To solve this issue and correctly identify the model originally described in Equation 3.1, many proposals are presented in the literature (e.g. Blume et al. () for a complete overview). In this chapter, we implement a multilevel model - one of the ways to change model specification and deal with the reflection problem.

Multilevel models, also called hierarchical models, conceptualize neighborhood effects on individual outcome in a different manner as detailed in the following model:

$$\begin{aligned} Y_{ig} &= \beta_{0g} + \beta_{1g} X_i + e_i \\ \beta_{0g} &= \beta_0 + \beta_2 Z_g + \beta_3 m_g \\ \beta_{1g} &= \beta_1 + Z_g * \Psi + m_g * \phi \end{aligned} \tag{3.3}$$

Basically this type of model adds cross products of variables in 3.1 to allow for non-linearity and thereby addresses the reflection problem discussed previously (BLUME et al.,). It is important to point out, though, that even if identification holds, the estimations from this model might be imprecise if m_g is highly correlated with individual determinants (BLUME et al.,).

Although multilevel models do not estimate causal impacts, as the selection bias can still be present and may render imprecise parameter estimations, they provide a robust methodology to estimate neighborhood effects with observational data. Moreover, multilevel models allow for the analysis of both between and within group variability and the estimation of effects of individual and neighborhood-level factors which are important

advantages for our study. Finally, and especially useful for our study, it allows units to be cross-classified as discussed in depth in section 3.5.

3.4 The current study

As we pointed out in chapter 1 of this thesis, there is a lack of evidence of neighborhood effects on children's development for developing countries. This study presents a quantitative analysis aimed at filling this gap. To do so, we test the existence of neighborhood effects on children's achievement in second grade. As previous evidence indicates that these results may vary according to the cognitive domain analyzed (VINOPAL; MORRISSEY, 2020; DUPERE et al., 2010), the analysis is implemented separately for mathematics and Portuguese language.

This study extends the existing literature by examining the association between neighborhood socioeconomic vulnerability and children's achievement in early elementary school, independent of children's characteristics and family income level, to better understand how where a child grows up affects their academic results and how this association varies by family income. Specifically, we use data from a standardized test, taken by all second graders in São Paulo, linked with corresponding data from the 2010 census to address the following research questions:

- • How does socioeconomic vulnerability of the children's residential neighborhood relate to the children's achievement at second grade? In line with the literature, our hypothesis is that children that were born in more vulnerable neighborhoods have lower scores than their peers that grew up in less vulnerable neighborhoods (VINOPAL; MORRISSEY, 2020; MORRISSEY; VINOPAL, 2018a; MCCOY et al., 2015; DUPERE et al., 2010).
- • Does this relationship vary by cognitive domain? Following the literature, our hypothesis is that there is a difference between the effects found for each cognitive domain, however the evidence is not conclusive about which cognitive domain is more affected by neighborhood vulnerability (VINOPAL; MORRISSEY, 2020; DUPERE et al., 2010).
- Does the access to public center-based care influence the neighborhood effect? Evidence on the role of child care centers as mediator of neighborhood effects are mixed (VINOPAL; MORRISSEY, 2020; MCCOY et al., 2015; DUPERE et al., 2010). But building on the early childhood literature, our hypothesis is that attending center-based care has a positive impact on children's educational outcomes, specially those for more vulnerable backgrounds, and because of that, it could mitigate negative neighborhood effects (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016).

- How does children's family income level influence the relationship between neighborhood socioeconomic vulnerability and their achievement? The literature does not provide a clear indication of the direction nor the significance of such effects (VINOPAL; MORRISSEY, 2020; DUPERE et al., 2010). But building on the literature regarding the impact of family poverty on children's educational outcomes, our hypothesis is that parental income can protect children from negative neighborhood effects.

3.5 Method

3.5.1 Data

The current study draws on multiple data sources. The measures of child cognitive ability, the children's proficiency in Portuguese language and mathematics, were collected from a specific standardized test taken by all second-grade students in municipal schools in 2018, the *Provinha São Paulo*[2]. The data set from this test contains individual results, including the percentage of correct answers and the proficiency on both tests, calculated using Item Response Theory. It also includes student characteristics such as special needs and the school and classroom attended. The data set has information for all the children enrolled in second grade in the municipal schools of São Paulo and it also includes an indicator of test participation as not all children enrolled took the test[3].

To learn the child's residential neighborhood and more individual characteristics, this data set is matched to an administrative dataset from the city's enrollment process as discussed in chapter 2, using a child ID code. From the administrative database, information was collected on the child's residential postal code at the time of registration as well as the child's month and year of birth, their race, gender and information regarding ECEC enrollment. This database also provided the parent's official ID, their *Cadastro de Pessoa Física* number.

To characterize the neighborhoods the children in our sample lived in, data from the 2010 Brazilian Census was used. This study uses a database constructed by Ipea and

[2] This test is part of a larger, city-wide initiative to evaluate students in the municipal school system, re-installed in 2017. This initiative,*Prova São Paulo* that consists of standardized tests for children between second grade and ninth grade in mathematics, Portuguese language and science (for grades 3-9). The *Provinha São Paulo* is a subset of the initiative, focused on children who are enrolled in second grade, and only tests children's proficiency in mathematics and Portuguese language. It was only universally implemented in 2018.

[3] It equals 1 if the child was enrolled and took the test and 0 if the child was enrolled but did not take the test. This second situation happens when the child is enrolled in school but either was not present on the test day or had to leave early and because of that, did not take the test. Therefore, there are children in our sample without valid non-missing mathematics and Portuguese Scores. Of the 46,713 children present in this database, 40.746 took the mathematics test and 40,773 took the Portuguese language one.

was discussed in chapter 2. It contains information on social vulnerability - measured by the IVS - of human development units[4]. The child's residential postal code, collected from the administrative data set, was used to assign each child to a human development unit. The human development unit that the child resided in when they first were registered in the municipal ECEC system was used as a proxy for the neighborhood the child grew up in[5]. Previous research indicates that the child's neighborhood in a given year is a reasonable proxy for the environment he or she is exposed to in childhood (MORRISSEY; VINOPAL, 2018a).

To characterize the schools the children attend, a database from the municipal Department of Education containing the *Índice de Desenvolvimento da Educação Básica* (IDEB) of municipal schools is used[6]. To merge this data to our sample, the official school ID is used.

To characterize the children's family, data on the parental income in the formal sector, on their years of schooling, as well as their gender and race was collected from the *Relação Anual de Informações Sociais*. This database is compiled by the federal government and has information on individual formal work contracts and salaries. To merge this information to our sample, the parents official ID was used.

3.5.2 Sample

In this study, we limited our sample to children who were present in the *Provinha São Paulo* database in 2018, did not have any special needs[7], requested a spot in the city's municipal early childhood educational system between 2010 and 2018[8], had a valid, non-missing residential postal code, located in the city of São Paulo[9] and had non-missing, positive parental hourly salary data[10]. The diagram in Figure 3.1 illustrates the sample

[4] This is a public database available at: http://ivs.ipea.gov.br/index.php/pt/planilha.

[5] For the children in second grade who were found in the administrative database, 60.26% requested a spot more than once: they registered first and then, drop out of the process or did not accept the spot they were offered, and requested a spot again. 74.62% of the first requests happened on or before 2013, when the children on our sample were three years old or younger. In our analysis, we only use the zip-code from the first request, independently of the child's age when it happened or the year it happened.

[6] This is a public database available at: http://dados.prefeitura.sp.gov.br/dataset/ideb-e-prova-brasil-na-rede-municipal-de-ensino.

[7] There were 1213 children with disabilities in the full sample. Among the type of special needs, autism and intellectual disability were the ones with the biggest prevalence.

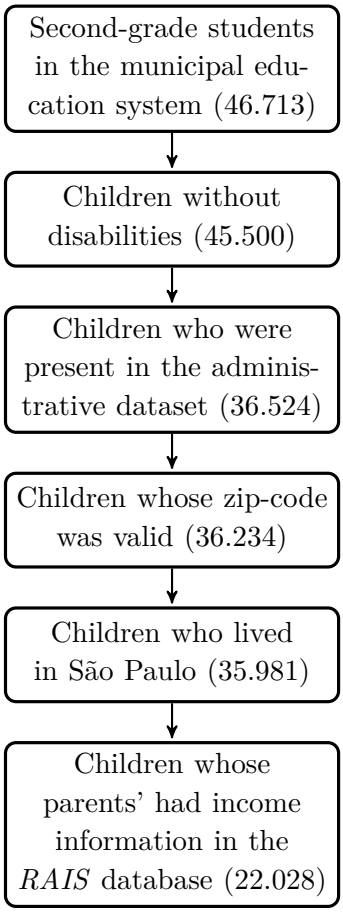
[8] This means we could find the child in the administrative database.

[9] By valid postal code, we mean that we could merge the postal code with the database for the human development units. 253 children were born outside the city of São Paulo, but within the Metropolitan Region. and were removed from our sample. Also, there were three zip-codes that were assigned to two HDUs. We chose randomly which HDU to keep.

[10] This means we could find the parents' official ID at least once in the *Relação Anual de Informações Sociais* database between the years of 2010 to 2017 with a positive hourly salary.

restriction process[11].

Figure 3.1 – Sample Selection



Source: Elaborated by the author.

The sample size varied by dependent variable, from 19,504 for Portuguese language to 19,521 for mathematics[12].

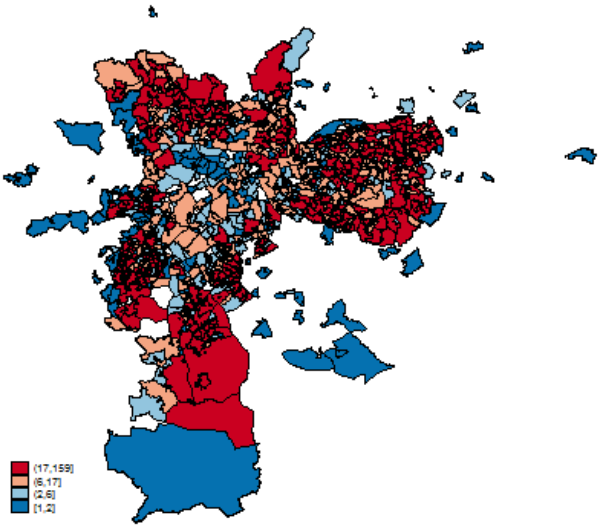
These children were spread across 1,159 HDUs, with a mean of 19.01 children per neighborhood and 553 schools, with a mean of 39.83 children per school. Our sample was not evenly distributed across the city as the map in Figure 3.2 presents and we had more observations outside the center of the city. The city center, a richer area, has a smaller concentration in our sample. This might be because children in this region do not rely on public ECECs or public elementary schools[13]

[11] Our final sample had 242 repeated parents' *official ID*. We could interpret these repeated observations as brothers and sisters.

[12] This difference is a consequence of the fact that there were children present in the *Provinha São Paulo* database who did not take the test. This happened because children were enrolled in Second Grade in 2018 but did not show up on the day of the test or did not take it for some other reason.

[13] There is also 434 HDUs within the city of Sao Paulo that are not part of our sample. No children in our sample lived in those HDUs.

Figure 3.2 – Distribution of Sample by Neighborhood



Source: Elaborated by the author based on a full sample.

3.5.3 Measures

Neighborhood Vulnerability To measure the neighborhood vulnerability, we opted to use the *índice de vulnerabilidade social* (IVS), discussed in previous chapters. As it uses data from the 2010 Census, it describes the neighborhood the child grew up in around the time they were born[14]. It is important to point out that this index considers neighborhood's characteristics of more than one domain of the model proposed by Goldfeld et al. (2015) and discussed in chapter 1. Its human capital, and income and labor dimensions consider characteristics that would be included in the socioeconomic domains and its infrastructure dimension, in the physical domain. Because of this, the results presented here are not directly compared with those systematized in chapter 1.

To analyze the distribution of the sample by vulnerability, we opted to separate the units by quartile of the distribution within the city: the cutoffs were 0.249, 0.313 and 0.368 and they were defined at the HDU level by analyzing the original IVS database[15]. In our sample, 10.07% of the children lived in human development units that were in the first quartile, 26.96% in the second quartile, 39.79% in the third and 23.18% in the fourth

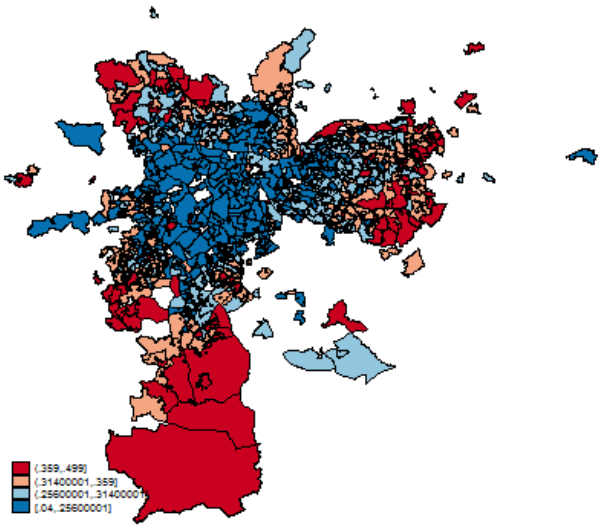
[14] Descriptive statistics indicate that children were around 7.5 years in October of 2018 - the time of the test - and so were born in early 2011.

[15] Following previous research, and considering that São Paulo is less vulnerable than most of regions analyzed - there were no children living in neighborhoods with vulnerability over 0.5 and the maximum IVS is 0.49. -, we opted to not use the classification of Ipea on vulnerability. Ipea has 5 levels of vulnerability: very low vulnerability, unities with index between 0 and 0.2; low vulnerability, unities with index between 0.2 and 0.3; median vulnerability, unities with index between 0.3 and 0.4; high vulnerability, unities with index between 0.4 and 0.5; very high vulnerability, unities with index between 0.5 and 1 (COSTA; MARGUTI, 2015).

quartile[16]. To estimate our model, we opted to use a z-score of the IVS, constructed with the mean and the standard deviation of the index within the city of São Paulo[17].

The spatial distribution of the data can be seen in Figure 3.3. As expected, the areas of higher vulnerability are outside the center of the city, in line with the discussion and the findings presented in the second section of this chapter.

Figure 3.3 – Spatial Distribution of IVS



Source: Elaborated by the author based on a full sample.

Children's achievement Separate measures for children's proficiency in Portuguese language and mathematics are the dependent variables of our study. Children answered 25 questions regarding Portuguese language literacy and 25 questions regarding mathematics literacy[18] in October of the child's second grade. We used the proficiency calculated by Item Response Theory as measures for the children's proficiency.

Covariates Individual, family and school level characteristics were used as covariates. For individual characteristics, the covariates used were child gender, age, race and their exposure to center-based care before age three. To construct this exposure variable, we used the information from the administrative database to indicate if a child was enrolled in the municipal ECEC system before preschool[19]. There was a high percentage of missing race information as 18.4% of the full sample did not have race information. To test the

[16] These descriptive statistics were calculated using the full sample of 22.028 children.

[17] The mean and the standard deviation were estimated in the HDU level with the *IVS* original database and it included all 1,593 HDUs.The mean IVS was 0.296 and the standard deviation, 0.10.

[18] For each subject, there are five questions that require the writing of numbers, letters or words and twenty multiple choice questions. More information can be found in Ordinance n. 7862.

[19] This variable was constructed using the information on the age group the child was first enrolled in the municipal ECEC system. If this age group was part of the daycare center division, the variable

sensitivity of our findings to this missing information, we imputed children's race based on parents' race information from the RAIS database[20].

For parental characteristics, we used parent's hourly income, calculated by averaging the parent's income found in the *RAIS* between 2010 and 2017[21] that correspond to the child's childhood years. We also controlled for the parent gender[22]. In our sensitivity analysis, we also controlled for parent's years of schooling[23] and race[24].

For school level covariates, we use the school's *Índice de Desenvolvimento da Educação Básica* (IDEB), an index from the federal Department of Education that considers the school's approval rate and proficiency results to compute a synthetic index of the school's quality. We use data from 2015 when the children in our sample would not have been in school yet to characterize the quality of the school. It is important to control for such characteristic as different levels of school quality can drive achievement gaps that could be wrongly attributed to neighborhoods (PEARMAN, 2019).

3.5.4 Analytic Plan

To investigate how neighborhood vulnerability may affect children's achievement in second grade (research question 1), we first estimated a cross-classified multilevel model. Cross-classified - or non-nested - multilevel models account for the dependency of the clustered data and model its structure by considering each observation part of two separate two-level hierarchies which cross each other - the school and the neighborhood (LECKIE, 2013). As neighborhoods are not perfectly nested within schools, nor schools

assumed the value 1. If this age group was in part of the preschool division, the variable assumed the value zero.

[20] To do so, we first analyzed the distribution of children by parent's race - divided in two categories, black and non-black. For white parents, 41.86% of their children were non-white and 58.14% were white. For non-white parents, 61.82% of the children were non-white and 38.18%, white. Children were classified as non-white when their race, as reported in the administrative database when they first entered the school system, was defined as black, brown, yellow or indigenous. There were also 4,047 children with missing race that had valid parent race, 1,937 with non-white parents and 2,110 with white parents. We used the proportion of black and non-black children by parent race to correct the missing children's race.

[21] The hourly income for each year was adjusted by inflation, using the *Índice Nacional de Preços ao Consumidor* (INPC) and the final average is on Brazilian *reais* from 2017. For each year, we used information from the parent's December salary and then, combined all the years and calculated a mean for the years between 2010 and 2017.

[22] This variable is equal to 1 if the person who enrolled the children in the system was a man. This person might not be the biological parent and may be the one responsible for the children. For simplicity however, we assume it is the mother or the father. This gender variable, in the *RAIS* database, had some issues: 3.71% of observations changed gender between 2010 and 2017. In this case, as we did not have any other source of information regarding the parent's gender, we assumed that the gender that was indicated the most times, in the years with valid information, was the valid one.

[23] We opted to use an average of parents' years of schooling, reported from 2010 to 2017. Most of the parents - 53.37% - did not change their years of schooling during this time. We did not include this covariate in the original model as this information in the *RAIS* database is self-reported and can, then, have a large measure error.

[24] We did not include this covariate in the full model because of the amount of missing information.

within neighborhoods, an usual nested multilevel model would lead to misleading conclusions (LECKIE, 2013). In our analysis, we consider a third level - families - as our sample has children who have the same parent. Again, this level is not perfectly nested within schools or neighborhoods and because of this, the cross-level structure is recommended[25].

Let Y_{ijkl} denote the outcomes of interest, the proficiency in Portuguese language and mathematics. Then, the estimated model can be described by the following equation:

$$Y_{ijkl} = \beta_0 + \beta_1 X_{ijkl} + \beta_2 T_l + \beta_3 Z_j + \beta_4 W_k + u_j + v_k + r_l + e_{ijkl}$$

$$\begin{aligned} u_j &\sim \mathcal{N}(0,\, \sigma_u^2), \\ v_k &\sim \mathcal{N}(0,\, \sigma_v^2), \\ r_l &\sim \mathcal{N}(0,\, \sigma_r^2), \\ e_{ijkl} &\sim \mathcal{N}(0,\, \sigma_e^2). \end{aligned} \tag{3.4}$$

in which i denotes the individual students, j the neighborhood,k the school the student attends and l the family the child belongs to, X_i, a set of individual characteristics, Z_j the index of socioeconomic vulnerability of the neighborhood the child grew up in, W_j, an index of school quality, and T_l, a set of parents' characteristics, u_j the random effect of neighborhoods, v_k, of schools, and r_l, of families. The error term is e_{ijkl}. We estimate this model for both mathematics and Portuguese language proficiency to answer research question 2.

To answer the third and fourth research questions and test if the effects of neighborhood vulnerability were mediated by center-based care and family income, we included cross-level interactions between neighborhood vulnerability and these covariates:

$$Y_{ijkl} = \beta_0 + \beta_1 X_{ijkl} + \beta_2 T_l + \beta_3 Z_j + \beta_4 W_k + \beta_5 T_l * Z_j + \beta_6 X_{ijkl} * Z_j + u_j + v_k + r_l + e_{ijkl}$$

$$\begin{aligned} u_j &\sim \mathcal{N}(0,\, \sigma_u^2), \\ v_k &\sim \mathcal{N}(0,\, \sigma_v^2), \\ r_l &\sim \mathcal{N}(0,\, \sigma_r^2), \\ e_{ijkl} &\sim \mathcal{N}(0,\, \sigma_e^2). \end{aligned} \tag{3.5}$$

in which β_5 and β_6 represent the parameters of interest as they indicate if the interactions between neighborhood vulnerability and parent's characteristics and between

[25] The structure of the data was verified by analyzing how students were allocated in families, schools and neighborhoods

neighborhood vulnerability and children's characteristics - respectively - are statistically significant. Additionally, to further examine whether the effects of individual and family characteristics vary across neighborhoods, we estimated a random slope model for both mathematics and Portuguese language proficiency. The estimated model can be described below:

$$Y_{ijkl} = \beta_0 + \beta_1 X_{ijkl} + \beta_2 T_l + \beta_3 Z_j + \beta_4 W_k + \beta_5 T_l * Z_j + \beta_6 X_{ijkl} * Z_j + u_{0j} + u_{1j} * T_l + v_k + r_l + e_{ijkl}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0}, \sigma_{u01} \\ \sigma_{u01}, \sigma^2_{u1} \end{pmatrix} \right\} \tag{3.6}$$

$$v_k \sim \mathcal{N}(0,\, \sigma^2_v),$$
$$r_l \sim \mathcal{N}(0,\, \sigma^2_r),$$
$$e_{ijkl} \sim \mathcal{N}(0,\, \sigma^2_e).$$

The random effects now assume a bivariate normal distribution with zero mean and unstructured 2×2 covariance matrices. Note that we make only parental characteristics random across units of neighborhood. We compared different model specifications through likelihood-ratio tests to reach this specific specification[26].

3.6 Results

3.6.1 Descriptive Results

Sample descriptive statistics are provided in Table 3.1[27]. The first section of this table presents the children's characteristics. There is an equal share of boys and girls and 48% of the children in the sample are non-white[28]. The children are around 7.6 years old, the expected age for October of their second grade school year[29], and more than half of them, 58.06%, attended at least one year of public center-based care[30]. Children's mean

[26] This was implemented with the command *ranova* in R which reduces or, when not possible, removes the random effect terms and compares the model reductions via a likelihood ratio test.

[27] All the results presented here are for a sub-sample of children with non-missing mathematics scores. A total of 2,507 children had missing mathematics scores; therefore, the sub-sample presented here is of 19,521 children.

[28] As discussed previously, there is a high number of children without race information; 31.6% in this sub-sample. In our sensitivity analysis, we impute children's race based on parent's race. These results are presented in the sensitivity analysis. Being classified as being non-white means that children's race, at the moment of registration by the parent into the system, was classified as either black, brown, yellow, or of indigenous background.

[29] In Brazil, to be in second grade in a given year, children should turn 7 before March 31st of that year. The standardized test occurred in October of that school year.

[30] In footnote 22, we detail how this variable was constructed.

proficiency was 157.99 in mathematics with a standard deviation of 42.65 and 159.06 in Portuguese language with a standard deviation of 37.85[31].

The second section of Table 3.1 presents parental characteristics[32]. There were 19,297 unique parental observations[33]. The parent's hourly salary was 7.56 Brazilian in 2017 *reais*[34] and had a mean of 6.42 years of schooling. Our data is comprised mostly of mothers as only 13% of the parents in our sample were male.

The third section of Table 3.1 has data on the child's elementary school. In the database with valid mathematics scores, children attended 553 different elementary schools and the median *IDEB* for the year of 2015 was 5.82[35]. For comparison, the *IDEB* for the city's municipal elementary schools was also 5.8 in 2015[36].

The final section of Table 3.1 has data on the vulnerability of the child's neighborhood in 2010. The table presents the z-score of the IVS, both the complete index as well as for each sub-dimension[37]. In the database with only valid mathematics scores, children lived in 1,147 different HDUs in the city of Sao Paulo. The z-scores indicate that the HDUs present in our sample had higher vulnerability in the infrastructure domain than the city as a whole but less vulnerability in the human capital and the income and labor dimensions than the city of São Paulo.

[31] There were 1,325 children with valid mathematics proficiency, but missing Portuguese language proficiency.

[32] We assume that the adult that registered the child into the system, and so shared their identification number, is the child's parent. This can be untrue: children might not have live with their parents or be orphans, for example. The legislation defines that a child should be registered by the adult legally responsible for them.

[33] Parents are identified by their national Identification Number. We are assuming that children with parents' with the same identification number are members of the same family and because of that, in our models, we control for parent's ID.

[34] This translates to a monthly wage of 1,204.80 in 2017 Brazilian reais considering a 40 hour work week. The Brazilian minimum wage was 937 *reais* in 2017.

[35] Six schools did not have a valid IDEB for 2015, therefore, there are 547 school represented in the table.

[36] Data from: https://qedu.org.br/cidade/2329-sao-paulo/ideb?dependence=3grade=1edition=2015.

[37] The z-score for the index and for the sub-indexes were calculated using the mean and the standard deviation calculated at the HDU level for the whole city of São Paulo.

Table 3.1 – Descriptives

	Mean	SD	Obs.
Student			
Female	0.50		19521
Non-white	0.48		13354
Age	7.57	0.51	19521
Exposure to center-based care before age 3	0.58		19521
Mathematics Results	157.99	42.65	19521
Portuguese Results	159.06	37.85	18196
Parents			
Parent's Hourly Income	7.56	6.39	19297
Parent's Yrs. of Schooling	6.42	1.15	19297
Parent's Gender	0.13	0.33	19297
School			
IDEB 2015	5.82	0.43	547
Neighborhood			
Z-Score - Vulnerability Index	-0.00	0.95	1147
Z-Score - Vulnerability Index - Infrastructure Dimension	0.01	0.98	1147
Z-Score - Vulnerability Index - Human Capital Dimension	-0.01	0.93	1147
Z-Score - Vulnerability Index - Income and Labor Dimension	-0.01	0.93	1147

Source: Elaborated by the author based on a sub-sample of children with non-missing math scores.

Table 3.2 presents the descriptive analysis of the relationship between vulnerability and proficiency. For both Portuguese language and mathematics, the table presents the mean score and the standard deviation in brackets by quartile of neighborhood vulnerability. As discussed previously, the quartiles were defined according to the distribution of the index in the city of São Paulo. Children in higher vulnerability neighborhoods had lower scores in both Portuguese language and mathematics. The slopes of both scores by neighborhood vulnerability are similar; if we compare children who live in neighborhoods in the top three levels of vulnerability there is a clear stepwise pattern, as neighborhood vulnerability is negatively associated with the children's score. The exception is the first level of vulnerability. When those in the 25% less vulnerable neighborhoods are compared to those in the second level of vulnerability, the relationship between neighborhood vulnerability and cognitive results is the positive: children in the median-low vulnerability neighborhoods scored higher than children in the low vulnerability neighborhoods.

Table 3.2 – Vulnerability and Proficiency

	Low Vulnerability	Median-low Vulnerability	Median-high Vulnerability	High Vulnerability
Portuguese Results	160.1	161.1	158.9	156.6
	(36.65)	(37.65)	(37.81)	(38.52)
Mathematics Results	159.3	160.7	158.3	156.3
	(41.72)	(42.33)	(42.90)	(42.70)

Source: Elaborated by the author based on a sub-sample of children with non-missing math scores.

Table 3.3 describes the relationship between neighborhood vulnerability and two indicators of neighborhood resources: exposure to center-based care before age three and school quality. For each of the four quartiles of neighborhood vulnerability, the mean value is presented along with the standard deviation in brackets. There is no clear relationship between the enrollment in the municipal center-based care and the vulnerability of the neighborhood. It is worth noting, however, that children living in median-low vulnerability neighborhoods attended center-based care in a higher proportion than their peers in any other type of neighborhoods. When school quality is analyzed, there is a clear negative relationship with neighborhood vulnerability; children who live in more vulnerable neighborhoods are exposed to worse quality schools.

Table 3.3 – Vulnerability, center-based care and school quality

	Low Vulnerability	Median-low Vulnerability	Median-high Vulnerability	High Vulnerability
Exposure to center-based care before age 3	0.571	0.592	0.577	0.577
	(0.495)	(0.492)	(0.494)	(0.494)
IDEB15	6.004	5.895	5.766	5.694
	(0.479)	(0.419)	(0.379)	(0.382)

Source: Elaborated by the author based on a sub-sample of children with non-missing math scores.

Figure 3.4 and Figure 3.5 describe the spatial distribution of both proficiencies in the city of São Paulo. For each HDU, the mean score for both mathematics and Portuguese language was calculated. There is no clear spatial pattern for the distribution of the median score. However, it is important to point out that the distribution of the sample is not uniform throughout the city nor by HDU and therefore the maps might be hiding some kind of spatial concentration.

Figure 3.4 – Spatial Distribution of Mathematics Proficiency

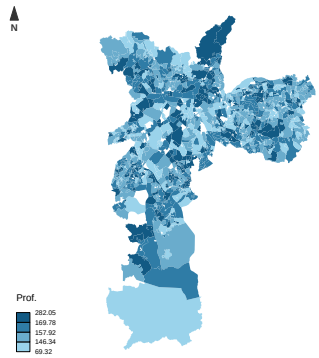
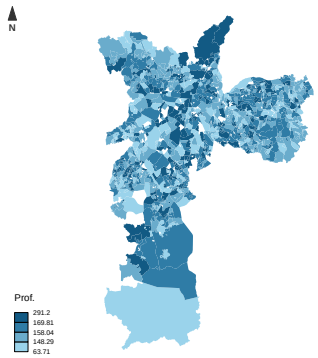


Figure 3.5 – Spatial Distribution of Portuguese Language Proficiency



Source: Elaborated by the author based on the sample without non-missing scores.

3.6.2 Regression Results

Multilevel Models Table 3.4 presents the results for the multilevel models, indicating the point estimates, the standard errors - in parenthesis - and super-scripted stars with the significance of each coefficient. Model 1 is the one described by Equation 3.3. Its results are presented for Portuguese language proficiency (1.1) and mathematics proficiency (1.2). Model 2 incorporates interaction terms between neighborhood vulnerability and child and parents characteristics to examine interactive effects. Again, results are presented for Portuguese language proficiency (2.1) and mathematics proficiency (2.2). All of these models were fitted through restricted maximum likelihood[38].

Both types of models fit the data significantly better than a one-level model or a two-level model taking into account only the neighborhood[39]. For models 1.2 and 2.2, likelihood tests indicated that adding a random slope to the model allowing the effect of parents' hourly income to vary between neighborhoods - to the model improved it[40].

[38] The estimations were done with the command *lmer* in R. To estimate p-values, the package *lmerTest* was used.

[39] Likelihood ratio tests were run between a one-level model without independent variables and a model with two levels, individual and neighborhood, without any independent variables. For both dependent variables, the results indicated that the neighborhood level improved the fit of the data. Subsequently, another likelihood test was implemented between the model using only the neighborhood level and a model with four levels: neighborhood, school, family and individual. Again, for both dependent variables, the results indicated that the second model fit the data better. As between-group differences may be revealed after covariates are added, the tests were re-run with the full models. The results hold; the null hypothesis is rejected for all four final models, 1.1, 2.1, 1.2 and 2.1, indicating that the multi-level is preferred both for a single level model and a two-level model.

[40] This was tested with the command *ranova*. For Portuguese language, this was implemented by the test indicated that it did not improve the model.

In addition, for all four models, the residuals were analyzed; residual plots were constructed to check for both homoscedasticity and normality. The graphs are in the Appendix, Figure 3.6 and Figure 3.7. An analysis of these indicates that although the residuals have mean zero and are approximately normally distributed, there is some evidence of the existence of heteroscedasticity.

At the bottom of Table 3.4, there is information on the models' fit; the Akaike information criterion (AIC) as well as the number of observations, both in the full sample and by group (of families, of neighborhoods and of schools). Models 1.1 and 2.1 were run with the sample of non-missing mathematics proficiency; models 1.2 and 2.2 used the sample with non-missing Portuguese language proficiency. Due to missing data in both race and the school quality index (IDEB15), the number of observations used in the estimation is smaller than the ones discussed previously[41]. The last six lines in table 1.4 provide information on the estimated random effects. As models 1.2 and 2.2 had a random slope added to them, there is extra information in the last two lines of the Table 3.4.

For Model 1, there is no statistically significant effect of growing up in a more vulnerable neighborhood. The coefficient for the z-score has the expected signal, negative, indicating that living in a more vulnerable neighborhood has a negative effect on proficiency in both mathematics and Portuguese language but it is not statistically significant for either dependent variable. Children's sex has a significant effect for both mathematics and Portuguese language proficiency, although for Portuguese language, the effect is much larger in size. Children's race and age also have significant effects; being non-white and younger, reduces proficiency in both mathematics and Portuguese language. For both covariates, the effects are superior when the dependent variable is mathematics proficiency. For exposure to center-based care before age three, there is a positive statistically significant effect only when the proficiency in mathematics is analyzed but this effect is small, 4.83% of a standard deviation. School quality only had a positive and statistically significant effect when Portuguese language was analyzed. Parent's hourly income, log-transformed, has a positive, statistically significant and large effect at around 30% of a standard deviation for both dependent variables.

For Model 2, the point also estimates the significance level of most independent variables as being quite similar. However, there is a negative and statistically significant effect for neighborhood vulnerability when the dependent variable is Portuguese language; there is a 5.52 reduction corresponding to 14.45% of a standard deviation, in Portuguese language proficiency for children who grew up in neighborhoods with a vulnerability index

[41] For the sample for models .1, from the total of 19,521 children with non-missing scores, 6,090 had missing race information, 163 were missing school quality information and 77 had both pieces of information missing. For the sample for models .2, from the total of 19,504, 6,149 had missing race information, 159 were missing school quality information and 76 had both pieces of information missing.

one standard deviation above the city's mean. The interaction term between parent's income and the IVS is statistically significant for Model 2.1 for children who grew up in neighborhoods with an IVS one standard deviation above the city's mean. The effect of parent's income on children's Portuguese language proficiency is even larger at 15.52, 40% of a standard deviation. The other interaction terms do not have a statistically significant result indicating that the neighborhood effect is not affected by children's characteristics. For mathematics, none of the interactions have statistically significant coefficients.

Table 3.4 – Statistical models

	Model 1.1	Model 2.1	Model 1.2	Model 2.2
(Intercept)	45.56***	45.97***	57.15***	57.07***
	(10.23)	(10.23)	(12.20)	(12.21)
Female	10.89***	10.73***	1.21·	1.16
	(0.62)	(0.65)	(0.68)	(0.71)
Non-white	−3.55***	−3.74***	−4.73***	−4.58***
	(0.64)	(0.67)	(0.70)	(0.73)
Exposure to center-based care before age 3	0.88	0.97	2.07**	2.06**
	(0.64)	(0.67)	(0.71)	(0.74)
Age	7.88***	7.89***	8.73***	8.73***
	(0.61)	(0.61)	(0.67)	(0.67)
Parent's Hourly Income (ln)	13.55***	13.33***	13.86***	13.76***
	(0.85)	(0.85)	(1.00)	(1.01)
Parent's Gender	−4.53***	−4.58***	−4.56***	−4.59***
	(0.96)	(0.96)	(1.06)	(1.06)
IDEB15	4.08**	4.10**	1.57	1.62
	(1.55)	(1.54)	(1.88)	(1.88)
Z-score of IVS	−0.75	−5.52*	−0.12	−2.66
	(0.54)	(2.32)	(0.58)	(2.65)
Parent's Hourly Income (ln)*Z-score of IVS		2.19*		1.39
		(1.05)		(1.22)
Exposure to center-based care*Z-score of IVS		−0.42		0.06
		(0.91)		(1.00)
Female*Z-score of IVS		0.76		0.30
		(0.89)		(0.97)
Non-white*Z-score of IVS		0.93		−0.70
		(0.91)		(1.00)
AIC	131217.77	131213.12	134678.70	134676.93
Num. obs.	13120	13120	13191	13191
Num. groups: families	12982	12982	13040	13040
Num. groups: neighborhoods	1096	1096	1095	1095
Num. groups: schools	547	547	547	547
Var: families (Intercept)	491.72	487.37	531.50	531.08
Var: neighborhoods (Intercept)	4.96	5.04	230.80	236.27
Var: schools (Intercept)	170.72	170.11	271.03	270.82
Var: Residual	724.85	729.07	949.48	950.03
Var: neighborhood, Parent's Hourly Income (ln)			58.67	60.26
Cov: neighborhood (Intercept), Parent's Hourly Income (ln)			−115.89	−118.88

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Source: Elaborated by the author based on a sub-sample of children with non-missing math scores.

The random effects are presented at the bottom of Table 3.4 and also in Figure 3.8, Figure 3.9, Figure 3.10 and Figure 3.11. For each model, the between-group variance is presented - for families, neighborhoods and schools - as well as the residual variance, which that is not explained by the grouping structure. The family's variance component, σ_u^2, measures the differences between neighborhoods, being adjusted for differences between

families and schools. The other variance components have similar interpretations.

The results suggest that both mathematics proficiency and Portuguese language proficiency vary between families, schools and neighborhoods, independent of individual characteristics. However, this variation is larger between schools and families, than between neighborhoods. This is especially true for Portuguese language: the value of σ_u^2 is very close to zero, indicating that there is almost no variation between neighborhoods, above and beyond differences in schools and families.

To interpret the relative magnitude of the variance components it is useful to compute variance partition coefficients that report the proportion of the variance of the outcome that lies in each level of the grouping structure. Table 3.5 presents the results and indicates only 4% of the variation in Portuguese language lies between neighborhoods; 35.3% lies between families and 12.3%, between schools. As this analysis is not valid for those models that include a random coefficient - as do models 1.2 and 2.1 -, Table 3.5 only has the results for the models with Portuguese language as the dependent variable.

Table 3.5 – Variance Partition Coefficients for Models 1.1 and 2.1

Group	Model 1.1	Model 2.1
School	0.123	0.122
Families	0.353	0.350
Neighborhood	0.004	0.004

Source: Elaborated by the author based on a sub-sample of children with non-missing math scores.

Another way to interpret the relative magnitude of the variance components is to calculate intraclass correlation coefficients (ICCs). For Model 1.1, the ICC is 47.9% of the variance and for model 2.1, 46.1%. This indicates that 47.9% of the variation in Portuguese language proficiency was attributable to differences between neighborhoods, schools and families in Model 1.1 and when interactions were added to the model, this proportion was reduced to 46.1% of the total variation.

In models 2.1 and 2.2, we allowed parents income to have a random coefficient at the neighborhood level; that is, we allowed the relationship between outcome, proficiency in mathematics and parent's income to vary across neighborhoods[42]. Because of this, the interpretation of their random effect results is a little different. At the bottom of table 1.4, for models 1.2 and 2.1, we have the value of σ_{u0}^2 and of σ_{u01}. The neighborhood slope variance indicates the mean variation, between neighborhoods, in the relationship between parent's hourly income and proficiency in mathematics. From the σ_{u01}, we can

[42] We tested if random coefficients for all the covariates improved the model, by using the command *ranova* that applies such test with likelihood tests; only for parent's income were there any positive results.

calculate the correlation between the neighborhood intercepts and slopes: it is -0.996. Its negative value indicates that neighborhoods with above average intercepts tend to have below average slopes or in other words, the relationship between parent's income and proficiency in mathematics tends to be less positive in neighborhoods with a high proficiency in mathematics.

Multilevel Models for IVS dimensions To understand if the effect of neighborhood vulnerability on proficiency varies according the type of vulnerability, we estimated the model for each one of the three dimensions of the IVS separately. Table 3.6 presents the results for model 1 and Table 3.7 in the Appendix, for model 2.

For model 1, the point estimates, as well as the statistical significance, are very similar to the results presented in Table 3.6 as expected, for both dependent variables. However, for the vulnerability index, negative and statistically significant effects are found for the human capital dimension in both Portuguese language and mathematics proficiency and for the infrastructure dimension for mathematics. For the human capital dimension, these results indicate that children who live in neighborhoods with vulnerability one standard deviation above the city's mean have lower scores in both mathematics and Portuguese language although the size of such effects are not very substantial; 4.8% of an standard deviation for Portuguese language and 4.1% of a standard deviation for mathematics. The effect for the infrastructure dimension of the vulnerability index is counterintuitive as it indicates that children who grow up in neighborhoods with higher infrastructure vulnerability have higher proficiency in mathematics although again, the size of the effect is small, 3.1% of a standard deviation.

The random effects structure is similar to the one discussed for the full vulnerability index. Again, for Portuguese language there is less variability between neighborhoods then for mathematics.

For model 2, estimations were done separately for each of the IVS dimensions. The results are presented in Table 3.7. The point estimates, and the significance levels, are pretty similar to those presented in Table 3.4. There are negative, and statistically significant, effects for all three dimensions of vulnerability when the dependent variable is Portuguese language. The interaction term between parent's income and the vulnerability is significant for two of the three dimensions - infrastructure and human capital. The size of the effects is also pretty similar to the one found in Table 3.4, when the full index was analyzed. When the dependent variable is mathematics, a statistically significant effect is found only for the income dimension. The interaction term was not significant for either dimension. The counterintuitive result found for Model 1 is not replicated here as even though the coefficient for the infra-structure index was positive, it is not statistically significant. Again, the results for the random structure are similar to those discussed

previously.

Table 3.6 – Statistical models

	Model 1.1	Model 1.2
(Intercept)	46.47***	58.41***
	(10.24)	(12.22)
Female	10.90***	1.23·
	(0.62)	(0.68)
Non-white	−3.52***	−4.70***
	(0.64)	(0.70)
Exposure to center-based care before age 3	0.86	2.06**
	(0.64)	(0.71)
Age	7.89***	8.76***
	(0.61)	(0.67)
Parent's Hourly Income (ln)	13.54***	13.84***
	(0.85)	(1.00)
Parent's Gender	−4.51***	−4.54***
	(0.96)	(1.06)
IDEB15	3.89*	1.29
	(1.55)	(1.88)
Z-Score - Vulnerability Index - Infrastructure Dimension	0.50	1.34*
	(0.61)	(0.66)
Z-Score - Vulnerability Index - Income and Labor Dimension	0.63	0.50
	(0.89)	(0.97)
Z-Score - Vulnerability Index - Human Capital Dimension	−1.83*	−1.76·
	(0.83)	(0.91)
AIC	131214.75	134672.78
Num. obs.	13120	13191
Num. groups: families	12982	13040
Num. groups: neighborhoods	1096	1095
Num. groups: schools	547	547
Var: families (Intercept)	490.57	531.50
Var: neighborhoods (Intercept)	4.78	231.12
Var: schools (Intercept)	171.00	272.08
Var: Residual	725.79	949.07
Var: neighborhood, Parent's Hourly Income (ln)		58.40
Cov: neighborhood, Parent's Hourly Income (ln)		−115.85

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Source: Based on a sub-sample of children with non-missing math scores.

3.6.3 Sensitivity Analysis

To test the sensitivity of our results to different samples and model specifications, we conducted five sets of analyses. All tables can be found in the Appendix.

First, we re-estimated the models with more complete data including parent's education and race. As there was a lot of missing information for parent's race[43], the

[43] For the full sample, 41.6% of the parents had missing race information.

sample size for this analysis was much smaller[44]. The results are presented in Table 3.8. The point estimates of most covariates, as well as their significance levels, are very similar to those presented in Table 3.4. Parent's education measured by years of schooling is significant in all models, but its effect size is much smaller than the effect of parent's income at around one third. Parent's race is statistically significant in the models where the dependent variable is proficiency in mathematics; its effect size is smaller than that of children's race[45]. The main difference is the size of the effect of the IVS in the more complete model with Portuguese language proficiency as the dependent variable where it becomes 26% of a standard deviation of the proficiency in Portuguese language. The random structure does not fit these models as well as it does the larger sample where we can see that the variance for the neighborhood is zero for the first three models indicating that proficiency did not vary between neighborhoods[46].

Using information on the parent's race, we imputed the children's missing race information[47]. This expanded the sample size significantly[48]. The results for this test can be found in Table 3.9. The point estimates are similar to those in Table 3.4 and indicate that the negative effect of the neighborhood vulnerability on children's Portuguese proficiency, in model 2.1, is robust to this imputation. One difference, however, should be pointed out. Considering this sample, we find statistically significant and positive effects to attending center-based care for all four models.

We also re-estimated the models using the sample before it was merged with the *RAIS* database. This sample is much larger in size but it does not have information regarding parents characteristics[49]. The results are in Table 3.10, We find a negative and statistically significant effect for the IVS in both models and for both dependent variables, indicating that, when we do not control for parent's income and other characteristics, there is a negative effect of growing up in a more vulnerable neighborhood. This result is in line with the literature and can be explained by the fact that, when we do not control for such covariates, we increased the likelihood that the found association is a consequence

[44] When the dependent variable is the mathematics proficiency, the sample size goes from 13,191 to 7,769.

[45] It is important to point out that the variable *Parent's Race* equals one when the parents are white.Therefore, even though the signal of the coefficients is opposite, the interpretation of both race effects, on the child and on the parent, is similar: being white has a positive association with the child's proficiency.

[46] It is worth mentioning also that we could not calculate ICC for models 3.1, 4.1 and 4.2, indicating problems with the random structure. However, because the objective here was to test the sensitivity of the results, we chose to keep the same model specification already used.

[47] We used the distribution of non-white and white children to non-white and white parents to do this imputation. For children with missing race, and whose parents also had this information missing, we could not do such imputation. Therefore, there were still 2,563 children with missing race information.

[48] When the dependent variable is mathematics proficiency, the sample size goes from 13,191 to 16,744.

[49] When the dependent variable is mathematics proficiency, the sample size goes from 13,191 to 21,050. As discussed previously, there was a lot of missing children's race information and therefore, the effective sample is smaller than the number presented in Figure 3.1.

of unobserved co-founders (VINOPAL; MORRISSEY, 2020). It is interesting to point out that, unlike the previous test, there is only a statistic significant effect for attending center-based care when the dependent variable is mathematics proficiency.

Our analyses relied on a strong hypothesis; that the neighborhood the children lived in when they first entered the municipal ECEC system corresponded to the neighborhood the child grew up in. However, children could have moved over the years; one or several times. To test this hypothesis, we analyze if the results change when we analyze non-movers. To do so, we define two sub-samples based on the notion that as schools should be near the child's home, we can assume that children did not move if their school is near its original residential postal code. The first sub-sample is composed of only children whose residential postal code was in the same district[50] as their school and the second is composed of only children whose postal code was in the same HDU as their school. The results are presented in Table 3.11. As expected the sample sizes are much smaller[51]. For the first sub-sample, the effects of the IVS on proficiency are not statistically significant in any of the models; the point estimates, however, are similar to the ones in Table 3.4. For this sub-sample, the effect of attending center-based care is statistically significant in three of the four models. For the second sub-sample, a puzzling result is found: the effect of IVS is positive for three of the four models, even though in neither model is statistically significant. For model 1.2, the effect is statistically significant and much larger than estimated before. When the effect of center-based care exposure before age 3 is analyzed, the results are similar to those presented in Table 3.4: there is a significant and positive impact when the dependent variable is the mathematics proficiency, but not when the proficiency in Portuguese language is analyzed.

Finally, using the same sample presented in Table 3.4, we re-estimated both models using the parent's maximum hourly income and the parent's minimum hourly income[52]. These results are presented in Table 3.12. For models 1.1 and 1.2, the results for the effect of neighborhood vulnerability are statistically significant for only for the minimum hourly salary. However, unlike the results presented in Table 3.4, in model 2.1, the effect of vulnerability, not the interaction term with parent's income, is statistically significant with either maximum or minimum hourly salary. When the dependent variable is the proficiency in mathematics, no statistically significant effect is found. Moreover, for these models, exposure to center-based care before age three is only significant when proficiency in mathematics is analyzed.

[50] Districts are an administrative division of the city of São Paulo and the HDUs can be aggregated into districts.

[51] When the dependent variable is the mathematics proficiency, the sample size goes from 13,191 to 7,483 in the first sample and 4,804 in the second sample.

[52] To find this maximum and minimum value, we analyzed the hourly income for each year between 2010 and 2017. It is important to point out that we did not consider incomes of zero and, within each year, we considered the maximum income level.

Taken together, the sensitivity tests mostly confirm the existence of neighborhood effects for Portuguese language proficiency. More research is needed to understand, however, the results for the exposure to center-based care before age 3, as well as to better understand the results presented in Table 3.11.

3.7 Discussion

The literature provides evidence of the existence of neighborhood effects in children's development, especially when the socioeconomic vulnerability of the neighborhood is analyzed and this thesis's chapter 1 has a systematic review of such findings. This study aims to fulfill the gap in the evidence for developing countries as it examines the association between neighborhood vulnerability and children's achievement in early elementary school in São Paulo. To do so, we estimated mixed models, controlling for the grouping structure of the data where children belong to families, to schools and to neighborhoods. By doing so, we are accounting for the fact that children who live in the same neighborhood or are enrolled in the same school, or belong to the same family, have shared backgrounds or experiences that might affect the results.

Our results indicate that there is an association between neighborhood vulnerability and student achievement in second grade. Growing up in a more vulnerable neighborhood is associated with lower proficiency in Portuguese language but not in mathematics. After accounting for child and family background characteristics and considering the interaction between those characteristics and neighborhood vulnerability, the effect size is 14.5% of a standard deviation of the children's proficiency in Portuguese language. The effect is smaller than the effect of parent's income or children's gender but similar to that of children's race.

These results corroborate the evidence found in the international literature of the existence of neighborhood effects on children's outcomes. The point estimates are not comparable to that of the international literature as most of it only analyzed socioeconomic vulnerability and the index used in this study incorporated more dimensions. It is worth nothing however, that effects on children's mathematical achievement and not language achievement, are generally found in the literature (PEARMAN, 2019).

Additionally, there is evidence of variation in children's proficiency between neighborhoods above and beyond differences in schools and families. In line with international evidence, we find that neighborhood vulnerability explains 4% of the variance in Portuguese language proficiency (LEVENTHAL; FAUTH; BROOKS-GUNN, 2005).

Our findings also indicate that neighborhood vulnerability and parent's income interact in a significant way. growing up in a more vulnerable neighborhood increases the effect of parents' income on children's proficiency in Portuguese Language in 16.4%.

This can be interpreted as a protective effect of parent's income in children's educational outcomes. Also, for the proficiency in mathematics, there is evidence that the relationship between parent's income and children's proficiency is more negative in neighborhoods with higher levels of children's proficiency. That is, in neighborhoods where children have better results, parental income matters more.

When the IVS is decomposed into its three dimensions, we find statistically significant results for two dimensions, human capital and infrastructure. Growing up in neighborhoods with worse human capital; higher child mortality, more children out of schools, less educated populations and more teenaged mothers, is associated with worse achievement in second grade for both cognitive domains. The results for the infrastructure domain are counterintuitive; growing up in a neighborhood with worse infrastructure has a positive association with proficiency in mathematics. Even though decomposing the IVS gives us clues to the aspects of the neighborhood that may influence the associations found in the study, more detailed exploration is needed to understand the mechanisms behind neighborhood effects.

Our findings also indicate that there is an association between attending center-based care and children's proficiency in second grade. Our estimations indicate that there is a significant statistical effect when proficiency in mathematics is analyzed however, this effect is quite small, 4.8% of a standard deviation. The size of the effect might be a consequence of the low quality of daycare system which in turn, can also be correlated with neighborhood vulnerability. There is also no evidence that center-based care mediated the neighborhood effect. These results are robust to different sample and model specifications. In the chapter 5, the possible impact of center-based care in cognitive outcomes is examined at length.

Taken together, these results indicate that growing up in a more vulnerable neighborhood, especially one with a less educated and more vulnerable population, can have both direct and indirect effects through increasing the effect of parental income on children's educational outcomes. These effects are however, were found for only one cognitive domain, Portuguese language. Our findings also indicate that attending center-based care has an effect on children's achievement in mathematics above and beyond family and neighborhood effects.

3.8 Limitations and Future Directions

Our analyses have several limitations. First, although multilevel models are the most recommended to deal with our cross-level data structure, and although we controlled for a myriad of family and individual characteristics that the affect selection process in the neighborhood, it is possible that other confounding factors - that were not controlled

for or that are non-observable - could have explained the results. Any claims of causality of the results - or interpretation as such - should be made with caution. From another standpoint, however, our choice to control for such family and individual characteristics might underestimate neighborhood effects as there are "indirect effects of neighborhoods that operate through time-varying family characteristics" - like family income for example (WODTKE; HARDING; ELWERT, 2011; WOLF; MAGNUSON; KIMBRO, 2017). To correctly solve the selection bias problem, an exogenous source of variation in the neighborhood would be necessary.

Second, this study examined a specific sample of children from the city of São Paulo. The results cannot be generalized to Brazil or Latin America. Not only is São Paulo less vulnerable than other metropolitan regions in Brazil and Latin America, but it might also have other specificities that make it impossible to compare with other contexts. Additionally, our results cannot be generalized for all second-grade students in the city of São Paulo as we analyzed only children who are enrolled in municipal schools corresponding to 30% of the total number of second graders in the city[53]. Finally, our sample restriction process means that we cannot even generalize our results for the cohort of children enrolled in municipal schools in second grade in 2018 because, as discussed previously, our final sample was statistically different from this full cohort of children.

Third, children might have moved during their childhood years and therefore, we might be allocating children to the wrong neighborhood. This incorrect allocation might also be caused by an incorrect report of residential postal codes at the moment of registration into the system although this is unlikely as parents need to provide proof of residence. The literature also indicates that even if children moved a lot, a child's neighborhood in a given year is "a reasonable proxy for his or her long-run environment" (MORRISSEY; VINOPAL, 2018a).

Fourth, we opted to use the HDU, an administrative division of the city, to define a child's residential neighborhood. However, the HDU might not necessarily correspond to the neighborhood the children and their family experience in their day-to-day lives. This is called the Modifiable Area Unit Problem (MAUP) and can affect results. Our findings can vary according to the scale of the areas and the interpretation of the results is only valid for the analyzed division of space (FLOCH, 2018). Moreover, our model relies on predefined neighborhood frontiers that may not correctly reflect the individual experiences and also it does not model the possible influence of individual neighbors on each other's outcomes (ZANGGER, 2019). As stated by Subramanian (2004), "identifying "true" neighborhood differences also requires identifying "true" neighborhoods". When researchers use predefined neighborhoods divided by administrative reasons or other exogenous factors, one

[53] Data from *Censo Escolar*, extracted from www.qedu.com.br. In total, in 2018, there were 153,981 second-grade students in São Paulo.

assumption is that people within such neighborhoods are experiencing the same set of neighborhood conditions and characteristics (COULTON; KORBIN; SU, 1999) - which might be a faulty assumption.

Fifth, we opted to use an index of vulnerability that aggregated different dimensions and therefore our results are not comparable with those from the international literature. Moreover, by using such measure, we are assuming that this index has a clear connection with the mechanisms that generate neighborhood effects, and this could be questioned (HARDING et al., 2010).

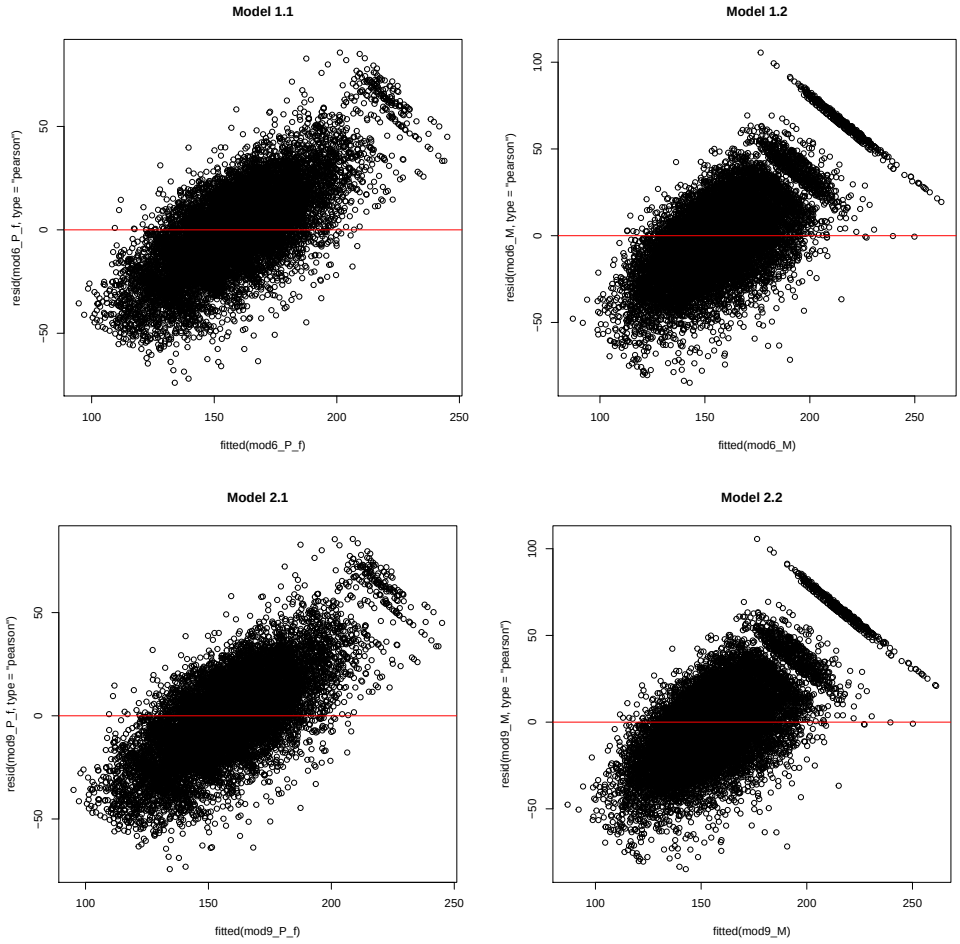
Even with these limitations, we understand that our results provide valuable evidence that indicates that the vulnerability of a neighborhood is related to a child's educational outcomes in late childhood and that there is an achievement gap caused by neighborhood vulnerability. These findings indicate that place-based policies might be useful in promoting better educational outcomes especially if they focus on children from more vulnerable neighborhoods. They also indicate that is important to intervene in neighborhoods to mitigate the achievement gap found in the early elementary school years.

Our findings also highlight the importance of advancing the methodological discussion on the estimation of neighborhood effects where more robust techniques that can have causal claims, are needed. It is important to understand if neighborhood vulnerability indeed has an impact, a causal effect, on children's outcomes. If it does, then place-based policies and other interventions that would reduce neighborhood vulnerability also became important education policies.

Finally, there is a need to research the mechanisms underlying the association between neighborhood vulnerability and student achievement and also the aspects of neighborhood vulnerability that generate neighborhood effects. Understanding such mechanisms and decomposing which aspects of neighborhood vulnerability truly matter can provide further insights as to where interventions should be aimed to improve children's outcomes.

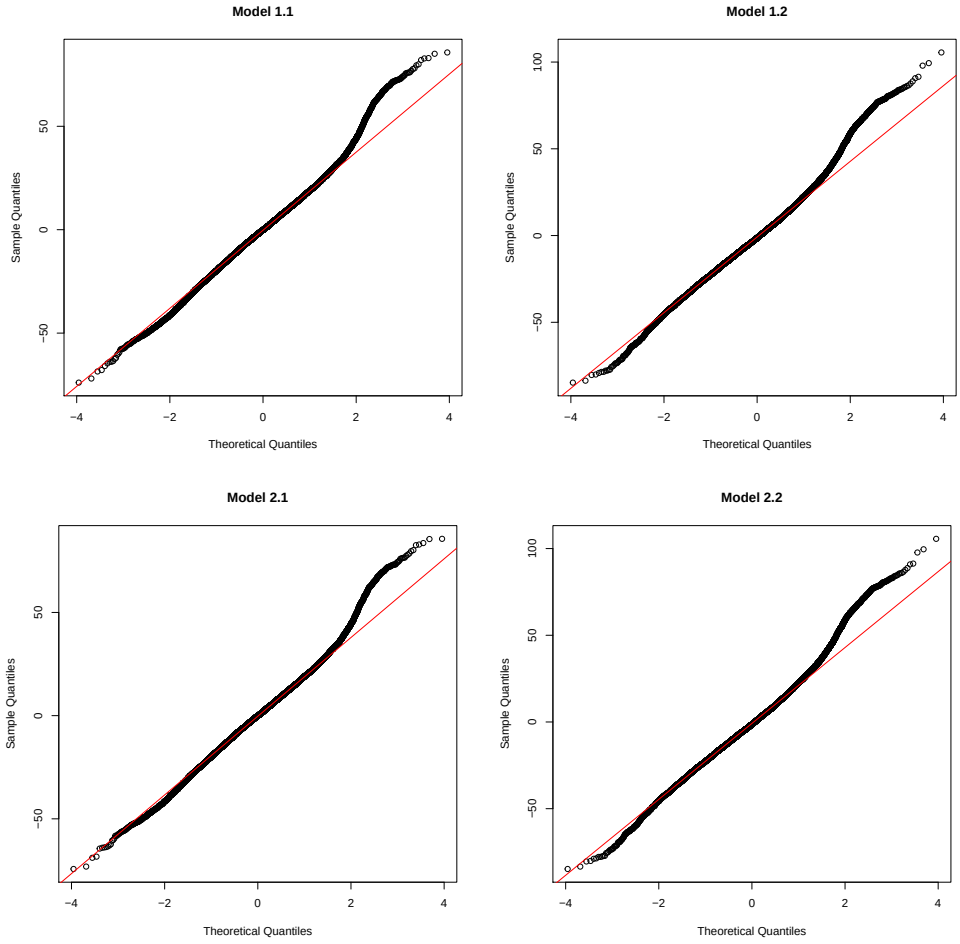
3.A Appendix A

Figure 3.6 – Residual Plots



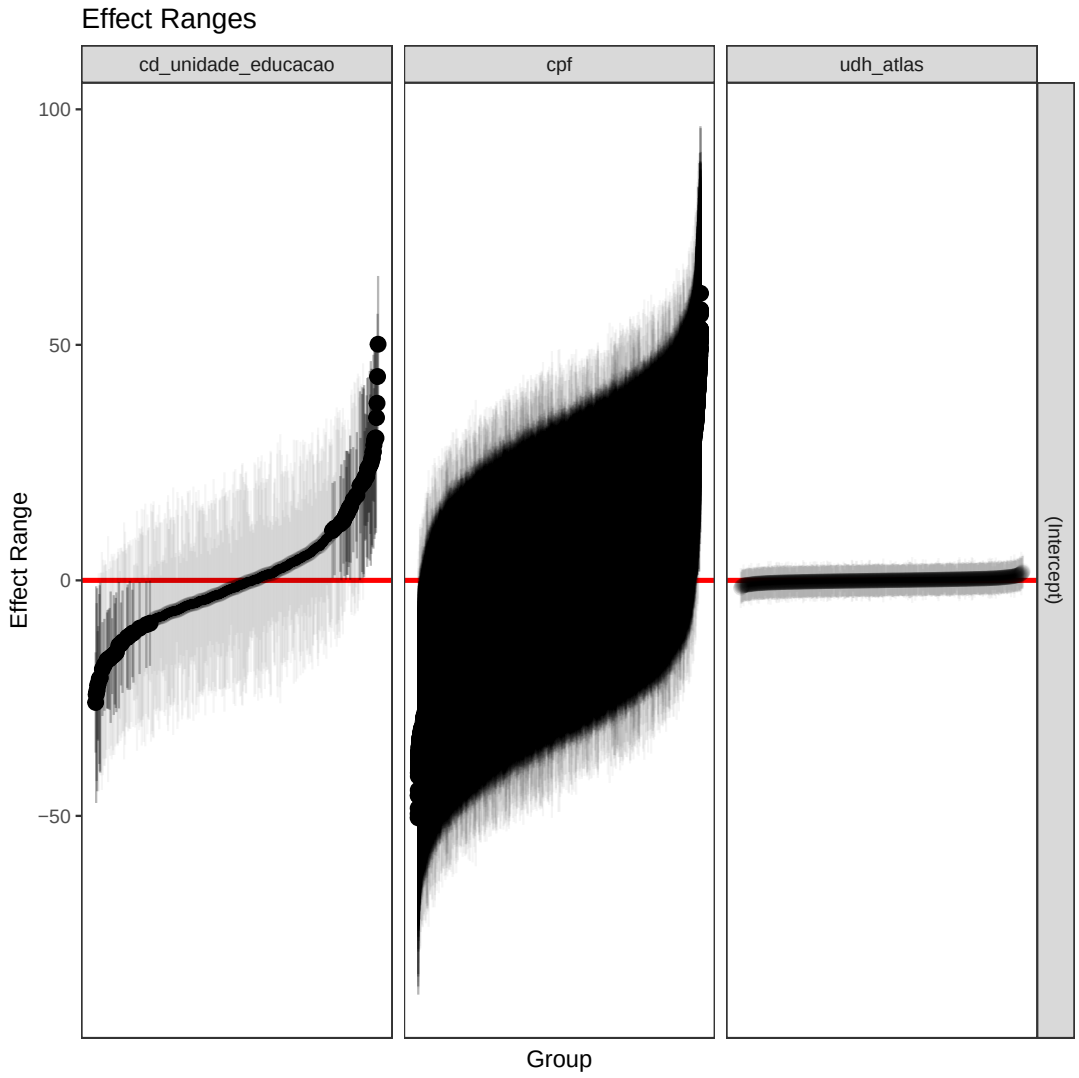
Source: Elaborated by the author.

Figure 3.7 – Normal Q-Q Plot



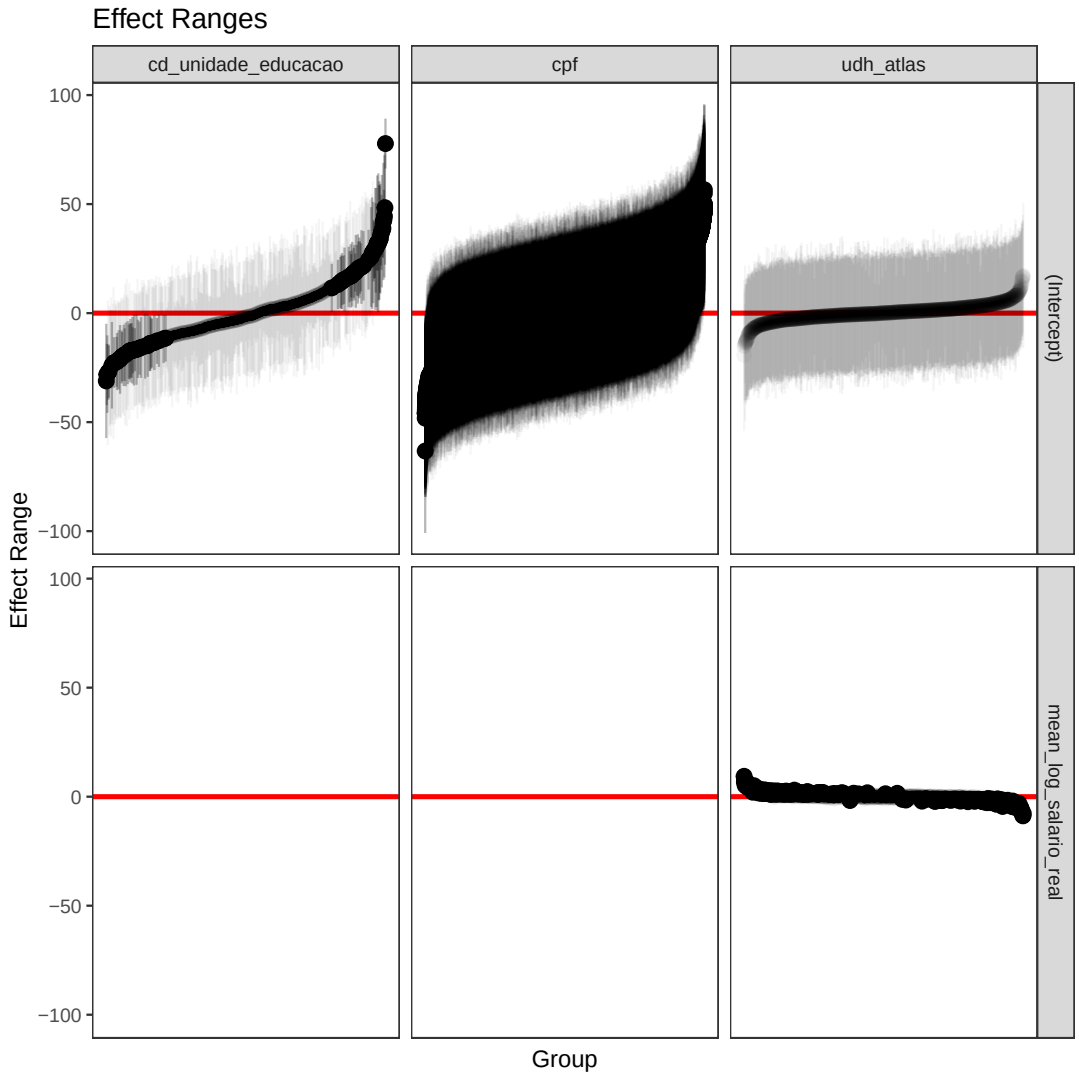
Source: Elaborated by the author.

Figure 3.8 – Random Effects by group for Model 1.1



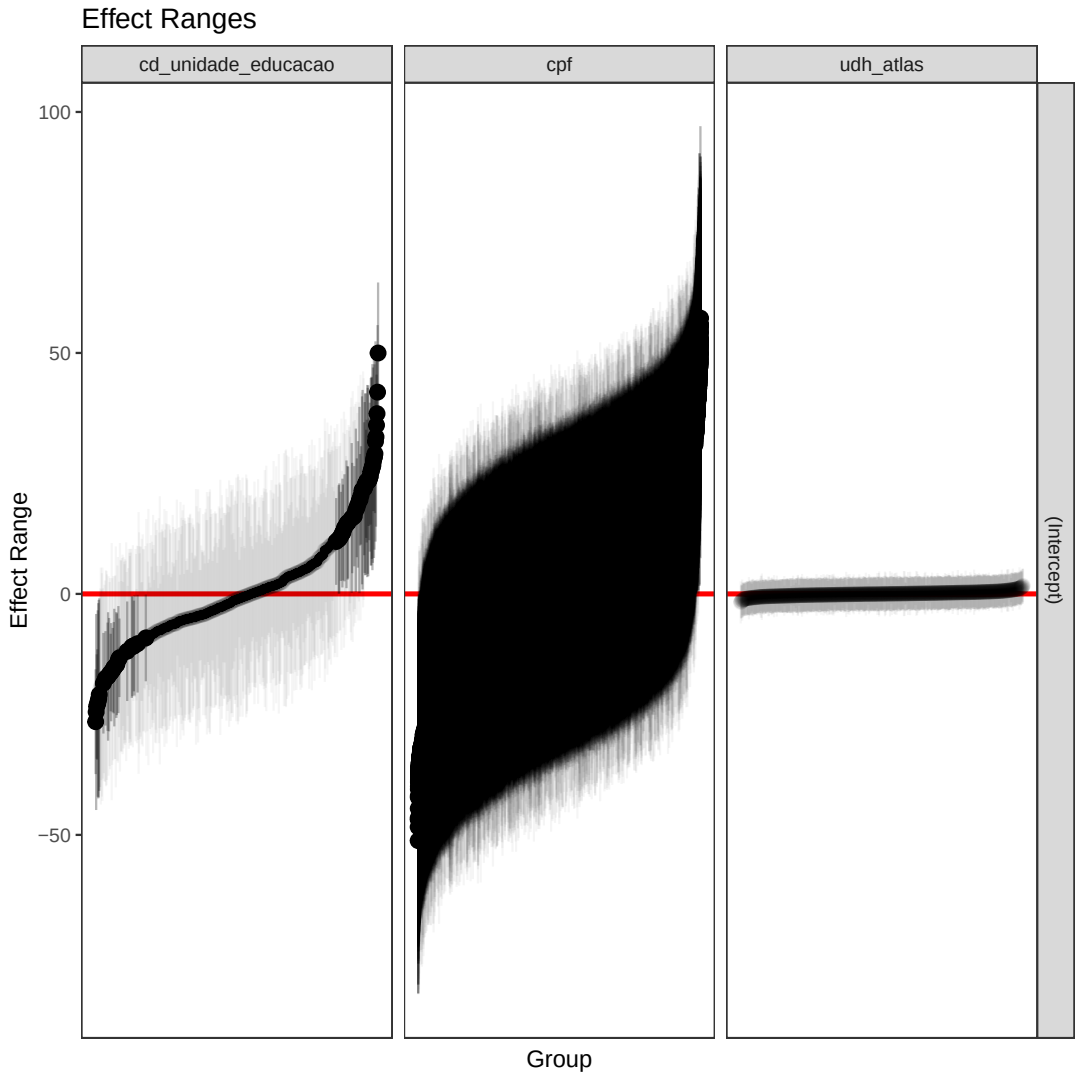
Source: Elaborated by the author.

Figure 3.9 – Random Effects by group for Model 1.2



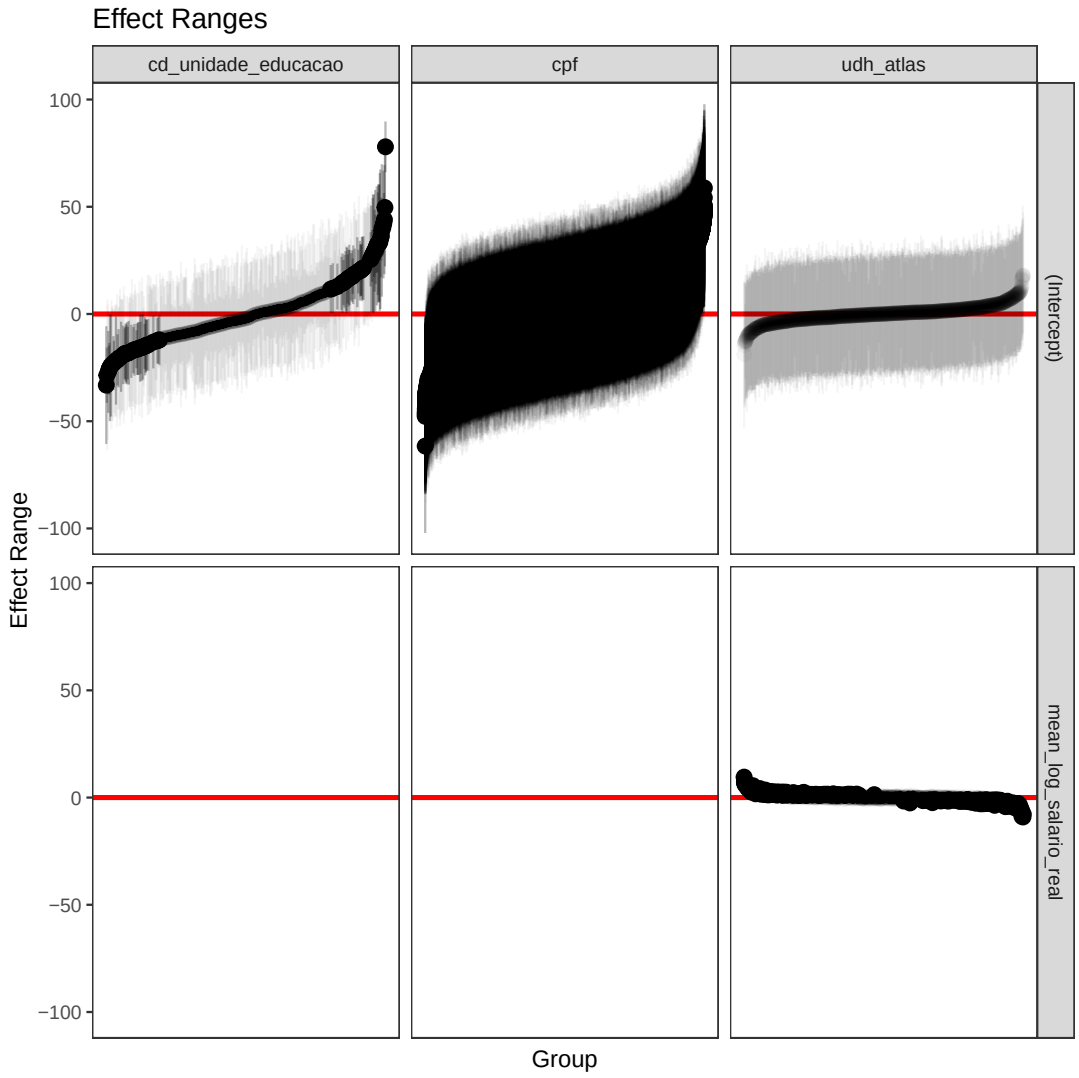
Source: Elaborated by the author.

Figure 3.10 – Random Effects by group for Model 2.1



Source: Elaborated by the author.

Figure 3.11 – Random Effects by group for Model 2.2



Source: Elaborated by the author.

Table 3.7 – Models 2.1 and 2.2 by IVS dimension

	Model 2.1	Model 2.1	Model 2.1	Model 2.2	Model 2.2	Model 2.2
(Intercept)	44.34***	47.08***	45.24***	55.69***	58.90***	57.35***
	(10.17)	(10.24)	(10.21)	(12.16)	(12.21)	(12.19)
Female	10.68***	10.84***	10.81***	1.29.	1.16.	1.13
	(0.66)	(0.63)	(0.63)	(0.72)	(0.70)	(0.69)
Non-white	−3.79***	−3.61***	−3.65***	−4.61***	−4.62***	−4.63***
	(0.68)	(0.65)	(0.65)	(0.74)	(0.72)	(0.71)
Exposure to center-based care before age 3	1.06	0.86	0.92	2.01**	2.08**	2.06**
	(0.68)	(0.65)	(0.65)	(0.75)	(0.72)	(0.72)
Age	7.89***	7.89***	7.89***	8.73***	8.75***	8.74***
	(0.61)	(0.61)	(0.61)	(0.67)	(0.67)	(0.67)
Parent's Hourly Income (ln)	13.29***	13.42***	13.50***	13.89***	13.74***	13.80***
	(0.86)	(0.85)	(0.85)	(1.01)	(1.01)	(1.01)
Parent's Gender	−4.56***	−4.57***	−4.56***	−4.57***	−4.59***	−4.57***
	(0.96)	(0.96)	(0.96)	(1.06)	(1.06)	(1.06)
IDEB15	4.37**	3.87*	4.14**	1.78	1.31	1.56
	(1.53)	(1.55)	(1.54)	(1.87)	(1.88)	(1.88)
Z-Score of IVS Infrastructure	−4.58*			0.37		
	(2.33)			(2.64)		
Parent's Hourly Income (ln)*Z-score of IVS Infrastructure	2.10*			0.31		
	(1.05)			(1.19)		
Exposure to center-based care*Z-score of IVS Infrastructure	−0.80			0.28		
	(0.94)			(1.03)		
Female*Z-score of IVS Infrastructure	0.88			−0.28		
	(0.93)			(1.01)		
Non-white*Z-score of IVS Infrastructure	0.93			−0.63		
	(0.95)			(1.04)		
Z-Score of IVS Human Capital Dimension		−4.88*			−4.25.	
		(2.20)			(2.55)	
Parent's Hourly Income (ln)*Z-score of IVS Human Capital Dimension		1.69.			1.82	
		(1.01)			(1.19)	
Exposure to center-based care*Z-score of IVS Human Capital Dimension		0.06			−0.09	
		(0.84)			(0.92)	
Female*Z-score of IVS Human Capital Dimension		0.37			0.40	
		(0.82)			(0.90)	
Non-white*Z-score of IVS Human Capital Dimension		0.62			−0.37	
		(0.84)			(0.92)	
Z-score of IVS Income and Labor Dimension			−3.97.			−3.09
			(2.27)			(2.63)
Parent's Hourly Income (ln)*Z-score of IVS Income and Labor Dimension			1.48			1.41
			(1.05)			(1.23)
Exposure to center-based care*Z-score of IVS Income and Labor Dimension			−0.29			0.02
			(0.86)			(0.95)
Female*Z-score of IVS Income and Labor Dimension			0.60			0.71
			(0.84)			(0.93)
Non-white*Z-score of IVS Income and Labor Dimension			0.81			−0.62
			(0.86)			(0.94)
AIC	131214.51	131213.14	131216.50	134676.63	134674.67	134676.54
Num. obs.	13120	13120	13120	13191	13191	13191
Num. groups: families	12982	12982	12982	13040	13040	13040
Num. groups: neighborhoods	1096	1096	1096	1095	1095	1095
Num. groups: schools	547	547	547	547	547	547
Var: families (Intercept)	487.71	488.34	489.64	532.46	530.20	530.80
Var: neighborhoods (Intercept)	4.99	4.85	5.27	233.08	240.91	236.78
Var: schools (Intercept)	169.70	170.71	170.07	271.17	270.97	270.68
Var: Residual	729.01	728.05	726.96	948.76	950.46	950.21
Var: neighborhood, Parent's Hourly Income (ln)				59.30	61.56	60.05
Cov: neighborhood, Parent's Hourly Income (ln)				−117.14	−121.36	−118.79

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{.}p < 0.1$

Source: Elaborated by the author.

Table 3.8 – Results for Sensitivity Analysis - Part I

	Model 3.1	Model 4.1	Model 3.2	Model 4.2
(Intercept)	36.01**	36.21**	51.34***	50.03***
	(11.57)	(11.55)	(13.87)	(13.90)
Female	11.52***	11.48***	1.37	1.43
	(0.80)	(0.84)	(0.89)	(0.93)
Non-white	−3.06***	−3.32***	−4.68***	−4.56***
	(0.84)	(0.88)	(0.93)	(0.97)
Exposure to center-based care before age 3	0.52	0.78	3.08***	3.33***
	(0.83)	(0.86)	(0.92)	(0.96)
Age	7.20***	7.24***	8.59***	8.62***
	(0.79)	(0.79)	(0.88)	(0.88)
Parent's Yrs. Schooling	3.52***	3.52***	2.96***	3.03***
	(0.40)	(0.41)	(0.44)	(0.46)
Parent's Hourly Income (ln)	11.14***	10.80***	11.44***	11.51***
	(1.25)	(1.26)	(1.38)	(1.48)
Parent's Gender	−2.32.	−2.45*	−2.30.	−2.45.
	(1.23)	(1.23)	(1.36)	(1.36)
Parent's Race	1.33	1.34	1.66.	1.66.
	(0.83)	(0.83)	(0.92)	(0.92)
IDEB15	3.21.	3.26*	−0.07	−0.01
	(1.64)	(1.64)	(2.04)	(2.04)
Z-score of IVS	−0.26	−9.93*	−0.65	−2.41
	(0.67)	(4.07)	(0.75)	(4.53)
Parent's Hourly Income (ln)*Z-score of IVS		4.71**		2.57
		(1.57)		(1.80)
Parent's Yrs. of Schooling*Z-score of IVS		0.09		−0.34
		(0.54)		(0.59)
Exposure to center-based care*Z-score of IVS		−1.57		−1.14
		(1.19)		(1.31)
Female*Z-score of IVS		0.16		−0.39
		(1.16)		(1.27)
Non-white*Z-score of IVS		1.48		−0.42
		(1.19)		(1.31)
AIC	77026.55	77009.29	79308.79	79304.76
Log Likelihood	−38496.27	−38484.65	−39639.39	−39630.38
Num. obs.	7715	7715	7769	7769
Num. groups: families	7633	7633	7681	7681
Num. groups: neighborhoods	990	990	991	991
Num. groups: schools	547	547	546	546
Var: families (Intercept)	406.50	398.38	464.48	452.24
Var: neighborhoods (Intercept)	0.00	0.00	0.00	312.39
Var: schools (Intercept)	305.62	154.57	276.55	275.17
Var: schools IDEB15	0.75			
Cov: schools (Intercept) IDEB15	−15.09			
Var: Residual	777.52	784.36	994.07	995.93
Var: neighborhood, Parent's Hourly Income (ln)				84.87
Cov: neighborhood, Parent's Hourly Income (ln)				−162.83

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{.}p < 0.1$

Source: Elaborated by the author.

Table 3.9 – Results for Sensitivity Analysis - Part II

	Model 1.1	Model 2.1	Model 1.2	Model 2.2
(Intercept)	39.01***	39.47***	48.64***	48.55***
	(9.67)	(9.67)	(11.62)	(11.63)
Female	11.03***	11.06***	1.22*	1.28*
	(0.55)	(0.57)	(0.61)	(0.63)
Non-white	−2.91***	−2.99***	−3.72***	−3.59***
	(0.56)	(0.59)	(0.62)	(0.65)
Exposure to center-based care before age 3	0.97.	0.99.	1.97**	1.94**
	(0.56)	(0.59)	(0.62)	(0.65)
Age	8.48***	8.48***	9.61***	9.61***
	(0.54)	(0.54)	(0.60)	(0.60)
Parent's Hourly Income (ln)	13.75***	13.50***	13.58***	13.51***
	(0.75)	(0.76)	(0.90)	(0.91)
Parent's Gender	−4.21***	−4.27***	−4.17***	−4.19***
	(0.84)	(0.84)	(0.93)	(0.93)
IDEB15	4.26**	4.28**	1.93	1.96
	(1.49)	(1.49)	(1.82)	(1.82)
Z-score of IVS	−0.76	−4.84*	−0.51	−1.70
	(0.48)	(2.08)	(0.53)	(2.40)
Parent's Hourly Income (ln)*Z-score of IVS		2.07*		0.79
		(0.95)		(1.11)
Exposure to center-based care*Z-score of IVS		−0.08		0.10
		(0.80)		(0.89)
Female*Z-score of IVS		−0.13		−0.28
		(0.79)		(0.87)
Non-white*Z-score of IVS		0.41		−0.56
		(0.80)		(0.89)
AIC	167039.51	167036.72	170882.61	170882.71
Log Likelihood	−83506.75	−83501.36	−85426.30	−85422.36
Num. obs.	16721	16721	16744	16744
Num. groups: families	16546	16546	16555	16555
Num. groups: neighborhoods	1137	1137	1133	1133
Num. groups: schools	547	547	547	547
Var: families (Intercept)	473.59	472.01	509.79	511.39
Var: neighborhoods (Intercept)	4.72	4.74	199.32	201.73
Var: schools (Intercept)	171.39	170.89	270.17	270.07
Var: Residual	735.57	737.14	974.69	973.36
Var: neighborhood, Parent's Hourly Income (ln)			50.81	51.53
Cov: neighborhood, Parent's Hourly Income (ln)			−99.58	−100.92

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{.}p < 0.1$

Source: Elaborated by the author.

Table 3.10 – Results for Sensitivity Analysis - Part III

	Model 1.1	Model 2.1	Model 1.2	Model 2.2
(Intercept)	65.63***	66.00***	67.50***	67.70***
	(9.20)	(9.21)	(11.20)	(11.20)
Female	11.36***	11.16***	1.83***	1.73**
	(0.50)	(0.52)	(0.55)	(0.57)
Non-white	−5.32***	−5.56***	−5.98***	−6.02***
	(0.51)	(0.53)	(0.56)	(0.59)
Exposure to center-based care before age 3	0.59	0.58	1.83**	1.77**
	(0.51)	(0.53)	(0.56)	(0.59)
Age	7.27***	7.28***	8.99***	8.98***
	(0.48)	(0.48)	(0.53)	(0.53)
IDEB15	5.59***	5.56***	3.70*	3.69*
	(1.44)	(1.44)	(1.78)	(1.78)
Z-score of IVS	−1.69***	−2.69***	−1.08*	−1.60·
	(0.43)	(0.77)	(0.47)	(0.84)
Exposure to center-based care*Z-score of IVS		0.04		0.30
		(0.72)		(0.78)
Female*Z-score of IVS		0.96		0.50
		(0.71)		(0.77)
Non-white*Z-score of IVS		1.07		0.21
		(0.72)		(0.78)
AIC	211048.02	211046.47	215464.85	215466.19
Log Likelihood	−105514.01	−105510.24	−107722.42	−107720.09
Num. obs.	21031	21031	21050	21050
Num. groups: neighborhoods	1171	1171	1173	1173
Num. groups: schools	547	547	547	547
Var: neighborhoods (Intercept)	2.26	2.26	0.00	0.00
Var: schools (Intercept)	166.54	166.54	270.33	270.33
Var: Residual	1275.23	1275.17	1551.40	1551.58

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Source: Elaborated by the author.

Table 3.11 – Results for Sensitivity Analysis - Part IV

	Sub-sample 1				Sub-sample 2			
	Model 1.1	Model 2.1	Model 1.2	Model 2.2	Model 1.1	Model 2.1	Model 1.2	Model 2.2
(Intercept)	53.16***	54.06***	63.33***	63.32***	37.60*	37.36*	58.45***	57.73***
	(12.36)	(12.39)	(14.81)	(14.84)	(14.66)	(14.68)	(17.51)	(17.54)
Female	9.89***	9.92***	1.72.	1.72.	10.72***	10.81***	−0.41	−0.52
	(0.84)	(0.92)	(0.91)	(0.99)	(1.03)	(1.10)	(1.16)	(1.23)
Non-white	−2.80**	−2.81**	−5.14***	−4.74***	−3.85***	−3.76***	−3.03*	−2.40.
	(0.86)	(0.94)	(0.93)	(1.01)	(1.06)	(1.13)	(1.18)	(1.27)
Exposure to center-based care before age 3	1.45.	1.10	1.99*	1.77.	1.64	1.83	2.37*	2.25.
	(0.88)	(0.95)	(0.94)	(1.02)	(1.07)	(1.14)	(1.20)	(1.28)
Age	7.16***	7.15***	8.26***	8.26***	6.89***	6.89***	7.75***	7.74***
	(0.84)	(0.84)	(0.90)	(0.90)	(1.02)	(1.02)	(1.14)	(1.14)
Parent's Hourly Income (ln)	15.61***	15.36***	13.55***	13.27***	15.55***	15.52***	15.38***	15.53***
	(1.16)	(1.19)	(1.33)	(1.36)	(1.45)	(1.47)	(1.65)	(1.67)
Parent's Gender	−5.74***	−5.80***	−4.30**	−4.37**	−5.00**	−5.02**	−6.13***	−6.06***
	(1.31)	(1.31)	(1.40)	(1.41)	(1.57)	(1.58)	(1.76)	(1.76)
IDEB15	3.17.	3.16.	1.29	1.40	6.09**	6.11**	2.36	2.43
	(1.79)	(1.79)	(2.22)	(2.22)	(2.08)	(2.08)	(2.55)	(2.55)
Z-score of IVS	−0.82	−4.62	−0.51	−4.68	0.67	0.74	−3.11.	0.57
	(0.85)	(3.63)	(0.93)	(3.92)	(1.30)	(4.72)	(1.59)	(5.42)
Parent's Hourly Income (ln)*Z-score of IVS		1.55		2.30		0.45		−1.56
		(1.65)		(1.77)		(2.10)		(2.37)
Exposure to center-based care*Z-score of IVS		1.37		0.74		−0.91		0.74
		(1.41)		(1.47)		(1.80)		(2.02)
Female*Z-score of IVS		0.00		0.05		−0.46		0.51
		(1.38)		(1.44)		(1.73)		(1.94)
Non-white*Z-score of IVS		0.01		−1.52		−0.36		−2.88
		(1.41)		(1.47)		(1.78)		(2.00)
AIC	72714.43	72710.30	76380.12	76373.99	48048.54	48043.86	49221.29	49213.51
Log Likelihood	−36344.21	−36338.15	−38175.06	−38168.00	−24011.27	−24004.93	−24595.65	−24587.76
Num. obs.	7257	7257	7483	7483	4800	4800	4804	4804
Num. groups: families	7183	7183	7406	7406	4758	4758	4750	4750
Num. groups: neighborhood	875	875	867	867	338	338	339	339
Num. groups: schools	527	527	528	528	491	491	491	491
Var: families (Intercept)	437.62	435.99	398.34	400.23	375.14	376.68	508.52	506.31
Var: neighborhoods (Intercept)	6.29	6.24	239.65	248.90	0.00	0.00	51.64	49.70
Var: schools (Intercept)	157.21	157.34	284.79	284.06	154.97	155.13	267.07	267.77
Var: Residual	792.98	794.98	1055.27	1053.35	836.74	836.09	997.64	1000.13
Var: neighborhood, Parent's Hourly Income (ln)			65.90	68.80			19.15	18.56
Cov: neighborhood, Parent's Hourly Income (ln)			−125.67	−130.86			−31.45	−30.37

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{.}p < 0.1$

Source: Elaborated by the author.

Table 3.12 – Results for Sensitivity Analysis - Part V

	With max. salary				With min. salary			
	Model 1.1	Model 2.1	Model 1.2	Model 2.2	Model 1.1	Model 2.1	Model 1.2	Model 2.2
(Intercept)	55.20***	55.35***	67.47***	67.25***	61.39***	61.43***	72.29***	72.12***
	(10.21)	(10.22)	(12.14)	(12.15)	(10.21)	(10.21)	(12.17)	(12.18)
Female	10.84***	10.67***	1.17.	1.10	10.74***	10.62***	1.10	1.07
	(0.62)	(0.65)	(0.69)	(0.71)	(0.62)	(0.65)	(0.69)	(0.71)
Non-white	−3.66***	−3.82***	−4.88***	−4.70***	−4.03***	−4.18***	−5.14***	−4.96***
	(0.64)	(0.67)	(0.70)	(0.73)	(0.64)	(0.67)	(0.70)	(0.74)
Exposure to center-based care before age 3	0.71	0.79	1.86**	1.85*	0.73	0.82	1.95**	1.93**
	(0.64)	(0.67)	(0.71)	(0.74)	(0.65)	(0.67)	(0.71)	(0.74)
Age	7.86***	7.86***	8.70***	8.70***	7.75***	7.77***	8.66***	8.66***
	(0.61)	(0.61)	(0.67)	(0.67)	(0.61)	(0.61)	(0.67)	(0.67)
Parent's Max. Hourly Income (ln)	21.38***	21.27***	20.90***	20.98***				
	(1.40)	(1.44)	(1.65)	(1.68)				
Parent's Gender	−4.15***	−4.15***	−4.03***	−4.02***	−2.29*	−2.23*	−2.50*	−2.49*
	(0.96)	(0.96)	(1.06)	(1.06)	(0.95)	(0.95)	(1.05)	(1.05)
IDEB15	4.31**	4.31**	1.85	1.87	4.84**	4.83**	2.24	2.27
	(1.55)	(1.55)	(1.88)	(1.88)	(1.55)	(1.55)	(1.89)	(1.89)
Z-score of IVS	−0.84	−1.75	−0.23	0.24	−1.02.	−0.83	−0.37	−0.17
	(0.54)	(1.67)	(0.59)	(1.88)	(0.54)	(1.41)	(0.59)	(1.59)
Parent's Max. Hourly Income (ln)*Z-score of IVS		0.55		−0.44				
		(1.83)		(2.09)				
Exposure to center-based care*Z-score of IVS		−0.38		0.08		−0.47		0.14
		(0.91)		(1.00)		(0.91)		(1.00)
Female*Z-score of IVS		0.80		0.38		0.58		0.15
		(0.89)		(0.97)		(0.90)		(0.97)
Non-white*Z-score of IVS		0.72		−0.86		0.68		−0.85
		(0.91)		(1.00)		(0.92)		(1.00)
Parent's Min. Hourly Income (ln)					12.95***	12.70***	15.34***	15.34***
					(1.29)	(1.30)	(1.49)	(1.54)
Parent's Min. Hourly Income (ln)*Z-score of IVS						−0.98		0.05
						(1.77)		(2.03)
AIC	131239.83	131238.27	134719.95	134718.28	131371.59	131367.43	134781.98	134780.48
Log Likelihood	−65606.92	−65602.14	−67344.97	−67340.14	−65670.79	−65666.71	−67375.99	−67371.24
Num. obs.	13120	13120	13191	13191	13120	13120	13191	13191
Num. groups: families	12982	12982	13040	13040	12982	12982	13040	13040
Num. groups: neighborhoods	1096	1096	1095	1095	1096	1096	1095	1095
Num. groups: schools	547	547	547	547	547	547	547	547
Var: families (Intercept)	490.67	489.15	535.10	535.09	492.32	496.23	531.46	531.92
Var: neighborhoods (Intercept)	5.53	5.57	74.76	74.09	20.11	4.54	44.77	45.03
Var: schools (Intercept)	172.02	171.87	270.84	270.97	171.74	171.59	273.43	273.54
Var: Residual	727.41	729.14	951.89	952.36	736.56	735.34	963.16	963.02
Var: neighborhood, Parent's Max. Hourly Income (ln)			129.51	128.32				
Cov: neighborhood, Parent's Max. Hourly Income			−96.56	−95.72				
Var: neighborhood, Parent's Min. Hourly Income					37.50		100.53	101.98
Cov: neighborhood, Parent's Min. Hourly Income					−25.00		−67.09	−67.77

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{.}p < 0.1$

Source: Elaborated by the author.

4 Demand for Early Childhood Education and Care in São Paulo

4.1 Introduction

In today's world, a multitude of actors participate, and are responsible, for the care of young children. Care itself also takes place in a multitude of settings: the child's home, in a relative's or friend's home or in more institutionalized settings, like daycare centers.

As our society becomes more urbanized and there is a rise in women's labor market participation, the need for affordable early childhood education and care by actors outside of the immediate family in institutional settings increases especially among lower income families (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; DUNCAN; MAGNUSON; VOTRUBA-DRZAL, 2014). Evidence also indicates that center-based care provides higher on average quality than unregulated arrangements, particularly for low income families (BASSOK; LATHAM, 2017; NICHD Early Child Care Research Network, 2002).

High-quality center-based childcare became a central public policy for guaranteeing optimal child development and reducing inequalities. However, the access to this is not equal and the poorest children have less access to center-based care and when they do have access, it is of lower quality (BURCHINAL et al., 2008; LEVENTHAL; BROOKS-GUNN, 2000).

To understand the access to center-based care for children between zero and five years old in the city of São Paulo, this chapter presents a descriptive analysis of those who tried to enter into the system. As discussed in chapter 2, there is an enrollment system that organizes the entrance to the ECEC centers. In this chapter, we use data from this system to analyze the characteristics of those children who were registered into the system between 2010 and 2019.

In Brazil, for example, it is mostly higher income families who use daycare. In 2015, among children from families in the top 10% of the Brazilian income distribution, 53% were enrolled in center-based care, whereas in the bottom 10%, this percentage was 16%[1]. This is also true for Latin America - where children whose mothers are highly educated are more likely to attend center-based care (BERLINSKI; SCHADY, 2015) - and in the USA, where toddlers from low-income families have a lower probability of being in center-based care, and if in care, have a higher probability of receiving lower quality care (RUZEK et al., 2014; NICHD Early Child Care Research Network, 2006).

[1] Data from Pesquisa Nacional por Amostra de Domícilios (PNAD/IBGE).

In addition to this introduction, the chapter is divided into six sections. The next section presents the findings in the literature on the use of early childhood education and care. The third section presents the research questions that guide this study. The fourth presents the data, sample and analytical plan. The fifth presents the results, followed by a discussion in the sixth section.

4.2 Literature Review

This study builds on a growing body of research that discusses personal and contextual factors that influence the type of early childhood education and care (ECEC) a child experiences. Different ECEC arrangements are possible including young children being cared for by their parents or by others in a myriad of settings such as in the homes of relatives, in home-based settings by non-relatives and in more formal settings such as childcare centers. Parental decisions on the arrangements for their child are shaped by their and their child's own characteristics, their preferences for particular characteristics in the type of care, their employment context, as well as the availability of different types of ECECs in the community, and constraints such as cost and access[2] (CHAUDRY; HENLY; MEYERS, 2010).

Parents' socioeconomic status as measured by family income or parent's level of education, is an important factor to understand both the use of non-parental care as well as the type of care arrangement chosen. Higher-income families use more non-parental care alternatives in the USA (NORES; BARNETT, 2014; MADILL et al., 2018) as in countries with comprehensive ECEC policies (SIBLEY et al., 2015). Even when lower income parents do use non-parental care arrangements, they are less likely to use center-based care than higher-income families (MADILL et al., 2018; NORES; BARNETT, 2014). This gap has been documented since 1968, and while still substantial, it has narrowed in recent decades (MAGNUSON; WALDFOGEL, 2016). When parental education is analyzed, a similar pattern is found and children whose parents have more years of education, particularly their mothers, are more likely to be enrolled in non-parental care and specifically in center-based care (CHAUDRY; HENLY; MEYERS, 2010; FRAM; KIM, 2008).

Maternal employment is also a key factor in explaining ECEC use especially for lower-income working mothers (LOEB et al., 2004; AHNERT; PINQUART; LAMB, 2006; SIBLEY et al., 2015). Family and household composition is another important factor: both marital status and the number of children in the household are predictors of ECEC utilization and are affect the type of care chosen for the children (CHAUDRY; HENLY; MEYERS, 2010; SIBLEY et al., 2015). Single-parent households are more likely to use

[2] Chaudry, Henly e Meyers (2010) present a conceptual framework, developed by the Child Care Policy Research Consortium that organizes all these factors.

relative care and families with more children are less likely to use any ECEC at all (CHAUDRY; HENLY; MEYERS, 2010; SIBLEY et al., 2015).

Child characteristics also are important when explaining ECEC use. Younger children tend to be in more informal alternatives of care when compared to older, preschool-age children (CHAUDRY; HENLY; MEYERS, 2010). In the USA, evidence indicates that race and ethnicity is associated with the type of ECEC used: Latino children tend to be in home-based care and Black children are more likely to be in center-based care (CHAUDRY; HENLY; MEYERS, 2010; FRAM; KIM, 2008). In addition to family and child characteristics, contextual factors affect both the use of ECECs, and the type of arrangement chosen. The availability of different types of care in the community is of special importance. Low income families need options that are easy to access and many times, they choose care arrangements according to logistical considerations and not their own preferences or quality concerns (CHAUDRY; HENLY; MEYERS, 2010).

However, there is greater availability of high-quality center-based care in more advantaged neighborhoods and not where most low income families live (BURCHINAL et al., 2008; DUPERE et al., 2010). This is concerning because the availability of childcare alternatives in neighborhoods is an important determinant of its use by low-income parents, corroborating the idea that location and easy access are important (HIRSHBERG; HUANG; FULLER, 2005). Additionally, parents learn a lot about childcare options through informal channels and therefore, living in neighborhoods where few families use formal ECECs can hinder the odds of learning about their options (FRAM; KIM, 2008).

The difference in availability by neighborhood has consequences on the quality of care that children end up receiving. There is evidence that children raised in more advantaged neighborhoods have higher quality ECEC experiences, even after controlling for family characteristics (DUPERE et al., 2010). This association seems to be particularly strong for children whose mothers do not have a high school education (BURCHINAL et al., 2008).

4.3 The current study

The municipal early education and care system in São Paulo expanded substantially in the past decade[3]. There was not only a growth in the number of spots in the system, but also an expansion in more vulnerable neighborhoods[4]. International evidence indicates that there is a positive association between the availability of child care in communities and its utilization of it by low-income families (HIRSHBERG; HUANG;

[3] This expansion was described at length in chapter 2.
[4] This was also presented in chapter 2.

FULLER, 2005). Based on this, the current study tests the hypothesis that the increase in availability of spots in the ECEC system in São Paulo expanded childcare access for more vulnerable families and subsequently modified the profile of those who attend.

The analysis of the characteristics of those who use the ECEC system is very important for guiding policies on childcare provision. And understanding if and how these characteristics are changing across time, is crucial to adjust these policies.

There is also some international evidence that access to quality center-based care is associated with neighborhood affluence (BURCHINAL et al., 2008; DUPERE et al., 2010). This study tests this hypothesis for São Paulo and contributes to this literature with evidence from a large city in the Global South from a free, publicly provided ECEC system. Understanding if this association exists in São Paulo is important to guide childcare policies.

By analyzing administrative data from the São Paulo municipal ECEC system from 2010 to 2018, this study answers the following questions:

- What are the characteristics of the children who were registered in the municipal ECEC system between 2010 and 2018?
  - Has the profile of those who enter the system changed in the past decade? Following the empirical evidence of an expansion in the availability of spots presented in chapter 2, along with the evidence from the literature on the association between availability and use of center-based care by low-income families (HIRSHBERG; HUANG; FULLER, 2005), our hypothesis is that children from more vulnerable backgrounds started to access the system more.
  - Does the profile of the children vary between those who register for care before age three and after? No evidence was found in the literature for this, but, as attending preschool is mandatory and center-based care for children younger than three is not, we expect that the two groups differ.
- What are the characteristics of the children who are enrolled in the municipal ECEC system?
- What factors influence the access to the ECEC system? We test if different parental and children's characteristics, along with neighborhood socioeconomic vulnerability are associated with the access to the ECEC system. We also compare the access to daycare centers and to preschools and the actual enrollment in either.

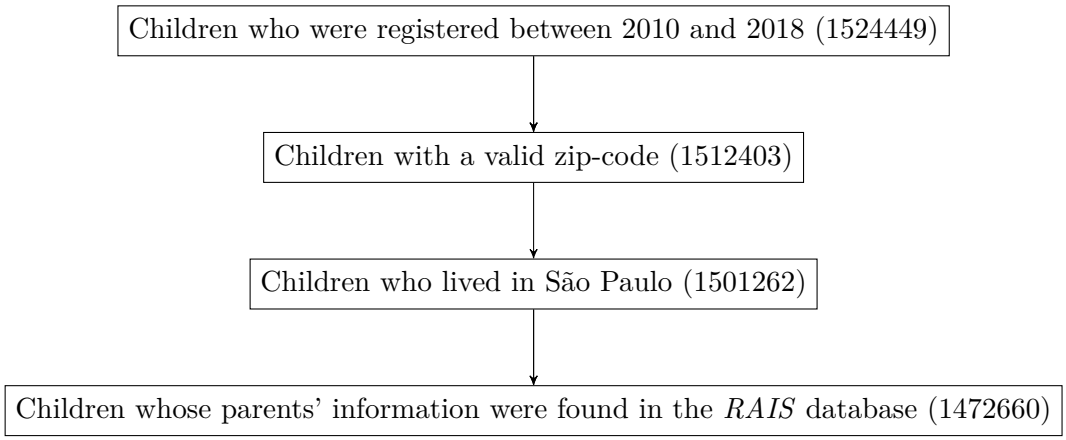
4.4 Method

4.4.1 Data and Sample

Data is drawn from databases already discussed at length in chapter 2 and chapter 3. The administrative database of the enrollment process is the main database. Using the parents official ID, this database was merged with the *Relação Anual de Informações Sociais* database. By using the residential postal code at the time of request, a merge with the Ipea database was implemented[5].

The sample used in this study is derived from the complete administrative database and includes all the children who were registered in the municipal early education and care system by their parents[6] for the school years of 2010 to 2018. To define the sample, a merge with both the *RAIS* database and the Ipea database was implemented and children who lived outside the city of São Paulo were excluded[7]. The final sample included 1,472,660 children. Figure 4.1 illustrates the restriction process that defined the sample[8]

Figure 4.1 – Sample Selection



Source: Elaborated by the author.

It is important to point out that the children in this sample were all registered in the municipal ECEC system. Their parents therefore, knew about this initiative, had an interest in it and actively tried to enroll their children. Thus, they are potentially different from other parents in the city of São Paulo who did not try to enroll their children for

[5] Following the discussion in chapter 3, we assume the human development unit is a proxy for the neighborhood.

[6] Throughout this thesis, I refer to the adult who registered the child as the parent even though they may not necessary be the parent, but an adult legally responsible for the child.

[7] There were 11,141 children whose residential postal codes were located in the metropolitan region of São Paulo. It is not clear how those children applied for a spot in the municipal system if they lived in another city; I opted to not analyze them.

[8] The child's residential postal code was defined as valid if it allowed for a merge with the *IVS* database. To merge the administrative database with *RAIS* database, I implemented the same merge as in chapter 3.

whatever reason. The results from this study cannot be generalized for all parents and young children who live in São Paulo.

4.4.2 Measures

In addition to the measures described in chapter 3 - including the child's residential neighborhood vulnerability, the child's characteristics and the parent's characteristics -, this study uses different measures of access to the municipal early education and care system. The ECEC system is divided into two main sections; one that serves children who are three years old or younger in daycare centers, and the one that serves children between ages four and five in preschools. The access measures used in this chapter, take into consideration this division. These measures also take into account the different steps in the process such as registering into the system indicating access to the system, but full access only happens when the children are enrolled into daycare centers or preschools. Five different measures of access were constructed to be used as independent variables in this study.

The first three measures cover the first step in the enrollment process: the registration. The first measure - R_t - indicates if the child was registered into the ECEC system. This equaled 1 for all the sample as all the children in it were registered into the ECEC system. The second measure - R_d - is an indicator variable that equals 1 if the child was first registered into an age group[9] that is part of the day-care center system. The second measure - R_p - is also an indicator variable that equaled 1 if the children were first registered into an age group that is part of the preschool system. To construct both measures, only the first attempt at registering each child was considered[10].

The other three variables use information on the enrollment into the system. To do so, only the last attempt at registering the child was considered[11]. First, an indicator variable E_d was constructed to indicate those children who enrolled into the day-care center division: it equaled one if the child was enrolled in an age group that is part of that division of the ECEC system. Another indicator variable E_p was constructed to indicate that the child was enrolled in an age group that was part of the preschool division. A third variable - E_n - equaled 1 if the child was registered into the system - in either division - but never enrolled in any.

[9] In chapter 2, there is an explanation of the different levels of the system.

[10] As discussed in other chapters, children can be registered into the system multiple times: in this sample, 42% did so, with some children entering the system more than 10 times throughout the years. To define the first attempt, I use the information from the day and time of registration and choose the oldest registration date and hour.

[11] When children register multiple times, it might indicate that they enrolled and dropped out. Because of that, we only consider the last attempt and suppose that it described the final result. We do not take into account that children might have dropped out.

4.4.3 Analytical Plan

To answer the first research question and its sub-questions, a descriptive analysis of the sample was implemented. To do so, all family, children and neighborhood variables available in the three databases were used. This descriptive analysis considered only the first attempt at registering each child in the sample.

The first sub-question requires the comparison between two sub-samples: those who registered for the 2010 school year and those who registered for the 2018 school year[12]. T-tests were used to compare to zero the difference in the means of continuous variables from the two sub-samples[13]. For the binary variables, child and parental race, and sex, a test on the equality of proportions using large-sample statistics was implemented[14].

The second sub-question requires the comparison between those who registered for daycare centers, that is, children with R_d equal 1, and those who registered for preschool - that is, children with R_p equal 1. Once again, t-tests and tests on the equality of proportions were implemented for all eleven variables were implemented.

To answer the second question, only the last attempt at registering each child was considered. Because children could have registered several times and we are interested only in the final outcome, it does not make sense to consider either the first solicitation or the other ones. The final attempt at registration is the best indicator of the child's final outcome in the ECEC system[15]. Then, we compare three groups of children: those who have variable E_d equaled to one, those who have variable E_p equaled to one and those who have variable E_n equaled to one. To compare the descriptive statistics, two tests were used: for the continuous variables, the Kruskal-Wallis H test and for categorical variables, the Pearson's Chi-squared test.

To answer the third question, logistic regressions were used. Two sets of models were estimated: in the first one, the independent variable equaled 1 if E_d or E_p equaled one; in the second one, the independent variable was E_d and only the children who actually enrolled, that is, with E_d equal to 1 or E_p equal to one, were considered. The dependent variables were the family, child and neighborhood characteristics examined previously. First, a model without the neighborhood variable was estimated with clustering at the family level[16]. We then estimated the model with the vulnerability index and clustering at the neighborhood level and the family level. We repeated this estimation and considered all three dimensions of the vulnerability index separately

[12] This does not mean that the date of registration is in 2010 or in 2018. A parent might register in 2009 for the following school year.

[13] This was done using the command *ttest*.

[14] This was done using the command *prtest*.

[15] It is not a perfect indicator as children might enroll and then, drop-out.

[16] As discussed previously, we considered two children to be in the same family if the parent had the same ID.

4.5 Results

Table 4.1 summarizes the characteristics of the 1,472,829 children who registered in the São Paulo ECEC system between the school years of 2010 and 2018. The first section of Table 4.1 presents the children's characteristics. There is an equal share of boys and girls with 49% of children being female, and of white and non-white children where 46% of children are non-white[17]. The children were around one and a half years old when they were first registered in the system.

The second section of Table 4.1 has information on the parental characteristics. The children in the sample are part of 841,078 families[18]. The parent's mean hourly salary was 4.67 in 2017 Brazilian *reais* and they had, on average, 6.59 years of schooling - almost equivalent to a complete middle school education. Most of the parents were mothers - only 13% of the parents in the sample were male - and white - 56% of the parents.

The third section of Table 4.1 has data on the social vulnerability of the child's residential neighborhood in 2010[19]. Children lived in 1,377 different neighborhoods[20] in the city of Sao Paulo. The table presents the z-score of the IVS, both for the full index as well as for each sub-dimension[21]. The z-scores indicate that the HDUs present in our sample had lower social vulnerability in all three domains, as well as in the full index, when compared with the city of São Paulo as a whole.

[17] As discussed previously, in chapter 3, there is a large percentage of children without race information in the administrative database. In this sample, 40.4% of the children had missing information.

[18] Following the analyses conducted in other chapters, this study considers that children belong to the same family if the adult who registered them had the same identification number. Throughout this thesis, this adult is called the child's parent.

[19] Data for neighborhood vulnerability was only available for 2010; it is important to point out that this data is not ideal to describe the social vulnerability of the neighborhood in more recent years. However, because the *IVS* uses data from the Brazilian Census, which happens every 10 years, the most recent information is from 2010.

[20] Throughout this thesis, neighborhoods are defined as the human development units that the child's postal code, given at the moment of registration, belonged to.

[21] The z-score for the index and for the sub-indexes were calculated using the mean and the standard deviation of the city of São Paulo calculated at the HDU level.

Table 4.1 – Descriptives of children registered in the enrollment process of the ECEC system, 2010-2018

	Mean	SD	Obs.
Student			
Female	0.49		1472660
Non-white	0.46		877440
Age	1.44	1.59	1472660
Parents			
Parent's Hourly Income	4.67	6.69	841078
Parent's Yrs. of Schooling	6.59	1.19	841078
Parent's Gender	0.13	0.33	841078
Parent's Race	0.56	0.50	436691
Neighborhood			
Vulnerability Index	-0.02	0.97	1377
Vulnerability Index - Infrastructure Dimension	-0.00	1.00	1377
Vulnerability Index - Human Capital Dimension	-0.02	0.96	1377
Vulnerability Index - Income and Labor Dimension	-0.02	0.96	1377

Source: Elaborated by the author considering only the first registration in the system.

Table 4.2 presents the descriptive statistics and comparison tests[22] between two sub-samples: children who were first registered for the 2010 school year and children who were first registered for the 2018 school year[23]. In the sample, 74,127 children were first registered for the 2010 school year and 162,712 children, for 2018.

There is a clear change in the profile of the children. Children who were registered for the 2018 school year are younger and there is a higher percentage of non-white children. The parents who first registered their child in 2018 received higher salaries and were more educated than those who first registered their children in 2010. Also, more mothers and non-white parents registered their child in 2018. The p-values confirm this descriptive analysis. However, there no statistically significant change in the social vulnerability of the residential neighborhoods of the children who registered into the system in 2010 and in 2018[24]. In the 2018 sub-sample, more neighborhoods are represented, 1,352, as compared to 1,325 in 2010.

[22] T-tests were implemented for continuous variables and a tests of proportion for the binary variables (sex and race).

[23] For this analysis, as the analysis in section 4.1, only the first attempt at registering the child was considered.

[24] As discussed previously, there is only data for vulnerability in 2010. This might be problematic as neighborhoods may have become much less vulnerable during the eight-year gap.

Table 4.2 – Comparison between children registered in 2010 and in 2018 for the enrollment process of the ECEC System

	2010		**2018**		
	Mean	Obs.	Mean	Obs.	P-value
Student					
Female	0.49	74127	0.49	162712	0.04
Non-white	0.43	47823	0.48	105959	0.00
Age	2.48	74127	1.01	162712	0.00
Parents					
Parent's Hourly Income	4.59	48961	4.69	107062	0.01
Parent's Yrs. of Schooling	6.39	48961	6.70	107062	0.00
Parent's Gender	0.13	48961	0.11	107062	0.00
Parent's Race	0.56	24663	0.55	54921	0.04
Neighborhoods					
Z-Score - Vulnerability Index	-0.02	1325	-0.02	1352	0.83
Z-Score - Vulnerability Index - Infrastructure Dimension	-0.00	1325	-0.01	1352	0.92
Z-Score - Vulnerability Index - Human Capital Dimension	-0.02	1325	-0.03	1352	0.78
Z-Score - Vulnerability Index - Income and Labor Dimension	-0.03	1325	-0.03	1352	0.85

Source: Elaborated by the author considering only the first registration in the system.

Table 4.3 presents the descriptive statistics and comparison tests[25] between two sub-samples: children who were first registered in the daycare center division of the ECEC system and children who were first registered in the preschool division of the ECEC system. For both sub-samples, we analyzed all the children who were registered between the 2010 and 2018 school years. In this study's sample, 1,287,771 children were first registered for daycare centers and 187,889 children for preschool.

As expected, children who were first registered for the daycare division were much younger, 1.01 years old compared to 4.38 years old. But the children also differ according to the other personal and family characteristics. More non-white children registered for the daycare center division of the ECEC system. These children had parents with higher incomes and more years of education. More mothers and non-white parents registered their children for this division. Once again, there was no statistically significant difference in the vulnerability of the neighborhoods these children come from.

[25] T-tests were implemented for continuous variables and a tests of proportion for the binary variables (sex and race).

Table 4.3 – Comparison between children registered for the daycare center and for preschool for the enrollment process of the ECEC System

	Daycare center		Preschool		
	Mean	Obs.	Mean	Obs.	P-value
Student					
Female	0.49	1284771	0.49	187889	0.00
Non-white	0.47	753858	0.44	123582	0.00
Age	1.01	1284771	4.38	187889	0.00
Parents					
Parents' Hourly Income	4.66	765085	4.40	112352	0.00
Parents' Yrs. of Schooling	6.61	765085	6.35	112352	0.00
Parents' Gender	0.12	765085	0.16	112352	0.00
Parents' Race	0.55	398738	0.57	54195	0.00
Neighborhood					
Z-Score - Vulnerability Index	-0.02	1377	-0.02	1360	0.91
Z-Score - Vulnerability Index - Infrastructure Dimension	-0.00	1377	-0.01	1360	0.92
Z-Score - Vulnerability Index - Human Capital Dimension	-0.02	1377	-0.02	1360	0.92
Z-Score - Vulnerability Index - Income and Labor Dimension	-0.02	1377	-0.03	1360	0.93

Source: Elaborated by the author considering only the first registration in the system.

Table 4.4 presents the descriptive statistics and comparison tests[26] between three sub-samples: those who were registered but did not enroll- that is, E_n equaled to one -, those who were granted a spot into the daycare center division - E_d equaled to one -, and those who were granted a spot into the preschool system - E_p equaled to one. In this sample, 246,500 (16.7%) children belonged to the first group, 802,378 (54.5%) to the second group and 423,782 (28.8%) to the third group.

There are statistically significant differences among the three groups of children when their personal and family characteristics are analyzed. In comparison with the children who entered the ECEC system during the preschool years, children who were enrolled in the daycare centers were younger and there was a higher proportion of female and non-white children. Their parents had higher hourly salaries and more years of schooling. In comparison with the children who were registered but never attended the ECEC system, the children who attended the daycare centers were younger and there was a higher proportion of female and non-white children. Their parents had lower salaries and fewer years of schooling.

In the last section of Table 4.4, neighborhood characteristics were analyzed and once again, there was not any statistically significant difference between groups when the neighborhood's social vulnerability was analyzed whether with the full index or when each dimension was analyzed separately.

[26] For the continuous variables, the Kruskal-Wallis H test was used to test if at least two samples were different. For categorical variables, the Pearson's Chi-squared test was used.

Table 4.4 – Comparison between sub-samples according to status of enrollment

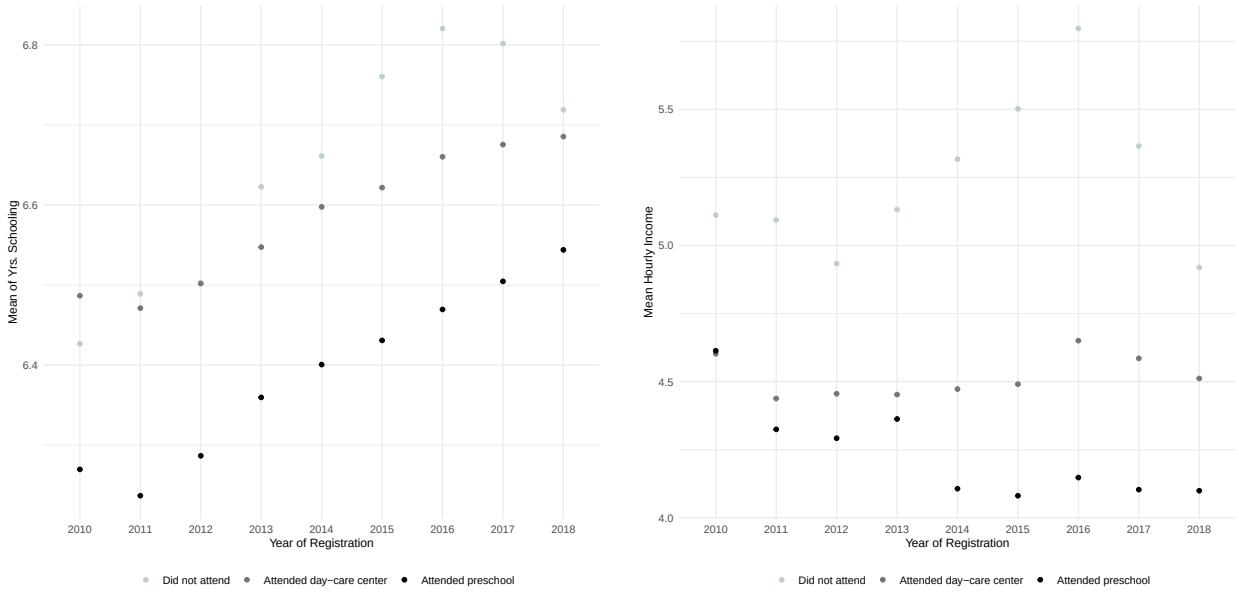
	Did not attend	Attended daycare	Attended preschool	P-value
Student				
Sex				$< 0.001^{1}$
Female	125940 (51.1%)	413160 (51.5%)	215075 (50.8%)	
Male	120560 (48.9%)	389218 (48.5%)	208707 (49.2%)	
Race				$< 0.001^{1}$
White	76506 (31.0%)	246414 (30.7%)	146678 (34.6%)	
Non-white	58588 (23.8%)	225768 (28.1%)	123486 (29.1%)	
Missing	111406 (45.2%)	330196 (41.2%)	153618 (36.2%)	
Age				$< 0.001^{2}$
Mean (SD)	2.29 (1.70)	1.13 (1.14)	4.10 (0.85)	
Obs.	246500	802378	423782	
Family	N=230448	N=696978	N=385241	
Parent's Hourly Income				$< 0.001^{2}$
N-Miss	70811	171061	120299	
Mean (SD)	5.18 (7.58)	4.52 (5.91)	4.20 (6.86)	
Parent's Yrs. of Schooling				$< 0.001^{2}$
N-Miss	70811	171061	120299	
Mean (SD)	6.66 (1.28)	6.61 (1.15)	6.42 (1.22)	
Parent's Gender				$< 0.001^{1}$
N-Miss	70811	171061	120299	
Female	140057 (87.7%)	468357 (89.1%)	228834 (86.4%)	
Male	19580 (12.3%)	57560 (10.9%)	36108 (13.6%)	
Parent's Race				$< 0.001^{1}$
N-Miss	152471	421422	253912	
Non-White	32850 (42.1%)	124682 (45.2%)	58662 (44.7%)	
White	45127 (57.9%)	150874 (54.8%)	72667 (55.3%)	
Neighborhood	N=1357	N=1373	N=1366	p value
Z-Score - Vulnerability Index				0.996^{2}
Mean (SD)	0.29 (0.10)	0.29 (0.10)	0.29 (0.10)	
Z-Score - Vulnerability Index - Infrastructure Dimension				0.995^{2}
Mean (SD)	-0.01 (0.99)	-0.00 (1.00)	-0.01 (1.00)	
Z-Score - Vulnerability Index - Human Capital Dimension				0.996^{2}
Mean (SD)	-0.02 (0.96)	-0.02 (0.96)	-0.02 (0.96)	
Z-Score - Vulnerability Index - Income and Labor Dimension				0.997^{2}
Mean (SD)	-0.03 (0.95)	-0.02 (0.96)	-0.03 (0.95)	

Note: 1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test
Source: Elaborated by the author.

Figure 4.2 presents two graphs with the behavior of parent's income and education throughout time, for the three sub-samples presented in Table 4.4. The pattern for parent's years of schooling is similar to the one discussed when Table 4.4 was analyzed; parents of children who do not attend the system have higher years of schooling than those who attended daycare centers and those who attended preschool are those whose parents have the fewest years of schooling. The mean years of schooling has risen among all three groups and the difference between groups has reduced through the time period. For parent's hourly income, the pattern is similar to that of parent's education but its evolution is but less consistent. The difference between groups varies throughout the period and not in a single direction.

Table 4.5 presents the results of multivariate logit regression models. The coefficients of these models were transformed into odds ratios. The standard errors were

Figure 4.2 – Parent's Hourly Income and Yrs. of Schooling by Year of Registration, 2010-2018



Source: Elaborated by the author.

calculated with adjustments for cluster at the family level for model 1 and at the family and neighborhood level for models 2 and 2.2.

In analyzing the first three models, whose outcome was enrollment into the ECEC system, several conclusions can be inferred. First, children's race and age affect the odds of enrollment; the odds of non-white children enrolling into the ECEC system are 14% higher than of white children and one year more of age reduces the odds of enrollment by 8%. Parental income and years of schooling also matter as an increase in one unit of either variable reduces the odds of enrollment by 2% and 8%, respectively. The parent's race has a significant effect as well; the odds of white parents enrolling their child is 5% lower than of non-white parents. The vulnerability of the residential neighborhood the child lives in positively affects the odds of enrollment into the system as an increase in one unit in the z-score of the IVS increases the odds of enrollment by 127%. When the three dimensions are analyzed separately, a different pattern emerges. All three dimensions have statistically significant coefficients but an increase in the human capital dimension leads to a reduction in the odds of enrollment.

When enrollment into the daycare center system in comparison to enrollment into the preschool system is analyzed, some different results appear. First, the coefficients for children's and parent's race are no longer significant. The coefficient for the child's age reduces its value; a one-year increase in the child's age reduces the odds of enrollment by 92%. The coefficient for parental income is still significant but its size decreases; a one-unit increase lowers the odds of enrollment by 1%. For parental education however, the coefficient changes; a one-unit increase increases the odds of enrollment into the daycare center division of ECEC. There is also a significant change in the coefficient for the z-score

of the IVS; a one-unit increase reduces the odds of enrollment by 74%. When all three dimensions are analyzed separately and comparing both outcomes, there is a change in the coefficient for the infra-structure division; a one-unit increase in this z-score reduces the odds of enrollment by 8%.

Table 4.5 – Statistical models

	Enrollment in the ECEC sytem			**Enrollment in daycare centers**		
	Model 1	Model 2	Model 2.2	Model 1	Model 2	Model 2.2
Female	0.98	0.98	0.98	1.02	1.02	1.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Non-White	1.15***	1.14***	1.14***	0.96*	0.98	0.98
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Age	0.92***	0.92***	0.92***	0.08***	0.08***	0.08***
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Parents' Hourly Income	0.98***	0.98***	0.98***	1.00*	0.99*	0.99*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Parents' Years of Schooling	0.92***	0.92***	0.92***	1.09***	1.08***	1.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Parents' Gender	0.93***	0.93***	0.93***	0.94*	0.94*	0.94*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Parents' Race	0.95***	0.95***	0.95***	1.04*	1.03	1.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Z-Score - Vulnerability Index		2.27***			0.26***	
		(0.31)			(0.06)	
Z-Score - Vulnerability Index - Human Capital Dimension			0.92**			0.74***
			(0.03)			(0.03)
Z-Score - Vulnerability Index - Infra-structure Dimension			1.04**			0.92***
			(0.01)			(0.02)
Z-Score - Vulnerability Index - Income and Labor Dimension			1.15***			1.27***
			(0.03)			(0.05)
Num. obs.	324015	324015	276733	276733		

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Source: Elaborated by the author.

4.6 Discussion and Limitations

As discussed in chapter 2, there was a substantial expansion in the municipal early education and care system in the past twenty years. This chapter analyzes if this expansion led to changes in the profile of registered and enrolled children. To do so, we implemented descriptive analyses of several child and family characteristics for different comparison groups. Logit models were used to understand which child and family characteristics affected the enrollment process.

Our findings indicate that the profile of children and families who registered in the São Paulo ECEC system changed after the recent expansion; more non-white children were registered in the system along with children whose parents had higher incomes and more years of schooling. These children, however, did not came from more vulner-

able neighborhoods as there was no significant change in the vulnerability of residential neighborhoods[27].

Taken together, our results indicate that the expansion of the ECEC system had mixed results in including a more diverse set of children. More non-white children were registered into the system along with younger children. There was, however, no indication that these children came from more vulnerable neighborhoods.

Another finding is that the profile of the children who were registered for daycare centers was different than those children who were registered for the preschool division of the ECEC system. Families who registered their child in the ECEC system before age three in the daycare center division are more likely to be non-white and to have higher incomes and more years of schooling. This pattern is maintained when only children who end up enrolling in the ECEC system are analyzed; those who enroll in the daycare center division have parents whose income and education level is higher than those who enroll in the preschool division. Once again, no difference is found for neighborhood vulnerability.

These findings allow us to conclude that the children who are first registered into the system in its preschool division are more vulnerable than the children who are first registered into the daycare center division. Several reasons could explain this pattern. Perhaps these parents have more access to information about the system itself or perhaps they understand that center-based care is important. These parents might also need center-based care more due to work or study commitments.

The findings also indicate that there is a difference among children who register but do not enroll, and those who ended up enrolling. The second group is more vulnerable, indicating that the less vulnerable families might be registering in the municipal ECEC system but not using it and opting perhaps, for the private system.

The analysis of the logit models indicates that parent's income and years of schooling are associated with the likelihood of enrolling in the ECEC system. However, the direction of such associations does not corroborate the findings described in the international literature. Our findings indicate that for enrollment in the ECEC system as a whole, the association is negative; an increase in income and education level reduces the odds of enrollment. However, this conclusion shifts when the enrollment into the day- care center division of the ECEC is analyzed. Here, a positive association is found; an increase in parental education levels increases the odds of enrollment in the daycare center division, a result in line with the international literature. More research is needed to better understand such results.

[27] This might be a consequence of the fact that there was only a slight expansion in the number of HDU present in our sample when we compare the 2010 and the 2018 sub-samples: from 1,325 to 1,352. It is also important to point out that this corresponds to almost the totality of the neighborhoods in the city; 1,593.

The vulnerability of the child's residential neighborhood is associated with enrollment in both divisions of the ECEC system. However, the direction of this association is not the same in both samples. When the whole ECEC system is analyzed, the association is positive; higher neighborhood vulnerability increases the odds of enrollment. When only the daycare center division is analyzed, the association becomes negative; higher neighborhood vulnerability decreases the odds of enrollment.

Our analyses have several limitations. First, as a result of the databases used, only some child and family characteristics could be analyzed. Important elements may have been left out of this analysis as a result. Marital status and number of siblings are two of the elements that are not part of this study but whose importance is noted in the international literature. Second, because of our sample selection process, which was a result of the data available, our results cannot be generalized to either São Paulo or even to children and families who register in the ECEC system. For example, only parents who were in the formal labor market were part of our sub-sample. Third, the logit models provide only an indication of the factors that are related to the choice of enrollment in the ECEC system. Their results cannot, and should not, be interpreted as causal but they do provide information on elements that are related to such choice. Fourth, it would have been important to understand the other care options available to parents in our sample as this would have allowed a more complete analysis of the choice process. Fifth, as in chapter 3, we used the HDU, an administrative division of the city, to define a child's residential neighborhood and this raises several questions which were discussed at length in the previous chapter.

Despite these limitations, we understand that our results provide valuable evidence on the profile of the children and their families who are benefiting from the ECEC system. Moreover, it also allows us to better understand which factors play an important role in the choice of enrollment into the municipal ECEC system. Taken together, our findings indicate that the profile of these beneficiaries of the ECEC system is changing but it does not seem to be including the most vulnerable. It also calls attention to the need to analyze in detail the characteristics of those entering the system to better guide public policy.

5 The impact of attending center-based care on cognitive skills

5.1 Introduction

There is inequality in access to center-based childcare. The poorest children have less access to childcare and when they do have access, it is of lower quality (RUZEK et al., 2014; BURCHINAL et al., 2008; LEVENTHAL; BROOKS-GUNN, 2000). Evidence for such inequality in the case of São Paulo was discussed in chapter 4.

The existence of such inequality is troubling as high-quality early childhood education and care can be an important tool in reducing the inequality of child development. There is evidence that attending center-based care has a positive impact on child development outcomes and later life results, specifically for children from disadvantaged backgrounds (BURCHINAL et al., 2008; DUNCAN; MAGNUSON, 2013). Moreover, there is evidence that the quality of this care is even more important for children from low-income backgrounds (DEARING et al., 2009; RUZEK et al., 2014).

In the past twenty years there has been a significant expansion in publicly funded ECEC systems in Brazil and Latin America. In 2015, between one-fifth and one-third of all Latin American children between the ages of zero and three attended childcare centers (BERLINSKI; SCHADY, 2015). In Brazil, from 2001 to 2015, the percentage of children younger than three years that were enrolled in daycare centers more than doubled: from 11% to 26%[1]. In 2019, 9 million children between zero and five years attended ECECs, a growth of 12.6% between 2015 and 2019 which was mainly due to an expansion in the enrollment of children between zero and three years which rose 23.2% in this period[2].

In the city of São Paulo, the expansion of center-based childcare was particularly substantial as discussed in chapter 2. An analysis of the effects of attending center-based childcare for children between zero and three years old in São Paulo is conducted in this study. To do so, the impact of center-based care enrollment on children's cognitive outcomes is estimated using a regression discontinuity design on a sample of eight-year-old children. In line with the literature, this study's main hypothesis is that attending center-based childcare has positive impacts on children's cognitive outcomes measured by achievement tests. We find evidence of a positive, marginally significant effect on the child's proficiency in mathematics on being offered a spot in *Berçário I*. No other impacts were found.

[1] Data from the Brazilian Household Survey (*Pesquisa Nacional por Amostra de Domicílios*)
[2] Data from the Technical Report of the 2019 Educational Census.

The remainder of this chapter is organized as follows. The next section presents an overview of the literature on the impact of early childhood care on child development. The research questions are presented in the third section. The fourth section briefly discusses the ECEC system in São Paulo, described at length in 1, and details its enrollment process. The data and the methodology are introduced in the fourth section, while the fifth section presents the results. The sixth section presents a discussion of the results and the limitations of this study.

5.2 The relationship between attending childcare centers and child cognitive development

Attending center-based care has positive impacts on child development, especially for those from more vulnerable backgrounds (CURRIE, 2001; CAMILLI et al., 2010; ENGLE et al., 2011; YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; BRITTO et al., 2017; MCCOY; WALDMAN; FINK, 2018). From randomized evaluations of small pilot programs such as the High Scope Perry Preschool Program and the Abcederian Project to newer quasi-experimental evidence from large-scale public preschool programs, there is robust evidence of positive and consistent short-term impacts on children's cognitive outcomes as measured by their results on achievement tests (WEILAND; YOSHIKAWA, 2013; DUNCAN; MAGNUSON, 2013; YOSHIKAW; WEILAND; BROOKS-GUNN, 2016). In examining a database of evaluation results from 84 programs, Duncan e Magnuson (2013) found a weighted average effect size[3] for early childhood education of 0.21 standard deviations on cognitive outcomes at the end of the treatment period. However, in the medium-term, this impact diminished; Duncan e Magnuson (2013) estimate a decrease in impact effect sizes of about 0.03 standard deviations per year.

Most of this evidence, however, is based on an analysis of preschool programs; programs for children between ages four and five. Less is known about the impacts on cognitive development of attending center-based care for infants and toddlers; children who are three years old or younger (RUZEK et al., 2014). There is reason to hypothesize that this impact could be different for younger children as the first couple of years of a child's life is a crucial period for the development of the relationship between the child and his/her main caregiver and receiving full-time low-quality center-based childcare might disrupt this process (BERLINSKI; SCHADY, 2015). Center-based childcare might also provide a more nurturing care than other, less formal childcare arrangements and there is evidence that corroborates such a hypothesis in the USA (FULLER et al., 2002). In the USA, evidence from experimental studies indicate large impacts on cognitive

[3] Evaluations had very different sample sizes and this weighted average was calculated by weighting the average treatment with the inverse of the squared standard errors in estimates (DUNCAN; MAGNUSON, 2013).

skills when pilot programs are analyzed such as the Abecederian Project as well as much smaller impacts when an at-scale program, the Early Head Start, is evaluated (DUNCAN; MAGNUSON, 2013). Non-experimental evidence from an American nationally representative database also indicates small positive effects of attending center-based childcare on children's cognitive outcomes but only for those children who attended medium- or high-quality centers (RUZEK et al., 2014).

In the UK, there is evidence that attending center-based childcare between the ages of zero and four is associated with higher cognitive skills (LOEB et al., 2007). Evidence from Germany indicates that attending center-based childcare has positive impacts on child development and that children from more vulnerable backgrounds benefited most (FELFE; LALIVE, 2013). In Norway, a large-scale expansion of childcare subsidies was studied by Havnes e Mogstad (2015) and the authors found large and positive long-term impacts on educational outcomes and future earnings.

In Latin America, evidence is much more mixed. In Colombia, analysis of a high-quality and intensive childcare program for toddlers finds positive impacts on children's cognitive outcomes (NORES; BERNAL; BARNETT, 2019). An evaluation of another program in Colombia based on an expansion of large daycare centers for children between ages zero and five, found a significant and negative impact on cognitive outcomes (BERNAL et al., 2019). In Chile, an evaluation of a publicly-funded childcare expansion found a negative effect on cognitive outcomes of children at two years of age (NOBOA-HIDALGO; URZúA, 2012). Another analysis of the Chilean case finds evidence that attending center-based care between ages two and four had positive effects on cognitive outcomes measured when children are ten years old and that children from middle-income backgrounds are those who benefit most (CORTáZAR, 2015). In Brazil, a randomized evaluation of the publicly funded daycare system in the city of Rio de Janeiro found evidence of positive impacts on children's cognitive outcomes (LIMA, 2019).

Taken together, this evidence indicates that in developed countries, attending center-based childcare before the age of four has positive impacts on children's cognitive outcomes but that these findings are not replicated in the Latin American context. This can be attributed to differences in either social policy and labor laws, or cultural differences such as living with extended family. All of these aspects can modify the type of care a child would receive if she/he was not in center-based care, counterfactual conditions for children in the different control groups.

However, this difference in findings can also be a consequence of the quality of care. Quality in early childhood development literature is generally defined in two dimensions; structural and process (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; BERLINSKI; SCHADY, 2015). Structural quality is related to resources present in the childcare center including human resources and process quality refers to the quality of interactions,

whether between children and teacher or among children (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016). There is evidence that the quality of center-based childcare for toddlers and infants is generally lower, a possible consequence of the fact that it is more expensive to provide high-quality care for the youngest children as they require more individualized attention (RUZEK et al., 2014).

The literature suggests that quality is an important determinant for the impact of center-based care on child development, whether in high-income countries or not (YOSHIKAW; WEILAND; BROOKS-GUNN, 2016; BRITTO et al., 2017; ENGLE et al., 2011). In the Latin American context, there is some evidence that the care in centers is of low quality (BERLINSKI; SCHADY, 2015). In the specific case of Colombia, Bernal et al. (2019) point out that the lack of quality in a center might be a plausible explanation for the results that were found as the analyzed centers did not have a structured curriculum nor any pedagogical orientation for the teachers (BERNAL et al., 2019).

5.3 The current study

There is mixed evidence on the impact on children's cognitive outcomes of attending center-based childcare before the age of four. The current study seeks to contribute to this literature by estimating the impact of enrolling in the publicly funded and free center-based childcare program in São Paulo on children's cognitive outcomes as measured by their proficiency in mathematics and Portuguese language in second grade. To do so, an analysis of data from the São Paulo municipal ECEC system, combined with data from a standardized test taken by all second-grade students who attend municipal schools is conducted in order to answer the following research questions:

- Did enrollment in center-based childcare before age four contribute to children's cognitive outcomes in primary school as measured by children's proficiency in mathematics and Portuguese language in second grade?
- Did this impact vary by cognitive domain?
- Did this impact vary by the number of years of care the child could potentially receive?

5.4 Center-based care in the city of São Paulo

The ECEC system in São Paulo was described in 2 of this thesis. In this study, the focus is on analyzing the center-based childcare for children between zero and three years of age - in daycare centers. In this division of the system, children are enrolled, according to their birthday, in four levels: *Berçário I*, *Berçário II*, *Mini-grupo I* and *Mini-grupo II*.

Independently of the child's age, children stay in the centers for up to 10 hours a day and they receive nutritious meals and participate in educational activities.

As discussed in chapter 2, in the past twenty years, there was an expansion in the number of open slots for this age group across the city of São Paulo.This expansion, however, was combined with large waiting lists for many years; a fact also described in chapter 2. Specifically, in the years analyzed in this study, 2010 to 2014, the size of these waiting lists was at their peak as shown in Figure 2.9. This fact, combined with the system's enrollment process, is central for the proposed identification strategy. In the next subsection, the system's enrollment process is explained in detail. The identification strategy is defined in the next section.

The enrollment Process To enroll a child in the São Paulo ECEC system, a parent[4] needs to go to an educational unit[5] and present basic identification documents and a proof of address[6]. This information is uploaded to a centralized system that automatically assigns the child a unique code and places her/him on waiting list queues according to her/his age and to her/his residential postal code. The system also records the date when the child was registered[7].

Waiting lists are not only specific for each daycare center but also specific to a child's age. Within a center, there can be waiting lists for the four different levels as defined in the previous section. For example, for the center *Yellow Home*, there can be one waiting list for *Berçário I*, another for *Berçário II* and two for each *Mini grupo*, I and II. The system uses the children's year and month of birth information to define the level they are eligible for[8]. TThe system then places children on waiting lists queues in all centers within a two kilometers radius of their home[9] that have these levels available[10]. In all waiting lists, children are identified by the same code[11] which defines a child's place in the queues. As the code depends on the date and time of request, children whose spots were requested first have smaller codes than children who requested later. That is, they are further ahead in the queues. For example, supposing a parent registers their child on the 7th of December of 2010 at 1:30PM and another parent registers their child on the same day, but at 2PM, the code for the first child would be smaller than the code for the

[4] Or the adult legally responsible for the child.

[5] Any center that is part of the municipal educational system, including primary schools.

[6] In this subsection, we describe how the system operated between 2010 and 2014 when the children analyzed could potentially be enrolled. There have been slight modifications in more current years; reducing the distance from the child's home to the daycare centers and implementing priority enrollment for children from vulnerable backgrounds. We do not consider these changes.

[7] This information is crucial for the empirical strategy used in this study. From here on, this date is defined as the date of registration.

[8] This is defined by city ordinances that are published annually.

[9] As defined in the following ordinances: Nº 4.801, Nº 5.033, Nº 5.741, Nº 6.542.

[10] There might be centers that do not offer all the levels.

[11] It is important to point out that this is not the child's ID code used in this thesis to merge the different databases.

second child, indicating that the first child should be in front of the second child in the waiting list.

Spots can open at any time of the year. However, the ECEC system follows the official school calendar and has classes from February to December, and therefore most enrollment occurs at the beginning of the school year prior to start of classes. When a spot opens up, the center has between three to five days to contact the next child on the waiting list; that is, the child with the smaller code on their waiting list. The date when this happens is also registered into the system. The parent then has then between five and ten days to accept the spot and enroll their child[12]. If the family does not enroll the child in ten to 15 days, he/she is removed from the waiting list. The family can also refuse a spot and opt to stay on the waiting list[13]. In this case, the child's code stays the same and because of that, the child does not go to the end of the waiting list, as her/his code does not change. IIf a child moves and needs a spot in a new daycare, they are again placed on the waiting list, but with the same code meaning that they are generally closer to the beginning of the list and do not have to wait for a spot for very long[14].

At any point in time, the parent can request to drop out of the queue. In this case, if they decide to request a spot again later on, their child's code stays the same in the system[15]. The parent can also check their child's spot in the waiting list queues through the Department of Education website or by going to any educational unit.

There are some instances in which this normal process is not followed. First, the parent can request a spot in a specific center and in this case, the child is placed only on its waiting list. Second, there are two instances that allow a child to "skip" the waiting list[16]: (i) judicial orders[17]; and (ii) children with special needs. These children have priority in enrollment, and they are offered spots as soon as possible.

[12] This is specified in the ordinances indicated previously. Again, the date of enrollment is also recorded into the system

[13] This might happen because, for example, a baby is called before the end of his mother's maternity leave; the mother can give up the spot and wait to be called again. The date of refusal is also recorded into the system.

[14] More information on all of these procedures, can be found in the ordinances indicated in footnote 9.

[15] This allows us to identify the different requests by child.

[16] Currently there is a third reason: extreme poverty. This was defined by ordinance Nº 6.770 established in 2014.

[17] There was a substantial judicialization of daycare enrollment in São Paulo. In 2014 for example, there was a waiting list of 160 thousand children and a request that 14,400 be enrolled as a result of judicial order (COSTA, 2016). For an analysis, see Costa (2016) and Oliveira, Silva e Marchetti (2018).

5.5 Method

5.5.1 Data

The current study draws on two principal data sources and two supporting databases. The main ones have already been discussed in this thesis: the administrative database from the ECEC enrollment process and the database for the 2018 *Provinha São Paulo*. The other two databases are: (i) a database from the municipal Department of Education with information on children who were registered into specific waiting lists, (ii) a database with all the ECEC centers in 2018[18].

The administrative database was described in chapter 2 and chapter 3. For this study, this database provides two crucial pieces of information; the first date the child was registered into the system and their residential postal code during this registration. This database also provides information on the child's personal characteristics such as their month and year of birth, their race and their sex[19].

The database from *Provinha São Paulo* was presented in chapter 3 and provides information on children's proficiency in Portuguese language and mathematics in second grade. It also indicates if the child had any special needs. The merge between this database and the previous one was implemented through the child's ID code[20].

The third database was used to identify requests that were made into a specific school's waiting lists[21]. This database had three main pieces of information; the child's ID code which is the same as the previous two databases, the time of entrance on to a specific waiting list and the time of exit from this waiting list. To merge this database with this study's sample, first, a match by the child's identification number was conducted. However, as there was more than one request per child in both databases, this was not sufficient[22]. We opted to merge the databases using the child's identification number and the date of request (to a spot in our original database or to a specific school's waiting list in the other database). We did not use the complete time of request because these did not match precisely and children were generally first registered into the system and then within a few minutes, on to the specific waiting list. The difference between the time of

[18] We used the version from 2018, as we wanted to understand the wait-lists in years between 2011 and 2014 and the 2018, was the oldest database avaliable. Available at: *http* : *//dados.prefeitura.sp.gov.br/dataset/cadastro − de − escolas − municipais − conveniadas − e − privadas*.

[19] As previously discussed, there could be multiple requests for the same child. It is also important to point out that children were present in the administrative database if there was some kind of update in the child's registration and enrollment process during the years of 2010 and 2018. The child could have registered or could have enrolled or could have been offered a spot, for example.

[20] This was generated by the municipal Department of Education and does not correspond to the child's official identification number.

[21] This information was not part of the version of the administrative database originally shared with us. We therefore had to solicit it from the municipal Department of Education in a separate database.

[22] As discussed previously, children could be registered into the system several times.

requests in both databases was then analyzed. If there was a large time difference of over 6 minutes between the time of request for a spot and the time of request for a specific waiting list, the request was not tagged as a request for a specific waiting list. Requests into the enrollment system were also not tagged as such if they happened after the child had exited the specific waiting list.

The fourth database was used to identify the existing ECEC centers that were part of the ECEC system in São Paulo and their precise locations. This database reflects the availability in 2018[23].

5.5.2 Analytic Plan

To answer our research question, our analysis implemented a regression discontinuity design using waiting list queues in the municipal ECEC system to construct a running variable. Two outcomes were analyzed: proficiency in mathematics and in Portuguese language. The design was implemented in four different samples, discussed at length in the next subsection.

To illustrate the idea behind the design, suppose there are two families who want to enroll their children in the same level for the following school year. Suppose one registers their child on a certain day and the other, one day later. By the end of the year, the child whose parent registered them first is offered a spot and the other child is not. Our hypothesis is that these children have very similar characteristics in addition to the fact that one of them received an offer and the other one did not. By comparing the outcomes of these very similar children, we can estimate the average treatment effect of being offered a spot in the ECEC system.

This is possible because there is an abrupt change in the probability of receiving this offer at a certain threshold. Children registered after a certain registration date, the threshold, do not receive offers and children who were registered just a day before do receive such an offer. Although the decision to register on waiting lists is endogenous as well as date of registration into the system, both are related to the parent's and child's characteristics; parents cannot predict the threshold nor manipulate their place in line. Parents do not know the size of all the waiting lists prior to registration[24] nor do they know the number of spots that will open in the following year for the level they are interested in at the schools they are on the waiting list for[25]. Because parents are unable

[23] This was the oldest database publicly available to us. It would have been interesting to use the database of each one of the school years analyzed - 2010 to 2014 - because there is a change of availability between years.

[24] Anecdotal evidence from interviews with mothers confirm that even in 2018, the size of the waiting lists were not common knowledge. Some knew they were large, but not the exact size or the expected waiting times for a spot.

[25] They learn the waiting list their child are in at the moment of registration, but the number of spots depends on the number of children who would continue from one age group to the next and if new

to manipulate, the discontinuous change in the treatment assignment can be used to learn the local causal effect of the outcome of interest by comparing children who were offered a spot, but barely, with children who were just missed being offered a spot.

A regression discontinuity design has three main elements: the running variable, the treatment assignment and the cutoff. In this type of design, the treatment depends on a known rule, which in turn is a function of the running variable and the cutoff. In our analysis, the treatment assignment is the offer of a spot in the daycare center system for a certain level, school year and center. Conceptually, the running variable is the distance between a child's date of registration and last day the child could have applied and still be offered a spot. The cut-off, because of the way the running variable was constructed, is zero for all children; if the running variable is equal or superior to zero, the child is assigned treatment.

To construct the running variable, we use data from the administrative database and apply an algorithm that is meant to reproduce the enrollment procedure discussed in section 5.4. The algorithm is intended to reproduce the automatic enrollment procedure as best as possible with the information we have available and following the procedures listed in the legislation. We did not have information for each child on the waiting list queues they were allocated to, nor did we know what the results of each queue were although this which would have allowed us to re-construct the queues. The algorithm is detailed below. It is first applied to the administrative database and then to the specific samples, whose construction is detailed in the next subsection.

- Using the administrative database[26] and considering only the first request for each child:

    1. For a center j, a level k and a school-year a, we identified the child who was offered enrollment before the 31st of December of the year $a-1$[27] and who had the latest date of registration, D_i. That is, we ranked all the children who were offered enrollment for the school-year a and for level k by their D_i and identified that with the most recent D_i.

        a) The date of registration of this child was defined as the center j cutoff date for the level k and school-year a: D_{jka} ;

- Using this study's samples:

spots would open.

[26] For this, children who had any special priority in enrollment or who applied only to a specific school where not considered.

[27] We tried to analyze the dates in which most enrollment referral for the following year happened. There was no clear pattern for the different age-groups. Because of that, we decided to use the 31st of December as our cut-off, based on the hypothesis that centers want to begin the school year with their full enrollment.

1. For each child i in the sample, we identified the 31 closest centers[28] to their homes[29].
2. For each child i and each of the 31 closest centers, we identified the cutoff date, D_{ka}, considering the level k and the school-year a, defined by each sample[30,31].
3. For each child i, we compared their date of registration - D_i - with the cut-off date for each of 31 centers:
   a) If $D_i \leq D_{jka}$, the child was offered a spot in center j for level k and school-year a;
   b) If $D_i > D_{jka}$, the child was not offered a spot in center j for level k and school-year a;
4. For each child i, we calculated the running variable: $Z_i = D_{jkai}$ - D_i in which D_{jkai} was
   a) the D_{jka} of the closest center to the child's home that offered him/her a spot;
   b) the D_{jka} of the closest center to the child's home that have a valid cutoff date[32] if the child was not offered a spot at any centers.
5. For each child i, we assigned the treatment, T_i based on the running variable:
   a) If $Z_i < 0$, then $T_i = 0$.
   b) If $Z_i \geq 0$, then $T_i = 1$.

To illustrate the first section of this algorithm, let's use one example extracted from the administrative database. For the level *Mini-Grupo I* and for the school year of 2014, in school 307771, seven children were offered spots. The latest date of registration, among those seven children, was 11/7/2013 - all other children were registered beforehand. This date became then the cutoff date for school 307771, for the level *Mini-Grupo I* and for the 2014 school year.

[28] We choose to analyze 31 schools because, for 75% of the sample, there were up to 31 schools within a 2 km radius from the centroid of their home's postal code.

[29] To do so, we used the residential postal code in their registration file and the centers addresses. This last information came from the fourth database, described in the previous section.

[30] Each of the four samples corresponded to specific levels and school-years, as detailed in the next section.

[31] In some cases, there was no cut-off date. This meant there were no children who were offered a spot in that center for that level and for that school year before the 31st of December of the previous year. This can happen for several reasons. First, we might be allocating children on non-existent waiting lists as the database with the addresses was from 2018 and might indicate centers that did not exist in previous years. Second, there might be no new spots open as all children stayed from one level to the next in the following year. Third, all the spots that existed were offered after the 31st of December of the previous year.

[32] As discussed in the previous footnote, there might not be a cutoff date for this specific combination of center/school-year/level.

To detail the model estimated, consider a child whose parents have registered her into the system. Using the potential outcomes framework, let Y_i denote the outcome of interest, the proficiency in Portuguese language and mathematics, and T_i, the assignment to treatment, the offer of a spot in the municipal daycare center system -, for child i. Each child has two potential outcomes, $Y_i(1)$, the outcome that would be observed under the treatment, and $Y_i(0)$, the outcome that would be observed under control conditions. As defined previously, T_i, the assignment for treatment, is a function of the running variable, Z_i and the cutoff[33]:

$$T_i = \begin{cases} T_i = 1, & \text{if } Z_i \geq 0 \\ T_i = 0, & \text{if } Z_i < 0 \end{cases} \tag{5.1}$$

The observed outcome is:

$$Y_i = (1 - T_i) * Y_i(0) + T_i * Y_i(1) = \begin{cases} Y_i(1), & \text{if } Z_i \geq 0 \\ Y_i(0), & \text{if } Z_i < 0 \end{cases} \tag{5.2}$$

We cannot observe both $Y_i(0)$ and $Y_i(1)$ for the same child i: the child is either under control or under treatment conditions. Moreover, we can only observe $Y_i(0)$ for children with strictly negative running variables and $Y_i(1)$ for children with zero or positive running variables. This is the fundamental problem of causal inference (CATTANEO; IDROBO; TITIUNIK, 2019).

Following on from the previous equation, we can write the observed average outcome as:

$$\mathbb{E}[Y_i|Z_i] = \begin{cases} \mathbb{E}[Y_i(1)|Z_i], & \text{if } Z_i \geq c \\ \mathbb{E}[Y_i(0)|Z_i], & \text{if } Z_i < c \end{cases} \tag{5.3}$$

The intention-to-treat effect could then be calculated by:

$$ITT = \mathbb{E}[Y_i(1)|Z_i = z] - \mathbb{E}[Y_i(0)|Z_i = z] \tag{5.4}$$

Again, the fundamental problem of causal inference presents itself: we can only observe $\mathbb{E}[Y_i(1)|Z_i]$ for those to the right of the cutoff and $\mathbb{E}[Y_i(0)|Z_i]$ for those to the left of the cutoff. It is not possible to observe $\mathbb{E}[Y_i(0)]$ and $\mathbb{E}[Y_i(1)]$ for the $Z_i = z$. In the context of the RD design, however, if we analyze those children with Z_i very close to the cutoff, it is possible to approximately calculate the average treatment effect using observed outcomes (CATTANEO; IDROBO; TITIUNIK, 2019). The reasoning behind this relies on the assumption that children with Z_i exactly equal zero, the cutoff in this study, and children with Z_i barely below zero are very similar except for their treatment status, that

[33] Because of the way the running variable was constructed, the cutoff, c, is equal to zero.

changed abruptly at the cutoff; comparing their observed outcomes would, then, provide an estimation of the average treatment effect (CATTANEO; IDROBO; TITIUNIK, 2019). To do so, we follow the literature and use a local linear point estimator:

$$\tau_{RD} = \hat{\mu}_+ - \hat{\mu}_- \tag{5.5}$$

in which $\hat{\mu}_+$ is an estimate of the point $E[Y_i(1)|Z_i = 0]$ and $\hat{\mu}_-$ is an estimate of the point $E[Y_i(0)|Z_i = 0]$. Both of this estimates are calculated by fitting local linear weighted regressions in a neighborhood defined by a bandwidth that optimizes the mean squared error of the local point estimator[34]. To find $\hat{\mu}_+$, a local linear regression is fitted for observations above the cutoff - that is to the right of the cutoff - and its intercept is the estimation we are interested in. To find $\hat{\mu}_-$, a regression, with the same parameters, is fitted for the observations below the cutoff - that is to the left of the cutoff.

The choice of a local linear regression follows the literature that considers that this estimator delivers "a good trade-off between simplicity, precision, and stability in RD settings" (CATTANEO; IDROBO; TITIUNIK, 2019). The observations are weighted by a triangular kernel function [35]. By using this type of kernel function, with a bandwidth that optimizes the mean squared error, we estimate point estimators that are both consistent and optimal (CATTANEO; IDROBO; TITIUNIK, 2019).

The running variable is a discrete variable, as it can only take a finite number of values. Therefore, there are values of the running variable that are shared by many observations (mass points). Our estimation method[36] is able to accommodate this as the number of mass points is large for all four samples[37].

A local point estimator is estimated for four different samples that are presented in the next subsection. As discussed in this section, we wanted to re-construct the enrollment process and to do so we had to identify which children could participate in the enrollment for different levels and school years. As children could stay on the waiting list for more than one school year, it was not enough to select the children who requested a certain level for a certain school year and therefore it was necessary to make certain assumptions and construct different samples; a process detailed in the next section.

[34] This approach optimizes the bias-variance trade-off that exists when selecting a bandwidth to define the neighborhood in which to estimate the local polynomial regressions (CATTANEO; IDROBO; TITIUNIK, 2019).

[35] This guarantees that observations closer to the cutoff receive more weight than those further from it.

[36] Our analysis was implemented using the command *rdrobust* in R.

[37] For 2011, in the sample with non-missing math scores, the number of mass points was 279 and in the sample with non-missing Portuguese language scores, 276. For 2012, the number of mass points was 857 for the sample with non-missing math scores and 853 for the sample with non-missing Portuguese language scores. For 2013, in the sample with non-missing math scores, the number of mass points was 394 and in the sample with non-missing Portuguese language scores, 396. For 2014, there were 393 mass points in the sample with non-missing math scores and 391 in the sample with non-missing Portuguese language scores.

The analysis is also done separately for each of the two outcomes: children's proficiency in Portuguese language and mathematics. These measures were assessed together in the October of the child's second grade in 2018 and were part of the *Provinha São Paulo* database.

As discussed previously, the local point estimator should be understood as an estimator of the "intention-to-treat" effect, the effect of the treatment assignment. That is, up to this point, we were analyzing the effect of the offer for enrollment in the daycare centers of the ECEC system on children's proficiency.

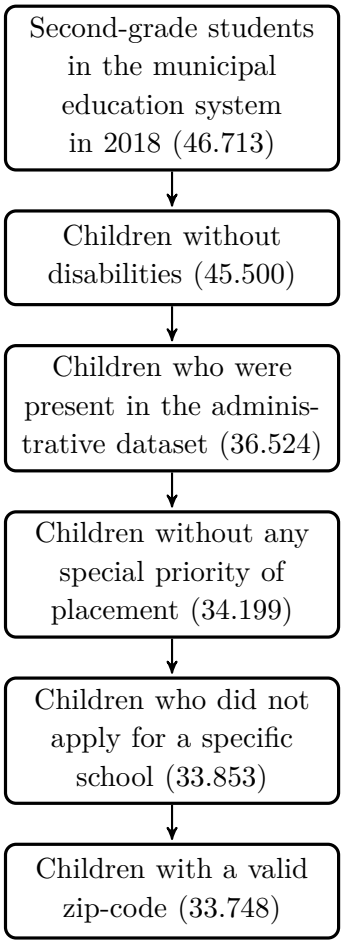
However, being offered a spot does not guarantee enrollment as parents might decide to not accept the spot because of outside care options such as a private daycare center or the fact that the mother is still on maternity leave. Additionally, children might leave the daycare center during the year, opening up new spots and children who were not called in at first would then be offered the spot. Children can also dropout after a while for whatever reason. There is no perfect compliance to treatment as those assigned to treatment do not necessarily receive treatment and those not assigned do not necessarily not receive treatment. Because of this, the ITT should not be understood as an unbiased estimator of the average treatment effect.

5.5.3 Sample

To define the sample for this study, following the procedure implemented in chapter 3, the two databases - presented in the previous section - were merged[38]. Several restrictions were then implemented as part of the sample selection process. Figure 5.1 describes this process.

[38] As discussed previously, an anonymized ID code was used for this merge. Children could be present more than once in the administrative database not only if their status in the system might have been updated more than once, but also children might have been registered more than once. After the merge, the 36,524 children who were present in both databases corresponded to 72,863 unique requests. More than 50% of these children had more than one request registered in the administrative database.

Figure 5.1 – Sample Selection



Source: Elaborated by the author.

Children who had any kind of disability were removed from our sample[39]. Requests with any kind of special priority in placement [40] or for placement in a specific school were also not considered[41]. Finally, requests without a valid residential postal code were not considered[42]. After all of these restrictions, the final sample was composed of 33,748

[39] There were 1,213 children with disabilities in the *Provinha São Paulo* database - 2.5% of the sample. Among the type of special needs, autism and intellectual disability were the ones with the biggest prevalence.

[40] 7,675 requests had some kind of priority in placement and were taken out of the sample as they were not part of the usual placement process. This represented 10.5% of total requests. When there is priority in placement, the enrollment process does not take the steps discussed previously and replicated in this study.

[41] 3,685 unique requests were taken out of our sample due to this or 5.65% of total requests. When a specific school is requested, children do not take part in the usual placement process and are only placed on one waiting list, that of the specific school and therefore, those students are not considered in this study. The merge between the sample and the database for specific waiting lists identified 3,767 possible requests for placement on a specific waiting list. We then analyzed the time of request for a spot on the specific waiting list to check if it was the specific request that was for a place on a specific waiting list and not another request for the same child on the same day. 82 unique requests were eliminated after this analysis as they had large time differences of over 6 minutes between the time of request for a spot and the time of request for a specific waiting list or they happened after the child had exited the specific waiting list.

[42] 243 requests were taken from our sample because of their residential postal codes were not valid.

children[43]. The sample varied by dependent variable; 29,731 children when we analyze Portuguese language to 29,732 children when we analyze mathematics[44]. T-tests were used to analyze the restriction process: only the first two steps significantly modified the sample in a statistically significant way[45].

To allow for the implementation of the regression discontinuity design - discussed in the next section and the main empirical strategy of this study -, the sample was further divided into four sub-samples:

- 2011 Sample: children registered into the system who were able to participate in the enrollment process for *Berçário I* in 2011;
- 2012 Sample: children registered into the system who were able to participate in the enrollment process for *Berçário I* in 2012 or *Berçário II* in 2012;
- 2013 Sample: children registered into the system who were able to participate in the enrollment process for *Berçário II* in 2013 or *Mini-grupo I* in 2013;
- 2014 Sample: children registered into the system who were able to participate in the enrollment process for *Mini-grupo I* in 2014 or *Mini-grupo II* in 2014;

The procedure to define such sub-samples had three main steps, illustrated in Figure 5.2. First, we identified the children who could participate in the enrollment process for a certain combination of age group[46] and potential year of enrollment[47]; that is, children who had the appropriate age for that age group[48] and who were registered into the system before the beginning of that potential year of enrollment[49] and who were still

There were 112 postal codes that were not valid: most of them did not have the necessary eight digits or indicated a missing information (1000999).

[43] The sample selection process, detailed in this paragraph and the respective footnotes, considered a database with repeated observations. As discussed at length, children could be present in the database more than once. Therefore, the final sample had 33,748 children and 61,260 unique requests. Only the first request for each child was kept in the sample.

[44] As discussed previously, there were children present in the database who did not take the test.

[45] T-tests were used to analyze if the proficiency in mathematics and Portuguese language had different means between the samples.

[46] In chapter 2,we discussed the different age groups that were part of the São Paulo ECEC. In this study, we were only interested in the age groups for children between zero and three years old. It is important to point out that when children are registered into the system, it places them into age groups according to their date of birth. However, if they stay on the waiting list for more than one school year, the age group they would be enrolled in changes, and this is not updated in the database. Therefore, we had to manually allocate children into age groups according to their date of birth and the legislation at the time.

[47] The term *potential* indicates that children were registered into the enrollment process and could potentially be offered a spot for that school year. It indicates that this result is not guaranteed as not enough spots were available and there were waiting lists.

[48] This information came from several ordinances from the municipal Department of Education of São Paulo - Nº 4.801, Nº 5.550, Nº 5.033, Nº 5.741, Nº 6.542, Nº 6.123, Nº 6.811, Nº 5.506, Nº 7.858.

[49] Children could have been registered in the year before or several years before. In this step, we did not consider the children's final outcome in the process - only their date of registration.

in the waiting-list[50]. For example, when we were analyzing the children who could have applied to be in *Berçário II* in 2013, we selected only children whose parents requested a spot before 2013, who were not taken out of the waiting list in the previous years, and who were born after 1st of January of 2012[51]. To define the combinations of age group and potential year of enrollment that would be analyzed, several ordinances from the city's Department of Education of São Paulo were analyzed to define, for each age group in the daycare center division, the months and the years of birth of those allowed to enroll[52] for each age group in the daycare center division. This information was combined with the children's expected date of birth[53]. This allowed us to conclude that if the children in the sample attended the daycare center division of the ECEC system, they should have been enrolled in the following age groups for the following school years: *Berçario I* in 2011 and 2012, *Berçario II* in 2012 and 2013, *Mini-grupo I* in 2013 and 2014, and *Mini-grupo II* in 2014. We did not analyze potential years of enrollment after 2014 because a change in the legislation[54] meant that children in our sample were old enough to attend preschool, which has mandatory attendance since 2009[55] and it is not the focus of this study. This procedure resulted in the division of the original sample of 33,748 into seven combinations of age group and potential year of enrollment[56].

The second step in the sample division procedure was the aggregation by potential year of enrollment. Independent of the age group the child was enrolled in, children who enrolled in the same year had the same amount of exposure to the ECEC system, as they all entered first grade in 2017[57]. Therefore, it made sense to aggregate the seven combinations according to the year of enrollment. For example, children who enrolled in 2013, for either *Berçário II* or *Mini-grupo I*, were part of the same sub-sample.

The final step involved taking out of the sub-samples observations that did not have a valid running variable[58].

[50] Children were considered still in the waiting-list if they had not given up a spot or enrolled in the previous school years.

[51] This came from ordinance Nº 5.741 from the Municipal Educational Department.

[52] This changed between 2010 and 2018 several times and because of that, several ordinances had to be analyzed. All the analyzed ordinances are in the footnote 9.

[53] As children in our sample were in second grade in 2018, according to the Brazilian legislation, they should have been born between the 1st of April of 2010 and 31st of March of 2011.

[54] Ordinance N. 6123 from October of 2014.

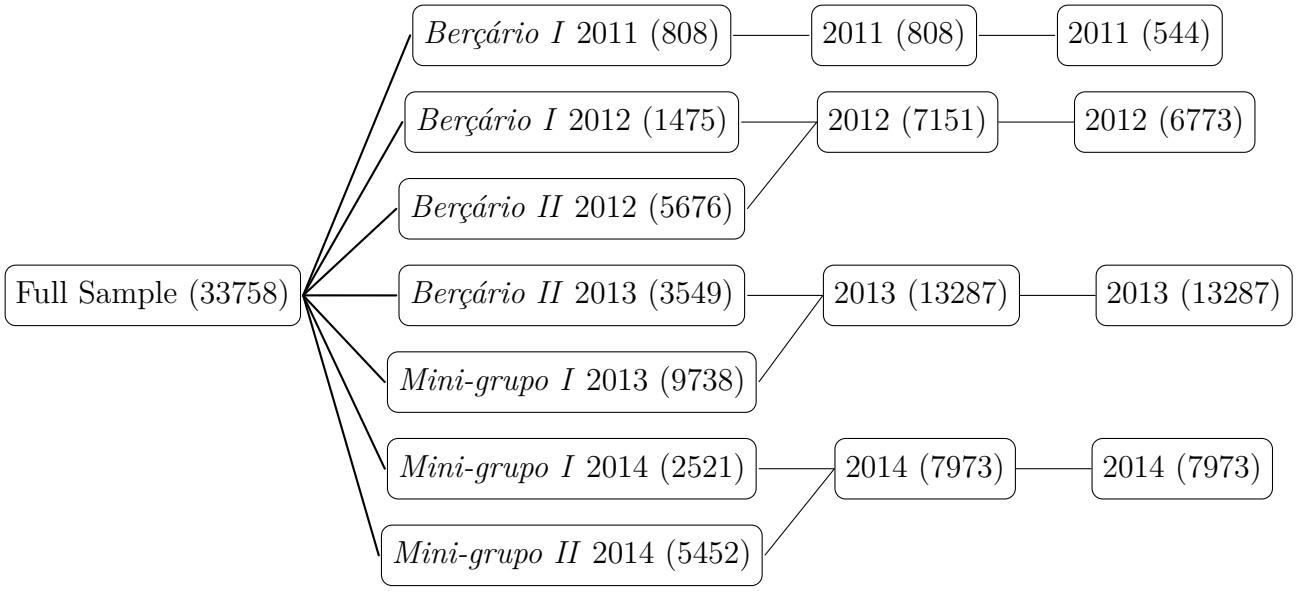
[55] The original law is *Emenda Constitucional* n.9 and it was implemented through Lei Nº 12.796, which gave until 2016 to all cities to comply.

[56] Children, in our sample, could not be in any of the age groups, as discussed in the end of this section. Also, children could be in multiple sub-samples.

[57] This is a consequence of the difference in cut-off dates between the ECEC system and the primary school system, both within the municipal education system. This difference was corrected in later years, but not for the sample we analyze.

[58] In the previous section, the construction of the running variable was explained in detail. In this step, we removed from the sub-samples the children we could not construct a running variable for. The fact that only in the sample for the school-years of 2011 and 2012 children are eliminated indicates that not being able to construct a running variable is probably a consequence of us using an newer database of centers. Unfortunately, older versions of the centers' database were not available to us.

Figure 5.2 – Sample Division Process



Source: Elaborated by the author.

Not all children from the original sample were part of the sub-samples depicted in Figure 5.2: 14,992 children were not part of any sub-sample. Of those, 59.48% applied after 2013 - in 2014, 2015, 2016 or 2017. Children who did apply before or in the 2013 school year were either removed in the restriction process as depicted in Figure 5.2, or were not on the waiting list queues at the appropriate time or did not have the appropriate age to be in the age groups analyzed[59]. The same child could also be part of multiple sub-samples[60].

5.6 Results

5.6.1 Descriptive Statistics

To understand if the children in our sub-samples were different from the universe of children who took the *Provinha São Paulo*, we did a two sample t-test for each sub-sample. The results indicate that there was a statistically significant difference between the sub-samples of 2011 and 2012 and the complete set of second graders for both mathematics and Portuguese language proficiency[61]. For the sub-samples of 2013 and 2014, there was

[59] 3.31% of these children were born before 2010.

[60] 29.87% of the total sample was part of only one sub-sample, 22.97% of two, 2.93% of three and 0.02% of all four.

[61] For the t-test between the 2011 sub-sample and the complete database, the p-value was 0.00 when mathematics proficiency was analyzed and 0.01 when Portuguese language proficiency was analyzed. For the t-test between the 2012 sub-sample and the complete database, the p-value was 0.0 when mathematics proficiency was analyzed and 0.04 when Portuguese language proficiency was analyzed.

no evidence of a statistically significant difference in either mathematics and Portuguese language when compared with the full set of second-graders who took the test in 2018[62].

Sample descriptive statistics are provided in Table 5.1 for each one of the sub-samples separately. It is important to point out that around one third of each sub-sample had missing data on race[63]. There is also a significant portion of children in each sub-sample for whom we do not have information on mathematics or Portuguese language proficiency.

The last column in this table presents the p-value of two different tests used to check if the results varied according to the sample analyzed. For the continuous variables, the Kruskal-Wallis H test was used to test if at least two samples were different. As all of the p-values of the Kruskal-Wallis H tests were inferior to 0.1, we can conclude there is a statistically significant difference in all the continuous variables between the different samples. For categorical variables, the Pearson's Chi-squared test was used and for race the result indicates that there is a statistically significant difference between sub- samples although the same cannot be said for sex. Considering all these results, we concluded there is evidence that the sub-samples are very different. All the results in the rest of this study are separated by sample.

[62] For the t-test between the 2013 sub-sample and the complete database, the p-value was 0.27 when mathematics proficiency was analyzed and 0.15 when Portuguese language proficiency was analyzed. For the t-test between the 2014 sub-sample and the complete database, the p-value was 0.16 when mathematics proficiency was analyzed and 0.08 when Portuguese language proficiency was analyzed.

[63] We classified the children's race as non-white when it was categorized as black, brown, yellow or indigenous. This information was collected when children first registered with municipal ECEC system. There was a high number of children without a defined race and for these children, the variable non-white was missing.

Table 5.1 – Descriptive Statistics

	2011 (N=544)	2012 (N=6773)	2013 (N=13287)	2014 (N=7973)	p value
Sex					0.232[1]
Female	279 (51.3%)	3467 (51.2%)	6767 (50.9%)	3961 (49.7%)	
Male	265 (48.7%)	3306 (48.8%)	6520 (49.1%)	4012 (50.3%)	
Race					< 0.001[1]
White	188 (34.6%)	2382 (35.2%)	4666 (35.1%)	2714 (34.0%)	
Non-white	189 (34.7%)	2468 (36.4%)	4661 (35.1%)	2586 (32.4%)	
(Missing)	167 (30.7%)	1923 (28.4%)	3960 (29.8%)	2673 (33.5%)	
Age					< 0.001[2]
Mean (SD)	7.93 (0.25)	7.62 (0.48)	7.55 (0.50)	7.52 (0.50)	
Prof. in Math					0.015[2]
N-Miss	64	677	1362	810	
Mean (SD)	162.73 (42.36)	157.14 (43.08)	156.35 (42.68)	156.63 (42.43)	
Prof. in Port. Lang.					0.028[2]
N-Miss	63	745	1417	881	
Mean (SD)	160.67 (36.13)	157.06 (37.89)	156.52 (38.19)	156.82 (38.74)	
Running Variable					< 0.001[2]
Mean (SD)	-58.44 (116.05)	-30.70 (193.94)	274.24 (185.73)	361.06 (252.48)	

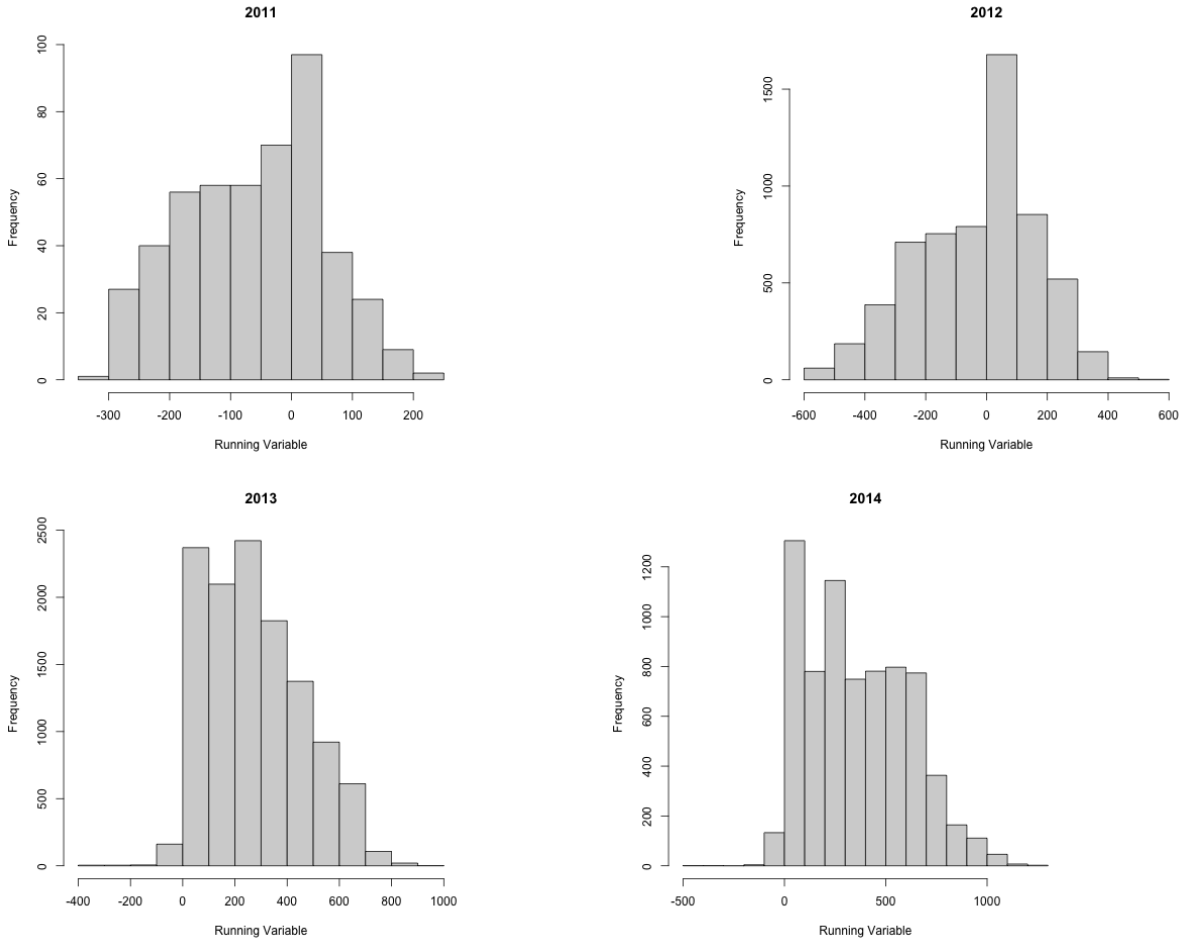
Note: 1. Pearson's Chi-squared test 2. Kruskal-Wallis rank sum test
Source: Elaborated by the author.

The running variable is measured in days and has a negative mean for the sub-samples of 2011 and 2012 and a positive one for the other two years and this provides the first indication that there might be a problem in the last two sub-samples. Figure 5.3 presents the histograms for the running variable in each one of the sub-samples. In analyzing the histogram for all four sub-samples, there is a clear difference between the first two sub-samples, 2011 and 2012, and the other two, 2013 and 2014. In the second sub-group, there are almost no observations to the left of the cut-off (with a negative running variable). This could be an issue for the posterior empirical analysis as we need enough observations to the right and to the left of the cut-off, within a certain bandwidth, to be able to estimate the local linear regressions.

To further analyze the distribution of the running variable, Figure 5.4 presents an analysis of the continuity of the running variable density around the cutoff for all the four sub-samples[64]. The plots for sub-samples of 2013 and 2014 provide more evidence of the imbalance between observations to the right and to the left of the cutoff - corroborating with the analysis of the histograms. The plot for the 2011 sub-sample provides evidence that there is continuity of the density of the running variable around the cutoff, but the plot for the 2012 sub-sample does not. In this last case, even though there is not this imbalance in the number of observations to the right and the left of the cutoff, there is evidence of a discontinuity in the density of the running variable around the cutoff.

[64] The plots were estimated considering no missing math scores. In Figure 5.7, the plots that consider no missing Portuguese language Scores are presented.

Figure 5.3 – Histograms of the Running Variable

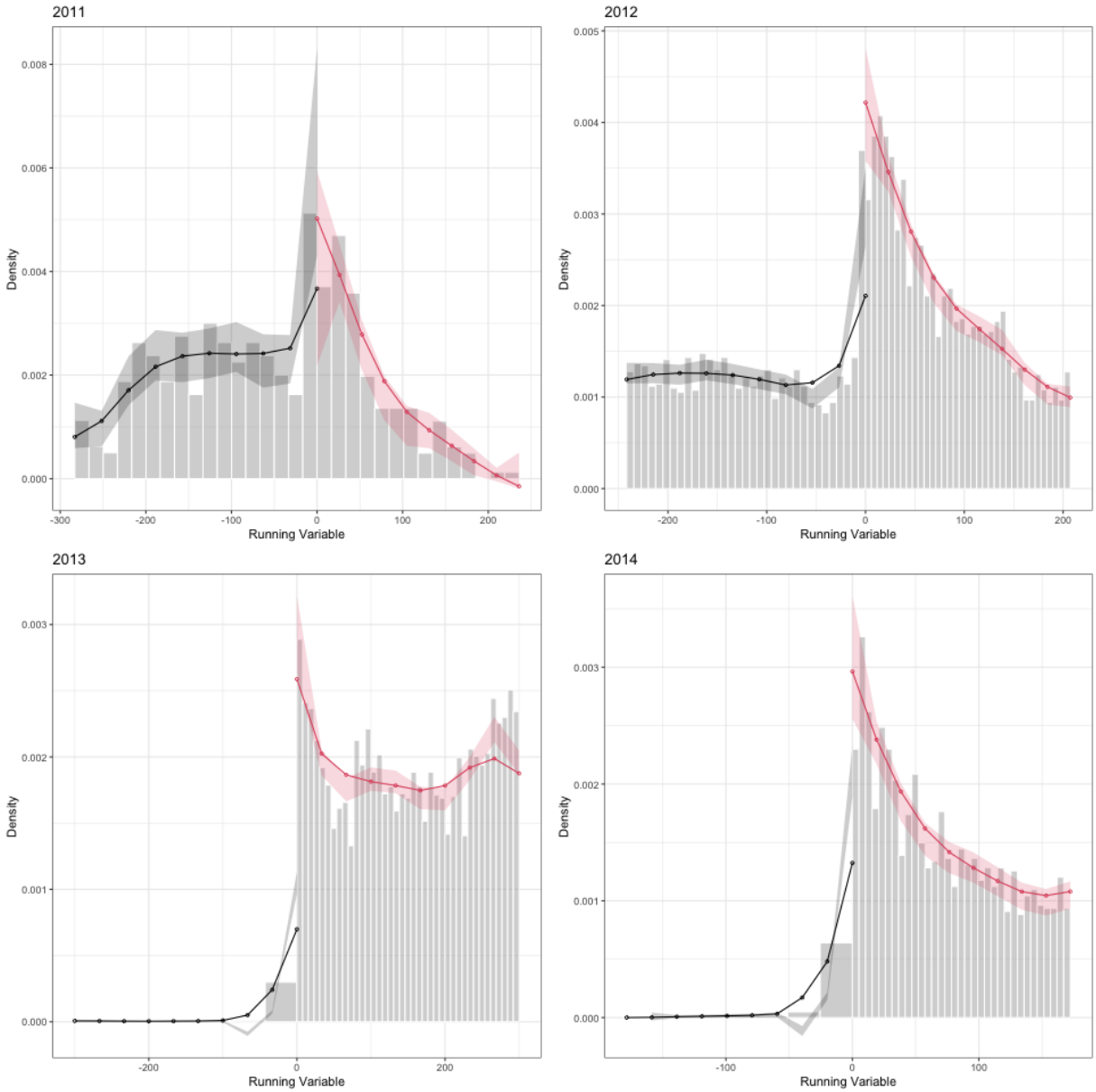


Source: Elaborated by the author.

In addition to the graphical analysis, a statistical test was also implemented. This test uses the the local polynomial density estimators proposed by Cattaneo, Jansson e Ma (2020); its null hypothesis is that there is a continuity of the density functions for control and treatment observations at the cutoff. The results for the test, consistent with the graphical analysis, indicate that only in the 2011 sample, is there evidence of continuity of the density of the running variable and for the other samples, there is no evidence of continuity[65].

[65] For the 2011 sub-sample, the value of the statistic was 0.6015 with a p-value of 0.5475. For the 2012 sub-sample, the value of the statistic was 6.4481 with a p-value of 0. For 2013 sub-sample, 18.3447 with a p-value of 0. And for the 2014 sub-sample, 10.6894 with a p-value of 0. All this considers the results for the sub-samples with non-missing mathematics scores. For the sub-samples with non-missing Portuguese language scores, the only difference is for the 2011 sub-sample.

Figure 5.4 – Estimated Density of the Running Variable



Source: Elaborated by author based on a sample of non-missing math scores.

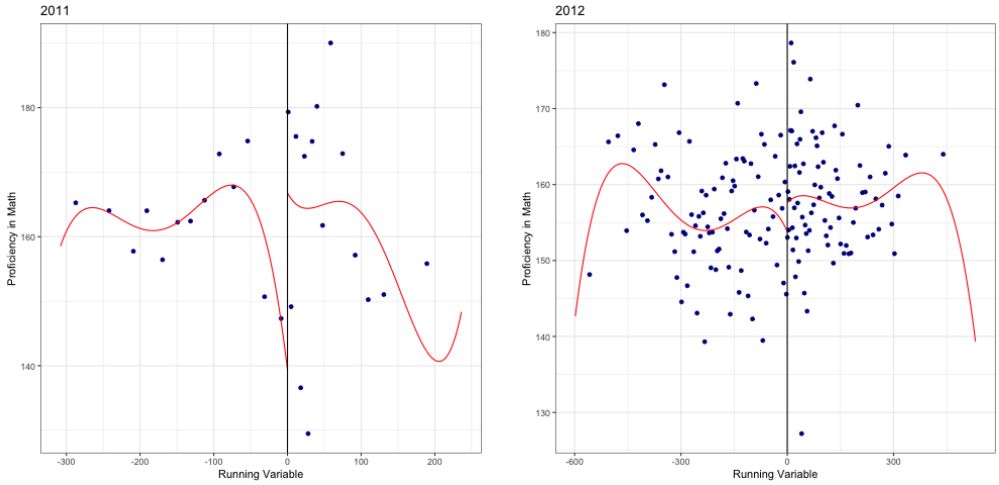
Taken together, these descriptive analyses support our decision to only implement the RD analysis for the 2011 and 2012 sub-samples. For the 2013 and 2014 sub-samples, not only is there a clear imbalance in the number of observations, but also the results of the density test provide evidence of a non-continuity in the density of the running variable around the cutoff that is worrisome. For the 2012 sub-sample, the density test is not favorable, but because there is no clear imbalance in the number of observations, we decided to still analyze its results.

5.6.2 Results of the Regression Discontinuity Design

Figure 5.5 presents plots for the relationship between the proficiency in mathematics and the running variable and Figure 5.6, for the relationship between Portuguese language and the running variable for the 2011 and 2012 sub-samples. These plots were constructed using quantile-spaced bins[66] to divide the number of observations[67]. These plots provide the first evidence of possible effects of having been offered spot in the municipal daycare center system.

The plots for the 2011 sub-sample present a discontinuity in the outcome when the running variable is at the cutoff and this discontinuity is more pronounced when we analyze math scores. For the 2012 sub-sample, there are many more bins, 168 for math and 167 for Portuguese language, a clear consequence of the larger sample size. However, there is no clear discontinuity for Portuguese language and a much less pronounced discontinuity for mathematics.

Figure 5.5 – Mimicking Variance RD Plot with Quantile-Spaced Bins

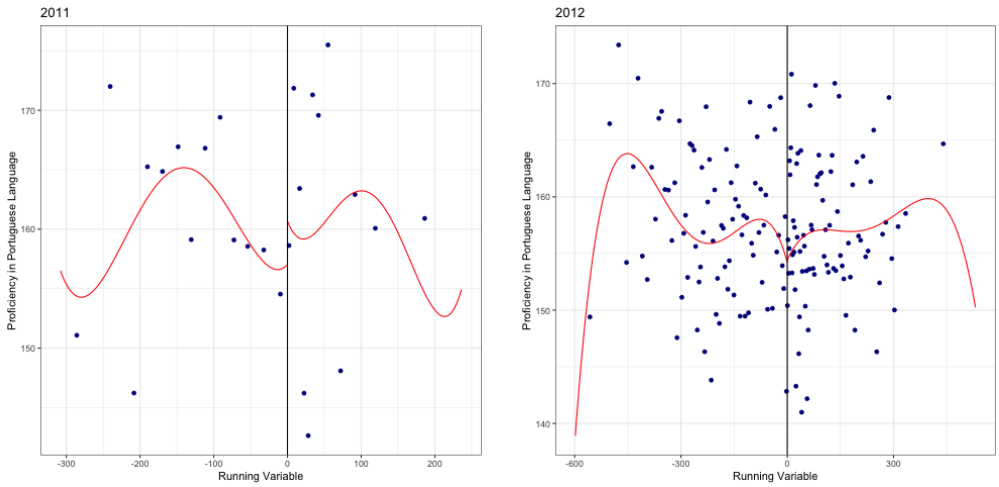


Source: Elaborated by the author by based on a sample of non-missing mathematics scores.

66 The bins were constructed with the same number of observations within in each treatment assignment status.

67 To select the number of bins, we followed Cattaneo, Idrobo e Titiunik (2019) and applied a method that defines the number of bins so that "the overall variability of the binned means “mimics” the overall variability in the raw scatter plot of the data".

Figure 5.6 – Mimicking Variance RD Plot with Quantile-Spaced Bins



Source: Elaborated by the author by based on a sample of non-missing Portuguese language Scores.

The results of the estimation are presented in Table 5.2. These tables present MSE-optimal local-linear point estimator of the RD treatment effect in the first line, followed by its p-value - constructed using a robust bias correction[68], the confidence interval[69], the size of the bandwidth, the number of observations, to the left and to the right of the cut-off point, used in the estimation and the mean of the running variable when only these observations are analyzed. These results are estimated with a correction for clustering at the school level[70].

For the sub-sample of 2011, there is a marginally statistically significant effect for mathematics proficiency, but not for Portuguese language proficiency. The point estimation of the effect in mathematics is 19.1, indicating that for children that were barely offered a spot in the municipal daycare center system, their proficiency in mathematics rises 19 points more than for children who were barely not offered a spot. This represents 12.25% of the mean mathematics proficiency for second-grade students in 2018. Such estimations were based on a sample of 199 observations, out of the 480 which were used, and the bandwidth was 65.7 days in length. The mean of the running variable is much closer to zero than in Table 5.1.

For the 2012 sub-sample, many more observations are used in the estimation - a clear consequence of the number of observations and their distribution. The mean of the running variable is higher: further from zero than that of the 2011 sub-sample. Neither of the estimations are statistically significant and their point estimations are much smaller.

68 For more information, see Cattaneo, Idrobo e Titiunik (2019).

69 Also constructed with a robust bias correction.

70 Because the errors can be correlated within groups, it is appropriate to employ variance estimators that take into account the nature of the structure of the data. The bandwidth selection process depends on variance estimators, and because of that, using cluster-robust variance estimator changes both the estimated standard errors and the point estimates. For the 2011 sub-sample, children were in 258 different schools and for the 2012 sub-sample, 543 different schools.

Table 5.2 – Results - Proficiency in Mathematics

	Mathematics		Portuguese Language	
	2011	2012	2011	2012
Coefficients	19.1	2.9	1.8	-0.7
P-value	0.12	0.47	0.97	0.99
Confidence Interval	[-5.2, 45.7]	[-5.8, 12.6]	[-22.5, 21.5]	[-6.9, 7.0]
Size of Bandwidth	65.7	149.6	79.4	165.5
Number of obs. to the left of the cut-off	69	1056	84	1161
Number of obs. to the right of the cut-off	130	2290	145	2391
Mean of the Running Variable	4.3	16.7	2.5	15.3

Source: Elaborated by the author by based on a sample of non-missing math scores.

5.6.2.1 Validity Tests

The assumptions that guarantee the validity of the RD design depend on non-observable, and consequently non-testable, hypotheses, that go beyond the existence of a rule to assign treatment. To indirectly test these hypotheses, empirical validation methods should be employed. Following Cattaneo, Idrobo e Titiunik (2019), four empirical tests are conducted: (i) the treatment effect on predetermined covariates; (ii) the exclusion of observations closest to the cutoff; (iii) the treatment effect on different bandwidths; and (iv) a change in the cutoff. These tests are implemented only for the 2011 sub-sample and for the mathematics proficiency outcome to check robustness of its results.

For the first test, we implemented two analyses, a graphical and a stastical one, on the following covariates: sex, age and race. For the graphical analysis, four graphs were constructed - that can be found in Figure 5.7. The statistical analysis was conducted using the same estimation procedure implemented in the previous section: a local linear polynomial estimation, correcting for clustering at the school level. Table 5.3 presents the results for all the covariates. The size of the bandwidth used as indicated in the third column, varies significantly between samples and covariates as expected as it depends on the variability of the covariate analyzed. The coefficients are, mostly, very close to zero, and most of the p-values are above any statistically significant threshold. The graphs are consistent with the statistical results as expected and in most of them, the local linear fits are very close to each other at the cut-off.

Table 5.3 – First Validity Test

Covariate	Bandwidth	Coef.	P-value	N_left	N_right
Age	81.56	0.03	0.73	86	141
Sex	73.23	0.12	0.51	77	139
Race	48.92	-0.25	0.11	33	83

Source: Elaborated by the author based on a sample of non-missing math scores.

For the second test, we re-estimate the treatment effect without the observations

closest to the cut-off, following the same strategy as for the main results. This test also allows us to understand how sensitive the results are to these observations which is important as the observations closest to the cutoff are most likely to influence the fitting of the local polynomials due to a triangular kernel that was used to weight the observations. The results of the re-estimation can be found in Table 5.4 where we re-estimated the treatment effect for different amounts of excluded units, varying the size of the interval excluded. The results show that the effect of being offered a spot in 2011 is not robust to excluding observations with a running variable module smaller than 3, 7, 10 or 14.

Table 5.4 – Second Validity Test

Radius of Excluded Obs.	Bandwidth	Coef	P-value	N_left	N_right
3	11.20	0.49	70.77	73	110
7	17.11	0.33	69.72	65	102
10	0.36	0.87	82.98	75	104
14	-22.91	0.41	40.79	25	58

Source: Elaborated by the author based on a sample of non-missing mathematics scores.

The third test consists on an exploration of the sensitivity of the results to the size of the bandwidth. Table 5.5 presents the results of an estimation with an MSE-optimal bandwidth, our default method presented in the previous section, with the results of estimations using a bandwidth that is double that of the MSE-optimal one, using a CER-optimal bandwidth[71] and one using double the CER-optimal bandwidth[72]. The empirical conclusions as presented in the previous section are robust to the change in bandwidths.

Table 5.5 – Third Validity Test

Type of Bandwidth	Coef	P-value	Size of Bandwith	N_left	N_right
MSE-optimal Bandwidth	19.09	0.12	65.73	69	130
Double the MSE-optimal Bandwidth	17.81	0.09	131.45	148	168
CER-optimal Bandwidth	19.94	0.12	49.34	50	115
Double the CER-optimal Bandwidth	17.98	0.09	98.69	108	153

Source: Elaborated by the author based on a sample of non-missing mathematics scores.

The final test checks if the results hold if we change the cutoff from zero to other values. To do so, we re-estimate using the same strategy as before and only changing the cut-off to 4, 7 and 14 days. As presented in Table 5.6, the results do not hold.

[71] This bandwidth is chosen to minimize an approximation to the coverage error of the confidence interval of the point estimate. The MSE-optimal bandwidth leads to more powerful hypothesis tests with larger size distortions than this bandwidth.

[72] We only explore these four bandwidths as bandwidths much larger than the MSE-optimal tend to lead to estimates with too much bias and those much smaller lead to effects with too much variance.

Table 5.6 – Fourth Validity Test

	Coef	P-value	Bandwidth	N_left	N_right
4	28.81	0.39	12.35	15	9
7	23.27	0.38	20.37	42	38
14	-38.52	0.19	26.28	44	58

Note: Elaborated by the author based on a sample of non-missing mathematics scores.

The validity tests provide mixed evidence. The first test provides evidence that children to the left and right of the cutoff are similar when their personal characteristics are analyzed. The second test indicates that the effect is very sensitive to the observations near the cutoff whereas the third test provides some evidence that supports the robustness of the effect of being offered a spot on 2011 on math scores. The fourth test can also be interpreted as providing some evidence of the robustness of the result as there is no statistically significant effect when the cutoff is artificially modified.

5.7 Discussion and Limitations

This study finds evidence of a positive, marginally significant effect on the child's proficiency in mathematics of being offered a spot in *Berçário I*, the first level of the ECEC system[73]. The effect has a substantial size of 12.25% of the mean proficiency in mathematics of second-grade students in São Paulo in 2018. This evidence is somewhat robust to the validity tests implemented.

No evidence of a significant effect is found for the other sub-sample analyzed, composed of children who could potentially enroll in the first or the second level of the ECEC system - *Berçário I* or *Berçário II* in 2012. This could be a consequence of the fact that children who enrolled in different levels were analyzed together but it could also indicate that starting center-based care before the first year of life has an impact on children's cognitive outcomes. This hypothesis is very interesting and further analysis is required to corroborate it. To test this, it is important to replicate this study for other cohorts of children such as those who took the *Provinha São Paulo* in 2019, for example.

Additionally, in line with the rest of this thesis, it would be important to analyze the heterogeneity of impact, and if and how impacts vary with neighborhood vulnerability. This would provide valuable evidence on the potential pathways between neighborhood vulnerability and child development.

Our analysis has several limitations. First, and foremost, it was necessary to construct the sub-samples and the running variable based on several non-testable assumptions

[73] As previously discussed, there are four levels in the daycare center division of the ECEC system; *Berçário I* is the first one. In the sub-sample for 2011 only children who could potentially enroll in this level were analyzed.

as we did not have access to the exact queue the child faced in the enrollment process, and we cannot verify if our choices are correct. The assumptions were based on our understanding of the enrollment process from documents analyzed and exchanges with personnel from the Department of Education. An important assumption is the one that defined the cut-off as the 31st of December of the previous year. Sensitivity tests could be run to better understand if our results are robust to different specifications.

Second, we do not analyze two sub-samples from 2013 and 2014 due to issues with the behavior of the running variable. This might indicate an issue with the running variable construction, that could potentially interfere with the results found for the other two sub-samples. However, this difference could also indicate that the enrollment process for children in older age-groups is simply different from the that of *Berçário I* and *Berçário II*. It would be important to try to replicate our analysis with other cohorts.

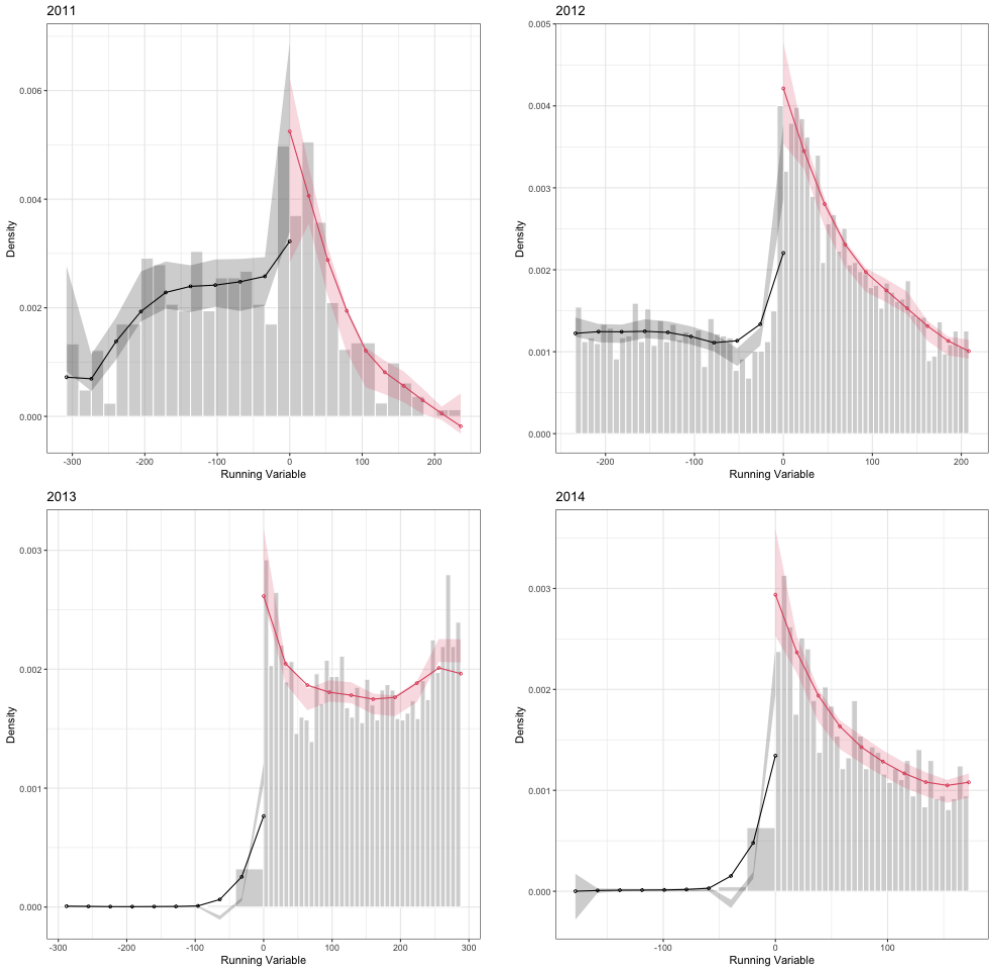
Third, there is a question of external validity, and our results should not be extrapolated for children outside our sample. Replication exercises would again be important to validate our results.

Even with these limitations, we understand that our results provide valuable evidence on the effect for very young children of attending center-based care in São Paulo. This finding indicates the importance of this policy in promoting better educational outcomes. It also calls attention to the need to study the quality of these centers to further understand our results.

Additionally, as we were analyzing publicly funded and free center-based care, the evidence found in this study corroborates the hypothesis that this policy might be an important tool in reducing inequalities. To further explore this, it would be important to understand the social-economic background of the children who benefit from such care.

5.A Appendix A

Figure 5.7 – Estimated Density of the Running Variable



Source: Elaborated by the author based on a sample of non-missing Portuguese language Scores.

6 Conclusion

To guarantee a healthy and integral development for all children, it is essential to allow them to be raised in a nurturing environment. Bronfenbrenner's seminal work on the socio-ecological model, described in Bronfenbrenner e Morris (1998), defends that to study an individual's development, one needs to take into account the different contexts in which this development occurs. This thesis analyzed two of these contexts, the neighborhood and the daycare center, and their relationship with a healthy child development.

In chapter 1, a systematic review of recent empirical literature on neighborhood effects was presented. It corroborates the theoretical hypothesis of Bronfenbrenner's model and concludes that both the socioeconomic conditions of the neighborhood as well as the characteristics of neighborhood social networks are related to child development.

The remainder of this thesis focused on one case, of São Paulo, and empirically tested different associations between child development and the contexts that child experiment during their development. Following the findings in chapter 1, the main neighborhood characteristic analyzed is the socioeconomic status, measured by a social vulnerability index (IVS).

In chapter 3, the existence of neighborhood effects for children's cognitive outcomes is tested. Our results provide valuable evidence that there is a significant association between the social vulnerability of the neighborhood a child was born in and their educational outcomes in middle childhood. These results are in line with the international evidence as systematized chapter 1 and contribute to the literature by providing evidence for the Brazilian context.

This result is especially important in a context of spatial segregation as described in chapter 2. The pattern of spatial segregation found in São Paulo, combined with the heterogeneity of the peripheral areas, underscores the importance of understanding the consequences of space to a person's opportunities and outcomes. In chapter 3, we sought to contribute to this understanding by analyzing one specific outcome, children's educational results in middle childhood.

Besides the neighborhood, the other context of development analyzed in this thesis is that of the daycare center. The provision of free, center-based care for young children from zero to six years old is arguably the main public policy for early childhood development in Brazil. In São Paulo, there was a recent substantive expansion of center-based care. However, as described in chapter 2, there is some evidence that access to center-based care is not equal throughout the city. Additionally, as presented in chapter 4, children's access to center-based care varies according to both children's family background and

neighborhood vulnerability. Taken together, the evidence discussed in these two chapters indicate that the expansion of center-based care in São Paulo had mixed results in including a more diverse set of children.

The importance of attending center-based care is highlighted by the results presented in chapter 5. Using an RD design, in this chapter, we find evidence of a positive, marginally significant effect on the children's proficiency in mathematics of being offered a spot in *Berçário I*, the first level of the ECEC system.

Taken together, those findings allow us to reach a troubling conclusion: center-based care could be an important tool in guarantee that children reach their full developmental potential, but, as children from more vulnerable neighborhoods are less likely to attend it, they are also not the ones benefiting from such policy. Combined with the existence of neighborhood effects, the lack of access to center-based care could have lasting consequences on children's lives.

The findings from this thesis point in two directions. First, place-based policies may be useful in promoting better educational outcomes particularly if they focus on children from more vulnerable neighborhoods. To design such interventions, it is important to advance the understanding of the relationship between neighborhood and children's educational outcomes with a specific focus on the mechanisms that explain this relationship and in which of neighborhood characteristics truly matter. Second, our findings affirm the importance of the availability of center-based care for young children. It also indicate the relevance of expanding access to this type of care in the city's more vulnerable neighborhoods.

The knowledge and insight gained from this thesis may foster learning both in the field of early childhood development and urban studies as well as in the field of public policy. Ultimately, this work can help policy makers develop better early childhood policies that will guarantee more equal opportunities for all. It is also in sync with movements like Urban 95 and child friendly cities that underscore the importance of studying neighborhood context and understanding the challenges of raising a child in an urban environment.

Bibliography

AHNERT, L.; PINQUART, M.; LAMB, M. E. Security of Children's Relationships With Nonparental Care Providers: A Meta-Analysis. *Child Development*, v. 77, n. 3, p. 664–679, maio 2006. ISSN 0009-3920, 1467-8624. Disponível em: <http://doi.wiley.com-/10.1111/j.1467-8624.2006.00896.x>.

AIKENS, N. L.; BARBARIN, O. Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, v. 100, n. 2, p. 235–251, 2008. ISSN 1939-2176, 0022-0663. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-0663.100.2.235>.

ANDERSON, S.; JOHNSTON, W.; LEVENTHAL, T. When neighborhoods matter: Developmental timing and youth reading achievement and problem behaviors. *Social Science Research*, v. 81, p. 1–11, jul. 2019. ISSN 0049089X. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0049089X18306987>.

ANDERSON, S.; LEVENTHAL, T.; DUPéRé, V. Exposure to Neighborhood Affluence and Poverty in Childhood and Adolescence and Academic Achievement and Behavior. *Applied Developmental Science*, v. 18, n. 3, p. 123–138, jul. 2014. ISSN 1088-8691, 1532-480X. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/10888691-.2014.924355>.

BAKER, C. E. Does Parent Involvement and Neighborhood Quality Matter for African American Boys' Kindergarten Mathematics Achievement? *Early Education and Development*, v. 26, n. 3, p. 342–355, abr. 2015. ISSN 1040-9289, 1556-6935. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10409289.2015.968238>.

BARBARIN, O. et al. Children enrolled in public pre-K: The relation of family life, neighborhood quality, and socioeconomic resources to early competence. *American Journal of Orthopsychiatry*, v. 76, n. 2, p. 265–276, 2006. ISSN 0002-9432. Disponível em: <http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login-.aspx?direct=truedb=psyhAN=2006-05892-015site=ehost-livescope=site>.

BARRY, S. J. E. et al. Mapping area variability in social and behavioural difficulties among Glasgow pre-schoolers: linkage of a survey of pre-school staff with routine monitoring data. *Child: Care, Health and Development*, v. 41, n. 6, p. 853–864, 2015. ISSN 1365-2214. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1111/cch-.12237>.

BASSOK, D.; LATHAM, S. Kids Today: The Rise in Children's Academic Skills at Kindergarten Entry. *Educational Researcher*, v. 46, n. 1, p. 7–20, jan. 2017. ISSN 0013189X. Disponível em: <http://search.proquest.com.ezp-prod1.hul.harvard.edu-/docview/1873917019?accountid=11311>.

BAYDAR, N.; AKCINAR, B. Ramifications of socioeconomic differences for three year old children and their families in Turkey. *Early Childhood Research Quarterly*, v. 33, p. 33–48, 2015. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve-/pii/S0885200615000472>.

BENSON, J.; BORMAN, G. D. Family, Neighborhood, and School Settings Across Seasons: When Do Socioeconomic Context and Racial Composition Matter for the Reading Achievement Growth of Young Children? *Teachers College Record*, v. 112, n. 5, p. 1338–1390, 2010. ISSN 0161-4681.

BERLINSKI, S.; SCHADY, N. (Ed.). *The Early Years: Child Well-Being and the Role of Public Policy*. New York: Palgrave Macmillan US, 2015. ISBN 978-1-137-53648-8 978-1-137-53649-5. Disponível em: <http://link.springer.com/10.1057/9781137536495>.

BERNAL, R. et al. The effects of the transition from home-based childcare to childcare centers on children's health and development in Colombia. *Early Childhood Research Quarterly*, v. 47, p. 418–431, 2019. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0885200618300899>.

BICHIR, R. M. Determinantes do acesso à infra-estrutura urbana no município de são Paulo. *Revista Brasileira de Ciências Sociais*, v. 24, n. 70, jun. 2009. ISSN 0102-6909. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextpid=S0102-69092009000200005lng=ptnrm=isotlng=pt>.

BLACK, M. M. et al. Early childhood development coming of age: science through the life course. *The Lancet*, v. 389, n. 10064, p. 77–90, jan. 2017. ISSN 01406736. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0140673616313897>.

BLAIR, L. M.; FORD, J. L. Neighborhood Context and the Risk for Developmental Disabilities in Early Childhood. *Maternal and Child Health Journal*, v. 23, n. 9, p. 1213–1219, set. 2019. ISSN 1092-7875, 1573-6628. Disponível em: <http://link.springer-.com/10.1007/s10995-019-02757-w>.

BLUME, L. et al. Identification of Social Interactions. p. 154.

BRIGGS-GOWAN, M. J.; CARTER, A. S.; FORD, J. D. Parsing the Effects Violence Exposure in Early Childhood: Modeling Developmental Pathways. *Journal of Pediatric Psychology*, v. 37, n. 1, p. 11–22, jan. 2012. ISSN 1465-735X, 0146-8693. Disponível em: <http://academic.oup.com/jpepsy/article/37/1/11/906563/Parsing-the-Effects-Violence-Exposure-in-Early>.

BRITTO, P. R. et al. Nurturing care: promoting early childhood development. *The Lancet*, v. 389, n. 10064, p. 91–102, jan. 2017. ISSN 01406736. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0140673616313903>.

BRITTO, P. R. et al. What implementation evidence matters: scaling-up nurturing interventions that promote early childhood development: What ECD implementation evidence matters. *Annals of the New York Academy of Sciences*, v. 1419, n. 1, p. 5–16, maio 2018. ISSN 00778923. Disponível em: <http://doi.wiley.com/10.1111/nyas.13720>.

BRONFENBRENNER, U.; MORRIS, P. A. The ecology of developmental processes. In: *Handbook of child psychology: Theoretical models of human development, Volume 1, 5th ed.* Hoboken, NJ, US: John Wiley & Sons Inc, 1998. p. 993–1028. ISBN 978-0-471-05527-3.

BROOKS-GUNN, J. Do You Believe In Magic?: What We Can Expect From Early Childhood Intervention Programs. *Social Policy Report*, v. 17, n. 1, p. 1–16, mar.

2003. ISSN 23793988. Disponível em: <http://doi.wiley.com/10.1002/j.2379-3988.2003-.tb00020.x>.

BROWNELL, M. D. et al. A population-based analysis of factors that predict early language and cognitive development. *Early Childhood Research Quarterly*, v. 35, p. 6–18, 2016. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0885200615300090>.

BUGNI, R. P.; JACOB, M. S. ÍNDICE DE VULNERABILIDADE SOCIAL: UMA ANÁLISE DA CIDADE DE SÃO PAULO. In: *Territórios em números : insumos para políticas públicas a partir da análise do IDHM e do IVS de UDHs e regiões metropolitanas brasileiras*. Brasília: IPEA; INCT, 2017. v. 2, p. 336. ISBN 978-85-7811-297-4.

BURCHINAL, M. et al. Neighborhood Characteristics, and Child Care Type and Quality. *Early Education and Development*, v. 19, n. 5, p. 702–725, out. 2008. ISSN 1040-9289, 1556-6935. Disponível em: <https://www.tandfonline.com/doi/full/10.1080-/10409280802375273>.

CALLAHAN, K. L. et al. Neighborhood disadvantage as a moderator of the association between harsh parenting and toddler-aged children's internalizing and externalizing problems. *Journal of Family Psychology*, v. 25, n. 1, p. 68–76, 2011. ISSN 1939-1293, 0893-3200. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0022448>.

CAMILLI, G. et al. Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development. *Teachers College Record*, v. 112, 2010.

CARPIANO, R. M.; LLOYD, J. E.; HERTZMAN, C. Concentrated affluence, concentrated disadvantage, and children's readiness for school: A population-based, multi-level investigation. *Social Science & Medicine*, v. 69, n. 3, p. 420–432, ago. 2009. ISSN 02779536. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S027795360900327X>.

CATTANEO, M. D.; IDROBO, N.; TITIUNIK, R. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. 1. ed. Cambridge University Press, 2019. ISBN 978-1-108-68460-6 978-1-108-71020-6. Disponível em: <https://www.cambridge.org-/core/product/identifier/9781108684606/type/element>.

CATTANEO, M. D.; JANSSON, M.; MA, X. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, v. 115, n. 531, p. 1449–1455, jul. 2020. ISSN 0162-1459, 1537-274X. Disponível em: <https://www.tandfonline.com-/doi/full/10.1080/01621459.2019.1635480>.

CAUGHY, M. O.; O'CAMPO, P. J. Neighborhood Poverty, Social Capital, and the Cognitive Development of African American Preschoolers. *American Journal of Community Psychology*, v. 37, n. 1-2, p. 141, 2006. ISSN 1573-2770. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1007/s10464-005-9001-8>.

Center on the Developing Child at Harvard University. *The Foundations of Lifelong Health Are Built in Early Childhood*. [S.l.]: Harvard University.

CHALITA, G.; CIPRIANO, E. Educação Infantil: histórias, avanços e deSafios. In: *São Paulo Carinhosa: O que grandes cidades e políticas intersetoriais podem fazer pela primeira infância*. São Paulo: Secretaria Municipal de Cultura, 2016.

CHAUDRY, A.; HENLY, J.; MEYERS, M. *Conceptual Frameworks for Child Care Decision-Making*. Washington, DC, 2010.

CHETTY, R.; HENDREN, N. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*. *The Quarterly Journal of Economics*, v. 133, n. 3, p. 1107–1162, ago. 2018. ISSN 0033-5533, 1531-4650. Disponível em: <https://academic.oup.com/qje/article/133/3/1107/4850660>.

CHETTY, R.; HENDREN, N.; KATZ, L. F. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, v. 106, n. 4, p. 855–902, abr. 2016. ISSN 0002-8282. Disponível em: <http://pubs.aeaweb.org/doi/10.1257/aer.20150572>.

CHRISTIAN, H. et al. Relationship between the neighbourhood built environment and early child development. *Health & Place*, v. 48, p. 90–101, nov. 2017. ISSN 13538292. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1353829216304646>.

CHRISTIAN, H. et al. The influence of the neighborhood physical environment on early child health and development: A review and call for research. *Health & Place*, v. 33, p. 25–36, maio 2015. ISSN 1353-8292. Disponível em: <http://ezp-prod1.hul.harvard-.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=psyhAN=2015-18217-006site=ehost-livescope=site>.

COELHO, V. S. P.; MARCONDES, L. M.; BARBOSA, M. Accountability e redução das desigualdades em saúde: A experiência de São Paulo. *Novos Estudos - CEBRAP*, v. 38, n. 2, p. 323–349, ago. 2019. ISSN 01013300. Disponível em: <http://novosestudos.uol.com.br/produto/114/5d698593c436f>.

COELHO, V. S. P.; SZABZON, F.; DIAS, M. F. Política municipal e acesso a serviços de saúde São Paulo 2001-2012, quando as periferias ganharam mais que o centro. *Novos Estudos - CEBRAP*, n. 100, p. 139–161, nov. 2014. ISSN 0101-3300. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextpid=S0101-33002014000300139lng=pttlng=pt>.

COLDER, C. R. et al. Temperament in context: Infant temperament moderates the relationship between perceived neighborhood quality and behavior problems. *Journal of Applied Developmental Psychology*, v. 27, n. 5, p. 456–467, set. 2006. ISSN 01933973. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0193397306000803>.

COLEY, R. L.; LYNCH, A. D.; KULL, M. Early exposure to environmental chaos and children's physical and mental health. *Early Childhood Research Quarterly*, v. 32, p. 94–104, 2015. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve-/pii/S0885200615000344>.

CORTáZAR, A. Long-term effects of public early childhood education on academic achievement in Chile. *Early Childhood Research Quarterly*, v. 32, p. 13–22, 2015. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0885200615000046>.

COSTA, M. A.; MARGUTI, B. O. *Atlas da vulnerabilidade social nas regiões metropolitanas brasileiras*. Brazil: IPEA, 2015. Num Pages: 1. ISBN 978-85-7811-254-7.

COSTA, S. H. D. Acesso à justiça: promessa ou realidade? Uma análise do litígio sobre creche e pré-escola no Município de São Paulo. *Direito processual constitucional*, p. 23, 2016.

COULTON, C. J.; KORBIN, J. E.; SU, M. Neighborhoods and child maltreatment: a multi-level study. *Child Abuse & Neglect*, v. 23, n. 11, p. 1019–1040, nov. 1999. ISSN 0145-2134.

COULTON, C. J. et al. Temporal effects of distressed housing on early childhood risk factors and kindergarten readiness. *Children and Youth Services Review*, v. 68, p. 59–72, set. 2016. ISSN 01907409. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0190740916301943>.

CROWDER, K.; SOUTH, S. J. Spatial and temporal dimensions of neighborhood effects on high school graduation. *Social Science Research*, v. 40, n. 1, p. 87–106, jan. 2011. ISSN 0049089X. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0049089X10000840>.

CURRIE, J. Early Childhood Education Programs. *Journal of Economic Perspectives*, v. 15, n. 2, p. 213–238, maio 2001. ISSN 0895-3309. Disponível em: <https://pubs-.aeaweb.org/doi/10.1257/jep.15.2.213>.

DEARING, E. et al. Do Neighborhood and Home Contexts Help Explain Why Low-Income Children Miss Opportunities to Participate in Activities Outside of School? *Developmental Psychology*, v. 45, n. 6, p. 1545–1562, nov. 2009. ISSN 0012-1649.

DUNCAN, G. J. Can Intensive Early Childhood Intervention Programs Eliminate Income- Based Cognitive and Achievement Gaps? p. 56.

DUNCAN, G. J. et al. School readiness and later achievement. *Developmental Psychology*, v. 43, n. 6, p. 1428–1446, nov. 2007. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0012-1649.43.6.1428>.

DUNCAN, G. J.; MAGNUSON, K. Investing in Preschool Programs. *Journal of Economic Perspectives*, v. 27, n. 2, p. 109–132, fev. 2013. ISSN 0895-3309. Disponível em: <https://pubs.aeaweb.org/doi/10.1257/jep.27.2.109>.

DUNCAN, G. J.; MAGNUSON, K.; VOTRUBA-DRZAL, E. Boosting Family Income to Promote Child Development. *The Future of Children*, v. 24, n. 1, p. 99–120, 2014. ISSN 1550-1558. Disponível em: <http://muse.jhu.edu/content/crossref/journals-/future_of_children/v024/24.1.duncan.html>.

DUPERE, V. et al. Understanding the positive role of neighborhood socioeconomic advantage in achievement: The contribution of the home, child care, and school environments. *Developmental Psychology*, v. 46, n. 5, p. 1227–1244, set. 2010. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037-/a0020211>.

ENGLE, P. L. et al. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, v. 378, n. 9799, p. 1339–1353, out. 2011. ISSN 01406736. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0140673611608891>.

FANTUZZO, J. W. et al. Early childhood experiences and kindergarten success: A population-based study of a large urban setting. *School Psychology Review*, v. 34, n. 4, p. 571–588, 2005. ISSN 0279-6015.

FEDOR, M. C.; BENDER, S. L.; CARLSON, J. S. Examining Risk and Protective Factors in Head Start Populations Located in High- and Low-Violence Communities:. *Infants & Young Children*, v. 23, n. 3, p. 209–217, jul. 2010. ISSN 0896-3746. Disponível em: <https://insights.ovid.com/crossref?an=00001163-201007000-00005>.

FELFE, C.; LALIVE, R. Early Child Care and Child Development: For Whom it Works and Why. *SSRN Electronic Journal*, 2013. ISSN 1556-5068. Disponível em: <http://www.ssrn.com/abstract=2214874>.

FERNANDES, F. S.; DOMINGUES, J. d. R. Educação infantil no estado de São Paulo: condições de atendimento e perfil das crianças. *Educação e Pesquisa*, v. 43, n. 1, p. 145–160, mar. 2017. ISSN 1517-9702. Disponível em: <http://www.scielo.br/scielo-.php?script=sci_arttextpid=S1517-97022017000100145lng=pttlng=pt>.

FLOCH, J.-M. Spatial econometrics - common models. In: *HANDBOOK OF SPATIAL ANALYSIS: Theory and practical application with R*. [S.l.]: INSEE Eurostat, 2018. p. 29.

FLOURI, E.; MIDOUHAS, E.; JOSHI, H. The role of urban neighbourhood green space in children's emotional and behavioural resilience. *Journal of Environmental Psychology*, v. 40, p. 179–186, dez. 2014. ISSN 02724944. Disponível em: <https://linkinghub-.elsevier.com/retrieve/pii/S0272494414000565>.

FRAM, M. S.; KIM, J. Race/ethnicity and the start of child care: A multi-level analysis of factors influencing first child care experiences. *Early Childhood Research Quarterly*, v. 23, n. 4, p. 575–590, out. 2008. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0885200608000343>.

FROILAND, J. M. EXAMINING THE EFFECTS OF LOCATION, NEIGHBORHOOD SOCIAL ORGANIZATION, AND HOME LITERACY ON EARLY COGNITIVE SKILLS IN THE UNITED STATES. p. 15, 2019.

FULLER, B. et al. Welfare reform and child care options for low-income families. *The Future of Children*, v. 12, n. 1, p. 96–119, 2002. ISSN 1054-8289.

GOLDFELD, S. et al. Neighbourhood Effects Influencing Early Childhood Development: Conceptual Model and Trial Measurement Methodologies from the Kids in Communities Study. *Social Indicators Research*, v. 120, n. 1, p. 197–212, jan. 2015. ISSN 0303-8300.

HANSON, M. J. et al. Neighborhood Community Risk Influences on Preschool Children's Development and School Readiness. *Infants & Young Children*, v. 24, n. 1, p. 87–100, 2011. ISSN 0896-3746.

HARDING, D. et al. *Unpacking Neighborhood Influences on Education Outcomes: Setting the Stage for Future Research*. Cambridge, MA, 2010. Disponível em: <http://www.nber.org/papers/w16055.pdf>.

HART, D.; ATKINS, R.; MATSUBA, M. K. The association of neighborhood poverty with personality change in childhood. *Journal of Personality and Social Psychology*, v. 94, n. 6, p. 1048–1061, 2008. ISSN 1939-1315, 0022-3514. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.94.6.1048>.

HAVNES, T.; MOGSTAD, M. Is universal child care leveling the playing field? *Journal of Public Economics*, v. 127, p. 100–114, jul. 2015. ISSN 00472727. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0047272714000899>.

HEBERLE, A. E. et al. The Impact of Neighborhood, Family, and Individual Risk Factors on Toddlers' Disruptive Behavior. *Child Development*, p. n/a–n/a, abr. 2014. ISSN 00093920. Disponível em: <http://doi.wiley.com/10.1111/cdev.12251>.

HECKMAN, J. J. Schools, skills, and synapses. *Economic Inquiry*, v. 46, n. 3, p. 289–324, jul. 2008. ISSN 0095-2583. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10-.1111/j.1465-7295.2008.00163.x>.

HIRSHBERG, D.; HUANG, D. S.-C.; FULLER, B. Which low-income parents select child-care? *Children and Youth Services Review*, v. 27, n. 10, p. 1119–1148, out. 2005. ISSN 01907409. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0190740905000034>.

HURT, H.; BETANCOURT, L. M. Turning 1 Year of Age in a Low Socioeconomic Environment: A Portrait of Disadvantage. *Behavioral Pediatrics*, v. 38, n. 7, p. 8, 2017.

Instituto Brasileiro de Geografia e Estatística, Coordenação de Geografia. *Tipologia intraurbana: espaços de diferenciação socioeconômica nas concentrações urbanas do Brasil*. Rio de Janeiro: IBGE, 2017. ISBN 978-85-240-4429-8.

Instituto Brasileiro de Geografia e Estatística, Coordenação de População; e Indicadores Sociais. *Síntese de indicadores sociais. 2019*. Rio de Janeiro: IBGE, 2019. (Estudos e pesquisas Informação demográfica e socioeconômica, 40). ISBN 978-85-240-4511-0.

JOHN, A. M. S.; TARULLO, A. R. Neighbourhood chaos moderates the association of socioeconomic status and child executive functioning. *Infant and Child Development*, set. 2019. ISSN 1522-7227, 1522-7219. Disponível em: <https://onlinelibrary.wiley.com-/doi/abs/10.1002/icd.2153>.

KERSHAW, P. et al. Toward a Social Care Program of Research: A Population-Level Study of Neighborhood Effects on Child Development. *Early Education & Development*, v. 18, n. 3, p. 535–560, out. 2007. ISSN 1040-9289, 1556-6935. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/10409280701610929>.

KIERNAN, G. et al. the school readiness of children living in a disadvantaged area in Ireland. *Journal of Early Childhood Research*, v. 6, n. 2, p. 119–144, jun. 2008. ISSN 1476-718X, 1741-2927. Disponível em: <http://journals.sagepub.com/doi/10.1177-/1476718X08088674>.

KINGSTON, S. et al. PARENT INVOLVEMENT IN EDUCATION AS A MODERATOR OF FAMILY AND NEIGHBORHOOD SOCIOECONOMIC CONTEXT ON SCHOOL READINESS AMONG YOUNG CHILDREN: Parent Involvement as a Moderator of School Readiness. *Journal of Community Psychology*, v. 41, n. 3, p. 265–276, abr. 2013. ISSN 00904392. Disponível em: <http://doi.wiley.com/10.1002/jcop.21528>.

KOHEN, D.; OLIVER, L.; PIERRE, F. Examining the effects of schools and neighbourhoods on the outcomes of Kindergarten children in Canada. *International Journal of Speech-Language Pathology*, v. 11, n. 5, p. 404–418, jan. 2009. ISSN

1754-9507, 1754-9515. Disponível em: <http://www.tandfonline.com/doi/full/10.1080-/17549500903085919>.

KOHEN, D. E. et al. Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development*, v. 73, n. 6, p. 1844–1860, 2002. ISSN 0009-3920. Disponível em: <http://ezp-prod1.hul.harvard-.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=psyhAN=2002-11165-016site=ehost-livescope=site>.

KOHEN, D. E. et al. Neighborhood Disadvantage: Pathways of Effects for Young Children. *Child Development*, v. 79, n. 1, p. 156–169, jan. 2008. ISSN 0009-3920, 1467-8624. Disponível em: <http://doi.wiley.com/10.1111/j.1467-8624.2007.01117.x>.

KUHNEN, A.; SILVEIRA, S. M. d. Como crianças percebem, idealizam e realizam o lugar onde moram^ipt. *Psicol. USP*, v. 19, n. 3, p. 295–316, set. 2008. ISSN 0103-6564. Disponível em: <pt>.

LECKIE, G. Cross-Classified Multilevel Models Using Stata: How Important Are Schools and Neighborhoods for Students' Educational Attainment? In: *Hierarchical Linear Modeling: Guide and Applications*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc., 2013. p. 311–332. ISBN 978-1-4129-9885-7 978-1-4833-8445-0. Disponível em: <http://methods.sagepub.com-/book/hierarchical-linear-modeling/n14.xml>.

LEVENTHAL, T.; BROOKS-GUNN, J. The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, v. 126, n. 2, p. 309–337, 2000. ISSN 0033-2909. Disponível em: <http://doi.apa.org/getdoi-.cfm?doi=10.1037/0033-2909.126.2.309>.

LEVENTHAL, T.; BROOKS-GUNN, J. Moving to Opportunity: an Experimental Study of Neighborhood Effects on Mental Health. *American Journal of Public Health*, v. 93, n. 9, p. 1576–1582, set. 2003. ISSN 0090-0036, 1541-0048. Disponível em: <http://ajph.aphapublications.org/doi/10.2105/AJPH.93.9.1576>.

LEVENTHAL, T.; FAUTH, R. C.; BROOKS-GUNN, J. Neighborhood poverty and public policy: A 5-year follow-up of children's educational outcomes in the New York City moving to opportunity demonstration. *Developmental Psychology*, v. 41, n. 6, p. 933–952, 2005. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi-.cfm?doi=10.1037/0012-1649.41.6.933>.

LIMA, L. *Essays on human capital investments in Brazil*. Tese (Doutorado) — FGV EAESP, São Paulo, 2019.

LOEB, S. et al. How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, v. 26, n. 1, p. 52–66, fev. 2007. ISSN 02727757. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0272775706000422>.

LOEB, S. et al. Child Care in Poor Communities: Early Learning Effects of Type, Quality, and Stability. *Child Development*, v. 75, n. 1, p. 47–65, jan. 2004. ISSN 0009-3920, 1467-8624. Disponível em: <http://doi.wiley.com/10.1111/j.1467-8624.2004.00653.x>.

LOVASI, G. S. et al. Chlorpyrifos Exposure and Urban Residential Environment Characteristics as Determinants of Early Childhood Neurodevelopment. *American Journal of Public Health*, v. 101, n. 1, p. 63–70, jan. 2011. ISSN 0090-0036, 1541-0048. Disponível em: <http://ajph.aphapublications.org/doi/10.2105/AJPH.2009.168419>.

MA, J. Neighborhood and parenting both matter: The role of neighborhood collective efficacy and maternal spanking in early behavior problems. *Children and Youth Services Review*, v. 70, p. 250–260, nov. 2016. ISSN 01907409. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0190740916303097>.

MA, J.; GROGAN-KAYLOR, A. Longitudinal associations of neighborhood collective efficacy and maternal corporal punishment with behavior problems in early childhood. *Developmental Psychology*, v. 53, n. 6, p. 1027–1041, jun. 2017. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/dev0000308>.

MA, J.; GROGAN-KAYLOR, A.; LEE, S. J. Associations of neighborhood disorganization and maternal spanking with children's aggression: A fixed-effects regression analysis. *Child Abuse & Neglect*, v. 76, p. 106–116, fev. 2018. ISSN 01452134. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0145213417304076>.

MA, J.; KLEIN, S. Does Race/Ethnicity Moderate the Associations between Neighborhood and Parenting Processes on Early Behavior Problems? *Journal of Child and Family Studies*, v. 27, n. 11, p. 3717–3729, nov. 2018. ISSN 1062-1024, 1573-2843. Disponível em: <http://link.springer.com/10.1007/s10826-018-1200-7>.

MADILL, R. et al. *Access to Early Care and Education for Disadvantged Families: Do Levels of Access Reflect States' Child Care Subsidy Policies?* [S.l.], 2018. 33 p.

MAGNUSON, K.; WALDFOGEL, J. Trends in Income-Related Gaps in Enrollment in Early Childhood Education: 1968 to 2013. *AERA Open*, v. 2, n. 2, p. 233285841664893, abr. 2016. ISSN 2332-8584, 2332-8584. Disponível em: <http://journals.sagepub.com-/doi/10.1177/2332858416648933>.

MANSKI, C. F. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, v. 60, n. 3, p. 531, jul. 1993. ISSN 00346527. Disponível em: <https://academic.oup.com/restud/article-lookup/doi/10.2307/2298123>.

MARQUES, E.; BICHIR, R. M. Investimentos públicos, infra-estrutura urbana e produção da periferia em São Paulo. *Espaço & Debates*, v. 27, n. 42, p. 9–30, 2001.

MARQUES, E.; REQUENA, C. O centro voltou a crescer?: trajetórias demográficas diversas e heterogeneidade na São Paulo dos anos 2000. *Novos Estudos - CEBRAP*, n. 95, p. 17–37, mar. 2013. ISSN 0101-3300. Disponível em: <http://www.scielo.br-/scielo.php?script=sci_arttextpid=S0101-33002013000100002lng=pttlng=pt>.

MARQUES, E. C.; BICHIR, R. M. Estado e espaço urbano: revisitando criticamente as explicações sobre as políticas urbanas. *Revista de Sociologia e Política*, n. 16, p. 9–28, jun. 2001. ISSN 0104-4478. Disponível em: <http://www.scielo.br/scielo-.php?script=sci_arttextpid=S0104-44782001000100003lng=pttlng=pt>.

MCCOY, D. C. et al. Neighborhood economic disadvantage and children's cognitive and social-emotional development: Exploring Head Start classroom quality as a mediating

mechanism. *Early Childhood Research Quarterly*, v. 32, p. 150–159, 2015. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0885200615000411>.

MCCOY, D. C.; WALDMAN, M.; FINK, G. Measuring early childhood development at a global scale: Evidence from the Caregiver-Reported Early Development Instruments. *Early Childhood Research Quarterly*, v. 45, p. 58–68, 2018. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0885200618300449>.

MCWAYNE, C. M. et al. An Emic, Mixed-Methods Approach to Defining and Measuring Positive Parenting Among Low-Income Black Families. *Early Education and Development*, v. 28, n. 2, p. 182–206, mar. 2017. ISSN 1040-9289.

MCWAYNE, C. M. et al. Employing community data to investigate social and structural dimensions of urban neighborhoods: An early childhood education example. *American Journal of Community Psychology*, v. 39, n. 1-2, p. 47–60, 2007. ISSN 1573-2770. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1007/s10464-007-9098-z>.

MILBRATH, C.; GUHN, M. Neighbourhood culture and immigrant children's developmental outcomes at kindergarten. *Early Childhood Research Quarterly*, v. 48, p. 198–214, 2019. ISSN 08852006. Disponível em: <http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=truedb=aphAN=137073453site=ehost-livescope=site>.

MINH, A. et al. A review of neighborhood effects and early child development: How, where, and for whom, do neighborhoods matter? *Health & Place*, v. 46, p. 155–174, jul. 2017. ISSN 13538292. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1353829216303525>.

MORRISSEY, T. W.; VINOPAL, K. Center-based early care and education and children's school readiness: Do impacts vary by neighborhood poverty? *Developmental Psychology*, v. 54, n. 4, p. 757–771, abr. 2018. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/dev0000470>.

MORRISSEY, T. W.; VINOPAL, K. M. Neighborhood Poverty and Children's Academic Skills and Behavior in Early Elementary School. *Journal of Marriage and Family*, v. 80, n. 1, p. 182–197, 2018. ISSN 1741-3737. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1111/jomf.12430>.

MORRISSEY, T. W.; VINOPAL, K. M. Neighborhood Poverty and Children's Academic Skills and Behavior in Early Elementary School. *Journal of Marriage and Family*, v. 80, n. 1, p. 182–197, 2018. ISSN 0022-2445.

NERY, M. B.; SOUZA, A. A. L. d.; ADORNO, S. Os padrões urbano-demográficos da capital paulista. *Estudos Avançados*, v. 33, n. 97, p. 5–36, dez. 2019. ISSN 1806-9592, 0103-4014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextpid=S0103-40142019000400005tlng=pt>.

NEVES, K. da R. et al. Growth and development and their environmental and biological determinants. *Jornal de Pediatria*, v. 92, n. 3, p. 241–250, maio 2016. ISSN 00217557. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0021755715001941>.

NICHD Early Child Care Research Network. Early Child Care and Children's Development Prior to School Entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, v. 39, n. 1, p. 133–164, mar. 2002. ISSN 0002-8312, 1935-1011. Disponível em: <http://journals.sagepub.com/doi/10.3102-/00028312039001133>.

NICHD Early Child Care Research Network. Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist*, v. 61, n. 2, p. 99–116, 2006. ISSN 1935-990X, 0003-066X. Disponível em: <http://doi.apa.org/getdoi-.cfm?doi=10.1037/0003-066X.61.2.99>.

NOBOA-HIDALGO, G. E.; URZúA, S. S. The Effects of Participation in Public Child Care Centers: Evidence from Chile. *Journal of Human Capital*, v. 6, n. 1, p. 1–34, mar. 2012. ISSN 1932-8575, 1932-8664. Disponível em: <https://www.journals.uchicago.edu-/doi/10.1086/664790>.

NORES, M.; BARNETT, W. S. *Access to High Quality Early Care and Education: Readiness and Opportunity Gaps in America*. New Brunswick, NJ, 2014. 30 p.

NORES, M.; BERNAL, R.; BARNETT, W. S. Center-based care for infants and toddlers: The aeioTU randomized trial. *Economics of Education Review*, v. 72, p. 30–43, out. 2019. ISSN 02727757. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0272775718305855>.

ODGERS, C. L. et al. Supportive parenting mediates neighborhood socioeconomic disparities in children's antisocial behavior from ages 5 to 12. *Development and Psychopathology*, v. 24, n. 3, p. 705–721, ago. 2012. ISSN 0954-5794, 1469-2198. Disponível em: <https://www.cambridge.org/core/product/identifier/S0954579412000326/type-/journal_article>.

OLIVEIRA, V. E. d.; SILVA, M. P. d.; MARCHETTI, V. JUDICIÁRIO E POLÍTICAS PÚBLICAS: O CASO DAS VAGAS EM CRECHES NA CIDADE DE SÃO PAULO. *Educação & Sociedade*, v. 39, n. 144, p. 652–670, mar. 2018. ISSN 1678-4626, 0101-7330. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextpid=S0101-73302018000300652lng=pttlng=pt>.

OLIVEIRA, Z. d. M. R. A creche no Brasil: mapeamento de uma trajetória. *R. Fac. Educ.*, v. 14, n. 1, p. 43–52, jun. 1988.

OLIVEIRA, Z. d. M. R.; FERREIRA, M. C. R. Propostas para o Atendimento em Creches no Município de São Paulo. *Cad. Pesq.*, v. 56, p. 39–65, 1986.

PALAMAR, J. J. et al. Family- and Neighborhood-Level Factors as Predictors of Conduct Problems in School among Young, Urban, Minority Children. *Behavioral Medicine*, v. 41, n. 4, p. 177–185, out. 2015. ISSN 0896-4289, 1940-4026. Disponível em: <http://www.tandfonline.com/doi/full/10.1080/08964289.2014.907770>.

PEARMAN, F. A. The Effect of Neighborhood Poverty on Math Achievement: Evidence From a Value-Added Design. *Education and Urban Society*, v. 51, n. 2, p. 289–307, fev. 2019. ISSN 0013-1245, 1552-3535. Disponível em: <http://journals.sagepub.com/doi/10-.1177/0013124517715066>.

PEI, F. et al. The influences of neighborhood disorder on early childhood externalizing problems: The roles of parental stress and child physical maltreatment. *Journal of Community Psychology*, v. 47, n. 5, p. 1105–1117, 2019. ISSN 1520-6629. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1002/jcop.22174>.

ROOS, L. L.; WALL-WIELER, E.; LEE, J. B. Poverty and Early Childhood Outcomes. *Pediatrics*, v. 143, n. 6, p. e20183426, jun. 2019. ISSN 0031-4005, 1098-4275. Disponível em: <http://pediatrics.aappublications.org/lookup/doi/10.1542/peds.2018-3426>.

RUZEK, E. et al. The quality of toddler child care and cognitive skills at 24 months: Propensity score analysis results from the ECLS-B. *Early Childhood Research Quarterly*, v. 29, n. 1, p. 12–21, jan. 2014. ISSN 08852006. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0885200613000732>.

SAMPSON, R. J.; MORENOFF, J. D.; GANNON-ROWLEY, T. Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology*, v. 28, n. 1, p. 443–478, ago. 2002. ISSN 0360-0572, 1545-2115. Disponível em: <http://www.annualreviews.org/doi/10.1146/annurev.soc.28.110601-.141114>.

SANBONMATSU, L. et al. Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment. *Journal of Human Resources*, XLI, n. 4, p. 649–691, 2006. ISSN 0022-166X, 1548-8004. Disponível em: <http://jhr.uwpress.org-/lookup/doi/10.3368/jhr.XLI.4.649>.

SHAW, D. S. et al. Transactional effects among maternal depression, neighborhood deprivation, and child conduct problems from early childhood through adolescence: A tale of two low-income samples. *Development and Psychopathology*, v. 28, n. 3, p. 819–836, ago. 2016. ISSN 0954-5794, 1469-2198. Disponível em: <https://www-.cambridge.org/core/product/identifier/S095457941600033X/type/journal_article>.

SHIN, E. K. et al. Association of Maternal Social Relationships With Cognitive Development in Early Childhood. *JAMA Network Open*, v. 2, n. 1, p. e186963, jan. 2019. ISSN 2574-3805. Disponível em: <http://jamanetworkopen.jamanetwork.com/article-.aspx?doi=10.1001/jamanetworkopen.2018.6963>.

SHONKOFF, J. P.; PHILLIPS, D. A.; (U.S.), N. R. C. (Ed.). *From neurons to neighborhoods: the science of early child development*. Washington, D.C: National Academy Press, 2000. ISBN 978-0-309-06988-5.

SIBLEY, E. et al. Do increased availability and reduced cost of early childhood care and education narrow social inequality gaps in utilization? Evidence from Norway. *International Journal of Child Care and Education Policy*, v. 9, n. 1, p. 1, dez. 2015. ISSN 2288-6729. Disponível em: <https://ijccep.springeropen.com/articles/10.1007-/s40723-014-0004-5>.

SMALL, M. L.; JACOBS, E. M.; MASSENGILL, R. P. Why Organizational Ties Matter for Neighborhood Effects: Resource Access through Childcare Centers. *Social Forces*, v. 87, n. 1, p. 387–414, set. 2008. ISSN 0037-7732, 1534-7605. Disponível em: <https://academic.oup.com/sf/article-lookup/doi/10.1353/sof.0.0079>.

SMALL, M. L.; NEWMAN, K. Urban Poverty after The Truly Disadvantaged: The Rediscovery of the Family, the Neighborhood, and Culture. *Annual Review of Sociology*, v. 2, p. 23–45, 2001. Disponível em: <http://www.jstor.org/stable/2678613>.

SUBRAMANIAN, S. The relevance of multilevel statistical methods for identifying causal neighborhood effects. *Social Science & Medicine*, v. 58, n. 10, p. 1961–1967, maio 2004. ISSN 02779536. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0277953603004155>.

TORRES, H. d. G. Políticas Sociais e Território. In: *São Paulo: segregação, pobreza e desigualdades sociais*. São Paulo: SENAC, 2004. p. 297–314.

TORRES, H. d. G.; BICHIR, R. Conseqüências da segregação residencial para as Políticas Públicas: o caso do atendimento básico em saúde em São Paulo. *Caderno CRH*, v. 20, n. 50, p. 245–259, ago. 2007. ISSN 0103-4979. Disponível em: <http://www.scielo.br-/scielo.php?script=sci_abstractpid=S0103-49792007000200005lng=ennrm=isotlng=pt>.

TORRES, H. d. G. et al. Pobreza e espaço: padrões de segregação em São Paulo. *Estudos Avançados*, v. 17, n. 47, p. 97–128, abr. 2003. ISSN 0103-4014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextpid=S0103-40142003000100006lng=pttlng=pt>.

VADEN-KIERNAN, M. et al. Neighborhoods as a Developmental Context: A Multilevel Analysis of Neighborhood Effects on Head Start Families and Children. *American Journal of Community Psychology*, v. 45, n. 1-2, p. 49–67, 2010. ISSN 1573-2770. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1007/s10464-009-9279-z>.

VILLANUEVA, K. et al. Local Housing Characteristics Associated with Early Childhood Development Outcomes in Australian Disadvantaged Communities. *International Journal of Environmental Research and Public Health*, v. 16, n. 10, maio 2019. ISSN 1661-7827. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6572259/>.

VILLAçA, F. São Paulo: segregação urbana e desigualdade. *Estudos Avançados*, v. 25, n. 71, p. 37–58, abr. 2011. ISSN 0103-4014. Disponível em: <http://www.scielo.br/scielo-.php?script=sci_arttextpid=S0103-40142011000100004lng=pttlng=pt>.

VILSAINT, C. L. et al. The Ecology of Early Childhood Risk: A Canonical Correlation Analysis of Children's Adjustment, Family, and Community Context in a High-Risk Sample. *The Journal of Primary Prevention*, v. 34, n. 4, p. 261–277, ago. 2013. ISSN 0278-095X, 1573-6547. Disponível em: <http://link.springer.com/10.1007/s10935-013-0305-4>.

VINOPAL, K.; MORRISSEY, T. W. Neighborhood disadvantage and children's cognitive skill trajectories. *Children and Youth Services Review*, v. 116, p. 105231, set. 2020. ISSN 01907409. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii-/S0190740920300669>.

WEILAND, C.; YOSHIKAWA, H. Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. *Child Development*, v. 84, n. 6, p. 2112–2130, nov. 2013. ISSN 00093920. Disponível em: <http://doi.wiley.com/10.1111/cdev.12099>.

WINSLOW, E. B.; SHAW, D. S. Impact of neighborhood disadvantage on overt behavior problems during early childhood. *Aggressive Behavior*, v. 33, n. 3, p. 207–219, 2007. ISSN 1098-2337. Disponível em: <http://onlinelibrary.wiley.com/doi/abs/10.1002/ab.20178>.

WODTKE, G. T. Duration and Timing of Exposure to Neighborhood Poverty and the Risk of Adolescent Parenthood. *Demography*, v. 50, n. 5, p. 1765–1788, out. 2013. ISSN 0070-3370, 1533-7790. Disponível em: <http://link.springer.com/10.1007/s13524-013-0219-z>.

WODTKE, G. T.; HARDING, D. J.; ELWERT, F. Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation. *American Sociological Review*, v. 76, n. 5, p. 713–736, out. 2011. ISSN 0003-1224, 1939-8271. Disponível em: <http://journals.sagepub.com/doi/10.1177-/0003122411420816>.

WOLF, S.; MAGNUSON, K. A.; KIMBRO, R. T. Family poverty and neighborhood poverty: Links with children's school readiness before and after the Great Recession. *Children and Youth Services Review*, v. 79, p. 368–384, ago. 2017. ISSN 01907409. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0190740917301731>.

YOSHIKAW, H.; WEILAND, C.; BROOKS-GUNN, J. When Does Preschool Matter? *The Future of Children*, v. 26, n. 2, p. 21–35, 2016. ISSN 1550-1558. Disponível em: <https://muse.jhu.edu/article/641241>.

ZANGGER, C. Making a place for space: Using spatial econometrics to model neighborhood effects. *Journal of Urban Affairs*, v. 41, n. 8, p. 1055–1080, nov. 2019. ISSN 0735-2166, 1467-9906. Disponível em: <https://www.tandfonline.com/doi/full/10-.1080/07352166.2019.1584530>.