Bilingualism: Visualizing the Outcomes for Latinx Students in the United States

Paul Polanco
Southern Methodist University, paul.polanco@gmail.com

Follow this and additional works at: https://scholar.smu.edu/simmons_dtl_etds

Part of the Bilingual, Multilingual, and Multicultural Education Commons, and the Labor Economics Commons

Recommended Citation
Polanco, Paul, "Bilingualism: Visualizing the Outcomes for Latinx Students in the United States" (2019). Department of Teaching and Learning Theses and Dissertations. 2.
https://scholar.smu.edu/simmons_dtl_etds/2

This Thesis is brought to you for free and open access by the Department of Teaching and Learning at SMU Scholar. It has been accepted for inclusion in Department of Teaching and Learning Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
Bilingualism: Visualizing the Outcomes for Latinx Students in the United States

Paul Polanco

Southern Methodist University

Advisor: Doris Luft de Baker

November 2019
Abstract

In this dissertation, I analyzed the outcomes of bilingualism for the growing Latinx community living in the United States. Outcomes were quantitatively analyzed from four different perspectives: educational outcomes, job market participation, income, and social capital engagement. Chapters 1 and 2 cover previous studies about bilingualism, the importance of including outcomes that are not purely related to income, and general characteristics of the Latinx community. To perform the analyses, I used the Educational Longitudinal Study of 2002/2012 (ELS:2002) dataset, a nationally representative dataset administered by the National Center for Educational Statistics (NCES) of the Institute of Education Sciences, U.S. Department of Education. To determine the population sample, I used coarsened exact matching between the selected sample and the control group. This statistical technique allows for the matching of individuals a priori on an array of characteristics that make analyses stronger and increase internal validity. After the matching, I conducted a series of ordinary least squares regressions, including fixed effects of location using zip codes to account for the location of the individuals. In general, results were significantly positive for bilingual Latinx compared to non-bilingual Latinx in the United States. Thus, it appears that bilingualism is an advantage for Latinx individuals in variables of education, job market participation, and income. However, results were not as precise for social capital engagement activities. Finally, I present a discussion of the results and direction of future studies.
Chapter 1: Introduction

Since the early days of the American colonies, promoting bilingualism through bilingual education has been a perennial issue of controversy, capturing the attention of politicians, policymakers, and parents (Crawford, 2004). Sometimes there have been waves of support of bilingual education as in Florida in the 1960s. Other times there has been strong opposition to bilingual education in public schools as in the late 1990s in places like California, Arizona, and Massachusetts. The support for bilingual programs by policymakers and the community appears to be a function of the political power and level of influence their stakeholders have. For example, in the 1960s, after the Cuban Revolution in 1959, Cuban immigrants who fled their country created a successful bilingual program focused on preserving both languages in Dade County, Florida. The success of this program has been attributed, in part, to the active support of professional Cuban parents and federal assistance through the Cuban Refugee Act (Baker, Basaraba, & Polanco, 2016).

Also, during this time, various policies led to the support of bilingual education. For example, the Bilingual Education Act of 1968 led to helping school districts with establishing educational programs for immigrant children who did not speak English (Baker, Basaraba, & Polanco, 2016). Bilingual education was the preferred method of literacy instruction for English learners (ELs) in the 1970s and 1980s (Slavin & Cheung, 2005). However, the 1990s experienced a period of retrenchment on bilingual education, culminating in English only state policies such as Proposition 227 in California 1998, Proposition 203 in Arizona in the year 2000, and Question 2 in Massachusetts in the year 2002. The purpose of these educational initiatives was to focus on English-only literacy and language development. With few exceptions, the general goal of bilingual programs in the U.S. has been to use the native language of students to
transition them into English. Thus the goal has been for students to become proficient in English, not for students to achieve proficiency in both languages (Polanco & Baker, 2018). Such policies are premised mainly on the monolingual argument that students should communicate in English to fully participate and enjoy the benefits of the American economy (Lyons, 1990; August & Shanahan, 2007). Furthermore, some researchers have argued that bilingualism may carry a penalty for individuals in the labor market (Chiswick & Miller, 2007; Shin & Alba, 2009).

Critics have increasingly argued against an assimilationist perspective of teaching language, contending that by ignoring the native language of students, the educational system creates a de facto disconnection between the personal and cultural identities of the students (Pease-Alvarez & Hakuta, 1992; Baker, Basaraba, Polanco, 2016). Thus, society fails to capitalize on the cultural and academic values that students already possess that can support academic knowledge and subsequently have an impact on the educational and economic outcomes of the students (Coyne, Kame’enui & Carnine, 2011; García & Kleifgen, 2018).

Studies that analyzed the academic results for bilingual students in states that supported English-only policies, such as California and Massachusetts, demonstrated that bilingual students did not perform better in English when they attended English-only programs compared to students who attended bilingual programs (Gándara, 2018). Moreover, an emerging body of research has demonstrated that bilingual education yields better academic performance and higher cognitive and language abilities than monolingual education. This effect is particularly strong for students with low socioeconomic status, where the flexibility and processing of bilingual minds can potentially offset the effects of poverty (Calvo & Bialystock, 2013; Prior & Gollan, 2011). Another criticism of a monolingual perspective is based on studies that demonstrate that there are
important cost benefits associated with bilingualism, where in general, bilingual individuals earned more than monolinguals (Gándara, 2018; Robinson-Cimpián, 2014; Agirdag, 2014a).

Given the ongoing policy relevance of the merits of bilingual education as a source of dual language development and not as a tool to develop only English skills, and the debate that surrounds whether bilingualism leads to better educational, economic, and social outcomes, there is a critical need for rigorous empirical research on the impact of bilingualism and bilingual education programs and their impact on students’ success in the educational system and in the labor market (Gándara, 2018; Polanco & Baker, 2016)

The purpose of this dissertation is to evaluate the educational, economic, and social outcomes for Latinx bilinguals in the United States after high school compared to Latinx non-bilinguals. Latinx are the largest immigrant group in the United States, and this ethnic group is projected to grow from 59 million in 2018 to 107 million by 2065 (Flores, 2017; Flores, Lopez & Krogstad, 2019). Moreover, Latinx individuals are often bilingual and bicultural, and this duality may have an impact on the outcomes. Therefore, it is essential to understand the impact of bilingualism in this specific population. I will explore the outcomes for bilingual and non-bilingual Latinx in three general areas. First, I look at the educational attainment of Latinx and the likelihood of them achieving a post high-school degree. Second, I look at their labor market outcomes, as defined by participation in the job market and salaries earned in the post-secondary years. Third, I look at the possibility of bilingualism being a tool to engage in social activities in the community. I hypothesize that language is a factor that can lead to higher educational attainment, higher job market participation, and greater involvement in the community.
Defining Bilingualism

Before selecting any individuals for the analysis, the first thing to do is to define who is considered bilingual. Grosjean (2010) proposes that “bilinguals are those who use two or more languages in their everyday lives.” However, this definition is too broad, and it does not differentiate between bilinguals with advanced language proficiency in both languages, and bilinguals with beginning or intermediate language proficiency in either language or in both. Defining bilingualism is a process that can be complex because one must consider when, where, and with whom an individual uses the languages. Moreover, assessments in the native language, such as language proficiency, comprehension skills, or language abilities, better capture various levels of domain in one language (Baker, 2017). However, this information is not often available in longitudinal datasets.

Another consideration is the context in which both languages are used because it helps determine the level of language proficiency. For example, an individual speaking the native language only at home, might not have a more advanced level of language proficiency in the native language compared to an individual who uses the native language in school and at work. Similarly, an individual who speaks a second language with family members at home would use a more colloquial language compared to using the second language in school or in a professional work environment. Moreover, there is a choice component where bilingual individuals use contextual clues to determine which language is more appropriate to use (Baker, 2017; Cummins, 2000).

For consistency, in this dissertation, I define bilingualism as an individual who speaks Spanish at home and whose English proficiency is fluent or above as determined by an objective measure of language proficiency assessment in the individuals’ student records. Also, I refer to
bilingual education as the formal instruction that uses the native language of the student for instructional purposes.

**Latinx in the United States**

Traditionally, in the United States, Hispanic is the common terminology used to define “Americans who identify themselves as being of Spanish-speaking background and who trace their origin or descent from Mexico, Puerto Rico, Cuba, Dominican Republic, Central and South America, and other Spanish-speaking countries” (Passel & Taylor, 2009). However, many people of Latin American descent in some areas of the country identify themselves as Latino or Latinx. These terms are considered more inclusive, as they involve everyone with a Latin American background, independently of their native language, which could also be Portuguese, French creole, and others (García & Kleifgen, 2018). In the literature, the terms Hispanic and Latino, Latina, or Latinx are often used interchangeably. In this dissertation, I will study the effects of Spanish-English bilingualism, so I will use the term Latinx to refer to individuals in the United States who speak Spanish and who were either born in Latin America, or who have a Latin American heritage (i.e., they were born in the U.S. of Latin American parents).

The Latinx population in the United States represents a critical demographic segment in terms of growth and participation in the economy. A recent study revealed that this population segment in 2018 was near 60 million, and it has been the principal driver of population growth in the country, accounting for 54% of the nation’s population growth between the years 2000 and 2014 Latinx. According to some estimates, the Latinx segment of the population is expected to represent about 30% of the nation’s population by 2060 (Moore, Fee, Ee, Wiley, & Arias, 2014; Flores, 2017; Flores, Lopez & Krogstad, 2019; Stepler & Lopez, 2016).
Economically, Latinx-owned businesses reported an estimated sales/receipt of $360 billion, and 67% of the Latinx ages 16 and older were employed in the civilian labor force. Moreover, Latinx individuals represent the youngest and largest racial and ethnic group, and it could become an essential sector of the U.S. labor market in the future (see Table 1; Flores, 2017; U.S. Census, 2014). The Latinx population, however, is not only composed of immigrant Latinx, but it also includes U.S.-born individuals of Latinx heritage. In fact, out of the 56 million Latinx in the U.S. in 2016, 66% are U.S.-born (Flores, 2017).

Table 1

Demographic Statistics of Latinx in the U.S.

<table>
<thead>
<tr>
<th>Population*</th>
<th>58 mill.</th>
<th>Latinx Business Sales †</th>
<th>$360 bill.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth/year</td>
<td>4.4%</td>
<td>Civilian Labor Force †</td>
<td>67%</td>
</tr>
<tr>
<td>Nation’s Growth Share</td>
<td>54%</td>
<td>Management, business, science &amp; arts †</td>
<td>21%</td>
</tr>
<tr>
<td>Population Share by 2060</td>
<td>30%</td>
<td>U.S.-born Latinx*</td>
<td>66%</td>
</tr>
</tbody>
</table>

* Adapted from Flores, 2017
† Census Data 2014

Regarding education, about 40% of Latinx ages 25 and older had some college experience in 2015, which represents a 10% increase from 2000. As for language, about 66% (37 million) of Latinx ages five and older speak Spanish at home, and 63% indicate that they are proficient in English (see Table 2; Flores, 2017; Krogstad & Lopez, 2017; U.S. Census, 2016).
Table 2

*Educational and Language Statistics of Latinx in the U.S.*

<table>
<thead>
<tr>
<th>Educational or Language Statistics</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some college</td>
<td>40%</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>15%</td>
</tr>
<tr>
<td>Spanish at home</td>
<td>66%</td>
</tr>
<tr>
<td>Proficient in English</td>
<td>63%</td>
</tr>
<tr>
<td>Speak only English at home</td>
<td>25%</td>
</tr>
</tbody>
</table>

Adapted from U.S. Census, 2016; Flores, 2017; Krogstad & Lopez, 2017

Finally, in terms of location, recent data indicates that 65% of Latinx are concentrated in five states, California, Texas, Florida, New York, and Illinois (see Figure 1). However, economic opportunities have promoted dispersion in recent years with significant growth in areas that have traditionally seen little Latinx presence such as North Carolina and Georgia (Stepler & Lopez, 2016).

*U.S. Latinx Concentrations*

*Figure 1. U.S. Latinx Concentrations (Adapted from Stepler & Lopez, 2016)*
Diversity within the Latinx population. When defining Latinx, it is important to highlight that this racial and ethnic group is not monolithic and researchers agree that a focus on individual achievement of a general group, conceals fundamental differences between the nationalities that this category represents (Moor et al., 2014; Lukes, 2015; Rumbaut & Komaie, 2010; U.S. Census, 2014, Flores, 2017). Mexico has historically been the country with the largest representation of immigrants in the United States. However, in recent years, there has been a rise in the number of immigrants from El Salvador, Guatemala, and Honduras (Cohn, Passel, Gonzalez-Barrera, 2017). Table 3 summarizes the major characteristics of the seven largest Latinx groups represented in the United States.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Population (in millions)</th>
<th>High School Attainment</th>
<th>College or Higher</th>
<th>English Proficiency</th>
<th>Income (in thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Latinx</td>
<td>56.477</td>
<td>27%</td>
<td>15%</td>
<td>69%</td>
<td>$44.8</td>
</tr>
<tr>
<td>Mexicans</td>
<td>35.758</td>
<td>28%</td>
<td>11%</td>
<td>69%</td>
<td>$44.2</td>
</tr>
<tr>
<td>Puerto Ricans</td>
<td>5.371</td>
<td>29%</td>
<td>19%</td>
<td>83%</td>
<td>$40.5</td>
</tr>
<tr>
<td>Salvadorans</td>
<td>2.174</td>
<td>25%</td>
<td>10%</td>
<td>51%</td>
<td>$47.6</td>
</tr>
<tr>
<td>Cubans</td>
<td>2.116</td>
<td>29%</td>
<td>27%</td>
<td>61%</td>
<td>$44.5</td>
</tr>
<tr>
<td>Dominicans</td>
<td>1.866</td>
<td>26%</td>
<td>17%</td>
<td>58%</td>
<td>$36.8</td>
</tr>
<tr>
<td>Guatemalans</td>
<td>1.384</td>
<td>23%</td>
<td>9%</td>
<td>46%</td>
<td>$40.2</td>
</tr>
<tr>
<td>Colombians</td>
<td>1.091</td>
<td>25%</td>
<td>34%</td>
<td>64%</td>
<td>$54.5</td>
</tr>
</tbody>
</table>

*Groups with a population greater than 1 million. Adapted from Flores, 2017.

It is important to highlight this heterogeneity in this demographic segment, as some groups might be more bilingual than others and have different educational and economic trajectories that impact the outcomes. However, this hypothesis will not be explored in this dissertation. In the next chapter, I discuss the theoretical framework that guides this study, and I review previous studies relevant to this dissertation.
Chapter 2: Theoretical Framework and Review of the Literature

Economics and Education

In social sciences such as economics, the term capital is often used in various analyses, including educational analyses. Capital is generally defined as resources that can be accumulated and drawn upon to promote productive activity. Human capital refers to individual and collective knowledge, skills, and other attributes that affect a person’s capacity to do productive work. Human capital is embodied in skills, where the changes in the skills and knowledge of individuals make them more productive (Bourdieu, 1986; Shultz, 1961). Other sources of capital include (1) physical capital, which refers to material resources or physical implements; (2) fiscal or economic capital, which consists of monetary resources or resources; (3) cultural capital, which is composed of long-standing theories, dispositions, and goods from cultural or linguistic heritage; and (4) political capital, which refers to power and influence (Bourdieu, 1986). Notice that at the core of each definition is the fact that each form of capital can be converted to promote a productive activity or tangible good. Each of these constructs is important when analyzing the outcomes of education using the theory of capital in economics.

The analysis of the economics of education dates back to the 1960s when education was first regarded as an investment in human capital (Lovenheim & Turner, 2018). This analysis proposes that education increases the productivity of individuals and, theoretically, their earnings (Becker, 1962; Brown & Sessions, 2007; Johnes & Johnes, 2007). Thus, an investment in education is similar to an investment in physical capital (e.g., machinery). As companies decide to invest in physical capital that makes them more competitive in the market, in the same way, individuals also decide to invest in their education to increase their knowledge and skills, or their human capital, subsequently becoming more marketable. These decisions involve an
examination of the costs and benefits that come from an investment in education (Brown & Sessions, 2007; Johnes & Johnes, 2007).

To thoroughly understand how education works in the economic context, one must first understand how economics contributes to education research and instruction, and how it impacts policy and decision making. Economics provides an insightful lens into the analysis of education in three ways. First, it helps researchers understand the decision-making process of individuals about what type of education to pursue and which one would be a better investment of the individuals’ time and money as it relates to outcomes. Second, economics helps us understand the reactions of individuals to changes in costs and benefits. In other words, analyzing increases in cost and, subsequently, a reduction in the benefits would change the appeal to investing in educational training. While this change can have monetary consequences such as making the pursuit of a specific career more expensive, costs and benefits are not always associated with money, but also with the value of time spent doing other activities. For example, parents may forgo an upward career move offering a higher salary because it may compromise time with their children. Third, causal inference and quantitative analysis techniques from economics provide an essential analytical tool of educational outcomes where researchers approach educational questions borrowing analytical techniques from economics (Lovenheim & Turner, 2018). Several educational issues are approached using statistical methodologies that are commonly used for econometrics.

In summary, the intersection of economics and education allows researchers to evaluate education decisions and their possible outcomes for the individuals regarding earnings through sophisticated quantitative analyses. However, when analyzing education, economic models traditionally do not consider life experiences, relationships and other factors that have an impact
on education. These factors could affect the economic outcomes of individuals through the use of the language spoken at home how its impact on family and community relations.

**Economics of language education.** Language skills are a form of human capital in which individuals and society, in general, would decide to invest in making their individuals more efficient (Grin, 2003). Moreover, people will decide to invest in a specific language as a function of its use in the community and the economic demand for such specific language (Callahan & Gándara, 2014). Economics of language education is the interaction of economic variables with language proficiency variables (Grin, 2008). Similar to other economic interactions, the economics of language education, refers to the study of the allocation of scarce resources of individuals to acquire a new language and how the labor market and other economic forces shape such allocation of resources, for what purposes and for whom. Since the 1990s, studies on the economics of language has considered the importance of language policy issues such as the instruction of languages other than English in schools or instruction in the native language of non-English speakers. These considerations affect the choices that parents, students, and policymakers make about the allocation of financial resources.

Grin (2003) proposes three areas in which economic (e.g., allocation of resources) and linguistic factors (e.g., use of one language over another) interact. First, economic factors affect the use of a language because the relative prices and costs associated with language-specific goods and services will affect how a language is used based on the consumer needs and desire to acquire goods or services provided in a specific language. Furthermore, economic factors affect the use of a language at a grander scale because patterns of international trade influence what language is used for interaction between the trading agents (e.g., suppliers and consumers). For
example, in some areas in the United States, there might be a higher incentive to learn Spanish if there is a need to conduct business with communities who speak Spanish.

Second, language variables affect economic variables by creating income differentials for those of a specific language group and favoring a language over another, creating a form of competition between languages (Grin, 2003, 2008). Similarly, a community where the presence of individuals with a specific language is prominent could stimulate the demand for goods and services in the community’s language. For example, there is a clear demand in the United States for Spanish services in the form of bilingual instruction for the large Latinx community and those jobs often come with a higher salary incentive for individuals (Callahan & Gándara, 2014; García & Kleifgen, 2018; Sanchez, 2017). Similar incentives also occur in other economic sectors such as healthcare, direct client services, public services, and law enforcement, where there are higher wages for those who speak more than one language (Callahan & Gándara, 2014).

Third, economic arguments can be used to either support or oppose language policies, and policymakers face the issue of dealing with costs and benefits while aligning with budget constraints. For example, when discussing the issue of teaching Spanish speakers in the United States through bilingual programs versus English only programs, there is an argument that the costs of bilingual education programs are too high. However, another group argues that the costs are justified as the results from such investment come later for the individual and society in general, in the form of a better prepared and able workforce (Callahan & Gándara, 2014).

In summary, the intersection between economics and language allows us to analyze the use of a language from three different perspectives (i.e., economic factors affecting language use, income differentials for members of different language groups, economic arguments supporting
or opposing language policies), and the results of these analyses can impact educational policy and initiatives that foster or deter the use of a specific language.

**Social capital and linguistic capital.** Social capital is defined as the shared values and understandings that allow individuals and groups to work together through various social engagement activities (Keely, 2007). Moreover, social capital is composed of different entities that are part of some social structure. These entities facilitate the actions of individuals within that social structure. Similar to other forms of capital, social capital makes the completion of certain activities possible. However, social capital is not interchangeable, for it depends on the action to be accomplished (Coleman, 1988).

Coleman (1988), presents that a fundamental difference between social capital and other forms of capital (i.e., physical capital and human capital) where the individual reaps the benefits of investing in such capital (e.g., an individual who invests in acquiring more human capital directly benefits from such investment with possibly better jobs, or more satisfaction) is that there is that the individuals who generate the social capital generally capture only a small part of its benefits. This happens because benefits are distributed among different groups in the community such as the neighborhoods, school volunteering, community organizations that want to improve the community.

Social capital has been applied most frequently to understanding the contributions of schools, families, and communities to the learning and development of children and youth. It has also been applied to analyses of how the social organization of schools relates to teacher learning and development. This conceptualization of social capital helps us understand how social relations in school organizations may promote or impede implementation and change (Smylie & Evans, 2006). In addition, theories of linguistic and cultural capital present language as an asset
that is valued in the labor market (Bourdieu, 1977) because it becomes a connector between groups of individuals in two ways. First, native language allows individuals to establish network connections that influence relations between families and community members, and these agents influence the decisions individuals make during their school life. For example, a mother can communicate with their children on a daily basis about the importance of education, only if both can effectively communicate by for example, using the same language or a combination of two or more languages (e.g., Spanish and English) (Baker, 2011). Second, language allows individuals to fully participate in the U.S. society and be involved in the community at large. This duality of the purpose of languages makes it an asset that increases the opportunities for connections with more than one group, thus increasing the opportunities for developing larger social networks impacting social capital, educational, and ultimately, economic outcomes.

Considering that language is a source of social capital, it is fair to conclude that the advantages of bilingual education go beyond material or financial benefits in the form of more intangible sociocultural benefits for individuals and society. These benefits could take the form of access to diverse resources in multiple languages, diverse educational opportunities, and a more extensive networking with the bilingual community. If these sociocultural assets are nurtured adequately, possibly through an education system that promotes parallel language development (e.g., through bilingual education programs), then those benefits could be potentially transformed into material and financial benefits. On the contrary, if these assets are not nurtured, there are social costs related to assimilation, such as reduced family interactions, community involvement, or value of ethnic and cultural heritage (Agirdag, 2014a, Portes & Hao, 2002). For example, educational and social benefits can come from constant family interactions.
when students can effectively communicate with their parents and parents can be engaged in the education of their children without many language barriers.

One more thing to consider is that language is a naturally occurring asset that individuals already possess by the time they arrive in schools. Individuals who are non-native English speakers already come to school with the advantage of having another language. This natural asset is used to establish relationships between individuals, and if fostered in combination to the development of English, it positively affects their outcomes including educational, economic, social, neurological, and even task-flexibility (Bialystok, Craik, & Luk, 2008; Callahan & Gándara, 2014; Agirdag 2014b; Agirdag 2015; Prior & Gollan, 2011).

**Economic benefits of bilingualism.** Several studies have looked at bilingualism as a predictor of economic benefits, often with mixed results (Agirdag, 2014a; Chiswick, & Miller, 2007; Gándara, 2018; Shin & Alba, 2002). It is then relevant to look at the types of benefits that having two languages provide.

The benefits of bilingualism can be categorized into two groups. First, as market value benefits such as higher earnings, or access to more desirable jobs. Second, non-market value benefits which include networking opportunities, access to material resources in a different language (e.g., educational or political literature), and sensibility to other cultures (Portes & Hao, 2002). To measure market-value benefits, researchers use traditional labor market participation indicators. However, determining the effects of the non-market value benefits is often difficult as most of the data collected are related to salary differentials and jobs performed, and not about other intangible benefits that come from bilingualism, such as socio-cultural connections and communication with parents that could help support the educational endeavors of their children (Baker, 2011).
Grin (2008) proposes a general framework to study individuals learning a language for those who are not native speakers of the dominant language in society and from a perspective that yields returns for the individuals. Through this framework, Grin (2008) proposes that two relevant results arise, (1) learning the dominant language of the receiving country has an advantage over those who do not, and (2) skills in a foreign language can have a positive effect on income. However, this effect is dependent on individual variables that include educational opportunities such as attending bilingual programs and the opportunity to live in a community that speaks another language other than the dominant language (e.g., Spanish in the U.S.). Grin’s (2008) framework is useful in understanding how learning a new language or in the case of this dissertation, maintaining one’s language, in addition to learning the dominant language (i.e., English), can be potentially transformed into an economic benefit. However, few studies have focused on the value of language for individuals who are maintaining a native language while adding the majority language of the country of residence (e.g., a non-English speaker living in the United States who learns English; Agirdag, 2014a). A reason for this could be that the value of a language is perceived by the class who is in power. Hence the value of a language is attributed to those in the dominant class (Bourdieu, 1986).

Other researchers have looked at other ways in which language could be a determinant of income differentials, such as level of bilingualism, meaning the level of proficiency in their native language and the second language they are acquiring. Recent studies demonstrate that individuals who are balanced bilinguals, or individuals who can communicate effectively in two languages, see better economic, educational, and social outcomes (Agirdag, 2014a; Baker, 2011; Callahan & Gándara, 2014). In fact, recent studies on the cost and benefits of bilingualism, demonstrate that in the U.S., individuals who were bilingual, independently of race, earned
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

between $2000 and $3200 more compared to similar individuals who were English-dominant (Agirdag, 2014a, Gándara, 2018). However, in these studies, it is not clear how the researchers defined balanced bilingualism as they did not use objective measures of proficiency for either English or native languages. Defining bilingualism is often a challenge because individuals use language differently in various contexts (e.g., at school vs. at home), and individual data about the level of proficiency in a language other than English are not often collected.

In summary, to understand the outcomes of bilingualism, it is important to look at various factors that influence the benefits of Latinx. First, understanding that there are benefits that are captured by traditional economic and educational measures and that they are influenced by contextual variables such as educational experiences and parental education that must be considered when analyzing the outcomes. Second, few studies in the U.S. have explored how maintaining the native language while adding the English language, benefits bilingual individuals. Moreover, traditional analysis of outcomes for individuals often focus on economic outcomes and often do not account for the importance of personal relations as a means to create trust, establishing expectations, and enforcing norms within the community (Black, Coats, & Goodwin, 1973; in Coleman, 1988). Considering that language is a form of embodied human capital that enhances the production of social capital, it is evident that traditional analyses of outcomes of bilingualism that do not account for the importance of social relations, is not sufficient to evaluate the value of a language, specifically for individuals who are bilinguals. Thus, when analyzing bilingualism, it is then necessary to include how language facilitates social interactions that could affect the individual’s educational and economic outcomes. Finally, it is important to understand the levels of language proficiency of a bilingual individual in either language to find significant outcomes for non-native English speakers.
To add to the literature on outcomes for Latinx bilinguals and to further understand the gap between language, education, economics, and social capital, I used Bourdieus contributions to linguistic capital. Bourdieus theory argues that linguistic competence is a source capital, but only as the market claims the need for a specific language. However, the groups in power dictate whether a language is an asset and hence, the push of some political groups toward a monolingual ideology that started after World War I, and that persists today in the U.S. educational system (Wiley, 2014). Bourdieu (1986), proposes that cultural capital exists in three different forms. First, as an embodied form, owned by the individual through knowledge and skill, allowing connections within groups with the same language. Moreover, language acts as a source of social capital that could prove useful for market-level opportunities. Language is then an embodied sociocultural capital. Second, sociocultural capital can be objectified in the form of books, media, and general information in the minority language (i.e., Spanish). This form of capital provides access to a unique source of information in a specific language and can be a source of specific information that increases the human capital of the individual, hence making the individual more effective in the labor market. In other words, a bilingual individual can access twice the number of sources of information, one in the native language (i.e., Spanish) and another in the second language (i.e., English). Third, linguistic capital can also be institutionalized in the form of academic credentials that are transformed into a future earning advantage. Figure 2, presents a graphic version of this explanation, highlighting the connection between social and cultural capital and economic advantage.
Given the findings from Bourdieu, I then conceptualize that linguistic capital is a form of human capital that increases the chances for relationships within the community and that can have an impact on economic and educational outcomes. Considering this additive perspective of language as a form of capital and its possible impact on economic outcomes, for this dissertation as part of the benefits of bilingualism, I include exploratory analysis of social capital outcomes for Latinx individuals. Specifically, I will explore whether being bilingual facilitates relationships for Latinx in their community at large. For the purposes of this dissertation, social capital will be operationalized as the involvement of individuals in volunteer activities and in civic engagement (Claridge, 2017; Fukuyama, 2000). Next, I review the literature on individual characteristics and outcomes of Latinx bilinguals in the United States.
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

Review of Studies Evaluating Outcomes of Being Bilingual

Traditionally, studies on the effects of bilingualism have been focused on cognitive, educational, and socioemotional outcomes (Agirdag, 2014). A new wave of studies has focused on examining the economic returns of bilingualism for immigrant groups (Agirdag, 2014; Fry & Lowell, 2003; Gándara, 2016; Isphording, 2013; Moore et al. 2014; Robinson-Cimpian, 2014; Rumbaut, 2014). However, there is limited evidence for labor market benefits for those who speak another language in addition to English, and these studies often focus on income as the predominant indicator of success.

Empirically, the relationship between language and income has been studied in various contexts. However, it is often a challenge to separate the effects of language either as membership of a specific ethnic group or as a form of human capital. Moreover, confounding variables (e.g., socioeconomic status, educational attainment, parental education) and limited data on objective measures of language abilities (e.g., native or English) make it difficult to conduct rigorous studies on benefits of bilingualism.

To review the literature on the economic outcomes of bilingualism, I followed the references from previous studies to understand the connection between bilingualism and economic outcomes. Then, I focused on articles that analyzed the outcomes for bilinguals, including Latinx bilinguals in the United States. I used various sources to perform my search, including databases and recent book publications that focused on studies on Latinx individuals. To explore the outcomes, I specifically focused on empirical studies starting in the early 2000s.

To understand how a non-dominant language affects economics, I started reviewing the work of Grin (2003), where he explains how linguistics and economics affect each other. In his review, Grin (2003) concludes that there are four groups of studies on language use that help
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

researchers analyze how language is used, hence helping us understand how value is assigned to a language (Grin, 2003). The first group of studies, focuses on the measurement of language-based discrimination according to the person's first language, confirming that there are earning differentials for people belonging to different language groups even after controlling for their skills.

The second group of studies focuses on the value of the second language when the language is dominant in the region, a typical case for immigrants whose second language is that of the host country. Immigrants who are fluent in the dominant language of their region see benefits. Jorge, Lipner, Moncarz, & Salazar-Carrillo (1983) demonstrated the benefit of bilingualism in areas where trade with other countries is common. Jorge et al. (1983) found that in areas where there are high concentrations of Latinx, being bilingual allowed the area to thrive in commerce with Latin American nations (Jorge, Lipner, Moncarz, & Salazar-Carrillo, 1983).

However, other group of researchers argue that the economic value of bilingualism is often obscured by the desire of other countries wanting to trade with the United States and being the responsibility of those who want to conduct trade with the country, to be fluent in English (Chiswick & Miller, 2007; Alarcón, Di Paolo, Heyman, & Morales, 2014). Alarcón, Di Paolo, Heyman, & Morales (2014a) conducted research in border states that have a Latinx heritage and that conduct commerce with Mexico. They concluded that being bilingual provided a wage differential for individuals. In other words, being bilingual in the United States appears to hold little economic advantage and sometimes even carries a penalty in the United States labor market, except for some areas where there is a need for bilingualism and for specific jobs.

The third group of studies has investigated the values of skills in a non-dominant language region. For example, French-speaking Canadians who have learned English. Results
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

from these studies indicate a positive wage differential for bilinguals (Agirdag, 2014a; Grin, 2003; Vaillancourt, 1996). This positive wage differential is a consequence of the perception of the value of these two languages in Canada, where French and English coexist and are both official languages of the country.

The fourth group of studies is the one of interest for this dissertation. These studies focus on the rates of return of languages of immigrants in their new country. The results for this group of studies are mixed. Grin (2003) found in his review that an immigrant language is often of little value in the country of residence. However, Grin (2003) also argues that an immigrant language might be an asset and not a disadvantage. Using this group of studies as a guide, I searched the literature focusing on studies that researched the economic outcomes of bilingualism for individuals in the United States. The following studies served as the foundation for this dissertation.

In a study by Fry and Lowell (2003), the authors analyzed the nationally representative data from the National Adult Literacy Survey (NALS) to determine wage differentials between bilinguals and monolinguals. This study is relevant because it was the first to provide estimates on wage differentials for bilingual groups. In this study, researchers grouped bilinguals into two groups, those who self-reported speaking English “well” and “very well”, and then compared these groups to English monolinguals. This study found that, indeed, bilinguals, in general, earned more than their monolingual counterparts. However, this positive effect seems to be explained by the higher educational level of bilinguals over monolinguals. Importantly, this study used only self-reported measures of proficiency, and the meaning of speaking English “well” or “very well” may vary by individual. Another limitation of this study is that the sample size from the NALS dataset is small and prevents analysis for different immigrant groups.
Finally, this study did not account for other observable variables that would have made the analysis more robust and account for observable differences between the groups being analyzed, making the comparison more straightforward.

Another study by Shin & Alba (2009) conducted group-specific analyses to examine the extent to which bilingualism, along with other characteristics, influenced earnings for major Latinx and Asian groups in the United States. The sample was restricted to individuals who were between the ages of 25-64, who worked at least 160 hours in the public and private sectors and who were in the 1.5 generation segment. Moreover, the authors used metropolitan-level variables to account for contextual variables in different areas of the country. In this study, researchers conducted a series of ordinary least square (OLS) models to determine the gap between English monolinguals and bilinguals. During these analyses, no individual observable variables were included. Researchers found that there was little sign of a general economic advantage for bilinguals over English monolinguals. In fact, there was a penalty for the groups analyzed. However, the authors argue that the retention of the native language is a proxy for maintaining other valuable assets within their culture. Similar to the Fry & Lowell (2003) study, this study used self-reported measures of English proficiency. Moreover, there was not a clear measure of proficiency for either language as proficiency was self-reported. Also, in this study, only a sample of each group analyzed was selected, Mexicans, Cubans, and Dominicans for Latinx, and Chinese, Filipino, and Korean for Asians. While these groups are representative of these two demographic segments, they differ on many characteristics, such as educational trajectory, immigration patterns, and educational opportunities. Finally, while the ordinary least square (OLS) regression models used in the study provide a strong analysis mechanism, it would be
important to account for differences between groups through other procedures that compensate for the lack of randomized control trials.

More recent studies have explored whether there was a benefit for a specific population of bilinguals. Robinson-Cimpian (2014) used data from the 2006-2010 American Community Survey (ACS) to analyze how labor market participation, employment rates, and income changes between bilingual and monolingual English speakers. The sample was restricted to Latinx U.S. citizens between the ages of 24-64 who were not enrolled in school. The analyses were conducted through a series of OLS models to estimate random intercepts using two levels, (1) individual levels and (2) using codes from the Public Use Microdata Area to account for differences by metropolitan areas. The results from these analyses indicated that bilingual males were less likely than monolinguals to participate in the labor market, while bilingual females were more likely to participate in the labor market than monolinguals. Regarding wages, Latinx bilingual males had a disadvantage in wages over monolingual Latinx. Moreover, Robison-Cimpián (2014) concluded that while bilingualism is a desired skill in the workplace, it is not rewarded at a premium by the market for the jobs analyzed. In this study, it is not clear how bilingualism was determined for the individuals as there is not an objective measure of English proficiency. Moreover, in the dataset used for analyses, there is not a clear way to determine other observable differences such as individual and parental educational levels that have an impact on the wage of individuals.

In two studies by Alarcón, Di Paolo, Heyman, & Morales (2014a, 2014b), authors used qualitative analysis and census data to explore how bilingualism plays a part in the workplace in various occupational roles in a specific geographic area where bilinguals and binational individuals live (i.e., U.S. borderlands with Mexico). For the quantitative portion of the study,
researchers used the level of bilingualism (i.e., English monolingual, limited English proficient, and fluent bilingual). Moreover, researchers included covariates such as age, gender, and interactions, marital status and number of children, place of origin, and educational attainment. Using multinomial regression models, researchers found that the level of bilingualism could have a positive impact on how individuals participate in the economy. However, there were limitations to how far individuals could climb professionally. For example, fluent bilinguals are more likely to have occupations that involve intensive oral interaction with the public but had a lower probability of holding professional and managerial positions. Also, limited English proficient speakers were more likely to hold service positions with minimal written and limited oral skills and physical labor jobs. In other words, fluent bilingualism allows individuals to work above the low-skill services and manual labor, but below credentialed occupational sectors even when controlling for education. Researchers concluded that bilingualism could be rewarded at a premium, but it occurs through informal labor market opportunities and it is related to the work role and not through deliberate organizational policies that reward bilingualism. The main limitation in the findings is the language variable, as it was self-assessed. Moreover, these findings are very specific to a region, and the analyses did not control for other observable variables that influence the outcomes of bilingual individuals. Finally, it did not provide a way to compare individuals based on a set of observable variables that account for systematic differences between the comparison groups.

Another study further explores whether there are positive outcomes for bilinguals. Agirdag (2014a) hypothesized that bilinguals will earn more because they can take positions in specific language minority subfields in addition to the regular market. For this quantitative study, the researcher used two data sources, the National Education Longitudinal Studies (1988) and the
Children of Immigrants Longitudinal Studies of 1991/2003. Agirdag (2014a) defined bilingualism through latent class analysis (LCA) and created three groups (1) limited bilinguals, (2) balanced bilinguals, and (3) English dominant. In the analyses, he included covariates such as educational attainment, cognitive ability, parental SES, geographic location, and national origin. This last variable was limited as only partial information was available from either dataset. Results indicated that there are significant costs associated with linguistic assimilation and that balanced bilinguals are more likely to be employed than English-dominant individuals. Moreover, balanced bilinguals earned between $2000 and $3200 more annually. Similar to previous studies, these analyses are limited in the use of the language variable, since it is self-reported for both languages. Also, this study only focused on individuals who are early in their career and not for the entire bilingual population, limiting the results for this group of workers. Finally, these results are not indicative of any specific population (e.g., Latinx) but for bilinguals in general. This lack of specificity can be problematic because there are systematic differences between immigrant groups (e.g., access to labor markets, educational background, parental education, socio-economic status) that will influence the income levels of each group.

Other researchers have found positive results associated with balanced bilingualism and educational, occupational, and economic attainment. However, researchers argue that the value of bilingualism may not come as a higher pay but rather in the hiring process, where a bilingual candidate may be chosen over a monolingual (Porras, Ee & Gándara, 2014; Rumbaut, 2014).

It can be argued that these results have promoted a resurgence of bilingual programs and a rise in the demand for dual language programs, where students develop their native language plus another language simultaneously (Polanco & Baker, 2018; Sanchez, 2017). Moreover, in recent years, there has been a political shift of the English-only policies. In November 2016,
California overwhelmingly voted to overturn Proposition 227. According to the new law, Proposition 58, schools have the flexibility to design their programs based on the needs of their bilingual students. The law also requires that these programs are discussed with the community (Hopkinson, 2017). Similarly, in November 2017, the state of Massachusetts, overturned a 15-year ban on bilingual education, allowing for more flexibility when teaching bilingual students (Vaznis, 2017). Interestingly, this demand has come from middle-class English-speaking parents who see an advantage of their children being bilingual in an increasingly more global world (Gándara, 2018; García & Kleifgen, 2018).

After reviewing these studies about the economic outcomes of being bilingual, there are several limitations in the analyses presented. First, there is evidence that some bilinguals are performing better in economic measures. However, given that these studies have used only census data, it is a challenge to define the bilingualism of individuals. While objective measures on native language proficiency are often limited in longitudinal datasets, a better assessment of the English language through objective measures would help us understand their domain of the national language. Previous studies have indicated that immigrant individuals who handle the native language of the host country, see better outcomes, so it will be important to have a better measure of English proficiency than a self-reported measure (Chiswick 1978; Grin, 2003, 2010).

Second, only a few studies reviewed controlled for geographic variables, and studies demonstrate that the effects of bilingualism systematically vary according to geographical areas where there are higher concentrations of bilingual individuals. These concentrations of bilingual individuals will push the market for more bilingual services. Moreover, it is important to account for the differences in areas where these individuals live, study, and work as there are systematic
differences between these locations. Hence, there is a need to investigate income differentials based on the geographic location of the home and the school.

Third, benefits are defined as a function of wages and not of other assets that the individual already possesses. Questions remain about outcomes beyond wages, such as access to higher education and social benefits. For example, it is not clear how credentialing affects access to jobs or reasons why bilinguals might be taking lower-paying jobs. Moreover, there are limits to how bilingualism is valued in the labor market in the United States. Since the evidence supports that education is a factor on wages, it will be important to know how bilinguals are performing in educational variables compared to monolinguals.

Fourth, it is not clear why individuals are earning less other than discrimination. However, there might be a desire of bilingual individuals to engage in activities that are perhaps beneficial to their community, but that do not necessarily represent a higher income. It will be important to address how individuals are engaged in their community and whether bilingualism facilitates this engagement in community-focused activities.

Fifth, when studying how bilingualism affects the outcomes of immigrant individuals, it is important to include parental-level variables such as parental education and household income, since these variables affect how the individuals perform in schooling and consequently in the labor market.

Finally, it is fair to conclude that the costs of language assimilation come not only in the economic form but also as strained family relations caused by poor parental communication with children, when the parents do not speak the second language, lack of access to possible jobs that require the domain of Spanish in specific areas, lack of network opportunities, and material resources in their native language that could foster a higher level of communication and sense of
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

belonging (Bourdieu, 1977; Porras, 2002; Stanton-Salazar & Dornbusch, 1995; Santibañez & Zárate, 2014). This lack of access can also create a lack of opportunities for success for bilinguals.

Recent studies have shown that when bilingualism is analyzed from an asset perspective, bilingual individuals earn more than similar individuals who are English-only speakers. In other words, when analyses value the native language skills that individuals possess as well as include the social capital benefits that language brings, there is a positive economic outcome for bilingual individuals (Agirdag, 2014a). For the Latinx population, the effects are similar. When Latinx are bilinguals, they earn more than the same Latinx who are monolinguals, which places them at an advantage in the labor market and only for certain careers, specifically those who were in an geographic area where language is considered a public good (Alarcón, Di Paolo, Heyman, & Morales, 2014b; Moore, et al., 2014; Robinson-Cimpián, 2014). However, this is still an area that merits more study to better understand the variables that have a direct impact on the outcomes.

When analyzing the effects of language, they are often framed as a direct correlation between language skills and income differentials. However, these analyses often overlook skills in either language that could impact other forms of labor market outcomes, such as rates of participation, types of jobs, or differentials by country of origin (Polanco & Richards, 2016; Robinson-Cimpián, 2014). Moreover, few studies have analyzed the social capital assets that are inherent to a language and the costs associated with language assimilation (Agirdag, 2014a; 2015). Finally, until recently, analytical methodology limited the generalizability of the conclusions as they did not account for other unobservable differences between the groups analyzed. With this dissertation, my goal is to expand the knowledge on participation rates of
Latinx in American society through the analysis of a large longitudinal dataset and using novel methodologies that increase the internal validity of the analyses.

When analyzing the outcomes from bilingualism, it is crucial to clearly define the levels at which individuals can communicate in either language. Similar to previous studies, I explore the economic outcomes of Latinx individuals, but I will also look at indicators of access to the labor market, such as educational attainment and job market participation. Moreover, I explore how Latinx individuals participate in other activities that could increase their social capital and possibly benefit them and their community. Unlike previous studies, in this dissertation, I use an objective measure of English proficiency and contextual variables for the native language as information on the latter is rarely collected. To analyze this set of outcomes, first, I match individuals using a novel matching technique to account for unobservable characteristics. Then, I include several contextual variables such as generational status, parental-level variables (e.g., parental education, parental language, and parental income), and geographic location using zip codes as fixed effects. In summary, I use a novel statistical matching technique to compare similar groups of Latinx and then analyze four different outcomes (1) educational attainment, (2) labor market participation, (3) income, and (4) social engagement operationalized as volunteering in social activities. Next, I explain the methodology used to perform these analyses.
Chapter 3: Method

Research Questions

In this dissertation, I analyzed the longitudinal effects of being bilingual on educational and economic variables. Moreover, consistent with Bourdieu’s linguistic capital theory, I explored whether language has an impact on the engagement of individuals on social capital development activities. Specifically, the questions for this study are as follows:

Q1. Are Latinx bilinguals more likely to participate in any post-secondary education programs compared to Latinx non-bilinguals?

Q2. Are Latinx bilinguals more likely to have a job compared to Latinx non-bilinguals?

Q3. Are there income differentials for Latinx bilinguals compared to Latinx non-bilinguals?

Q4. Are Latinx bilinguals more likely to be engaged in volunteering activities compared to Latinx non-bilinguals?

In this dissertation, bilingualism is perceived as an asset. Hence, after controlling for a series of characteristics, I hypothesize that Latinx bilinguals will be more likely to have post high school degrees compared to Latinx non-bilinguals. Moreover, I hypothesize that they will be more likely to have a job and a greater income. Finally, I hypothesize that Latinx bilinguals will be more engaged in their community as they can communicate with a larger group of individuals than Latinx non-bilinguals.
To answer the four research questions, I used the Educational Longitudinal Study of 2002/2012 (ELS:2002) dataset, administered by the National Center for Educational Statistics (NCES) of the Institute of Education Sciences, U.S. Department of Education. This nationally representative sample represents over 3.4 million individuals in 10th grade starting in 2002. The students were followed for over ten years and information about their transition from high school into postsecondary years, and the workforce was collected. The data also include surveys from parents, mathematics and English Language Arts teachers, and school administrators. The last survey collection was conducted in 2012 and included information about student families, their employment status, income, and community involvement. The dataset also includes information about parental educational level, individual English proficiency, and native language, as well as whether students engaged in activities that developed their social capital, such as volunteer activities.

To determine English language proficiency, I used an objective measure of proficiency, based on school records that measured the level of proficiency of students. I considered students who were fluent or above on this measure as the English language proficient. As for native language proficiency (i.e., Spanish), there was limited information on students’ level of communication. The only available sources of Spanish proficiency were the student records indicating whether individuals and their parents spoke Spanish at home or not.
Dependent Variables

I used the following dependent variables: any post-secondary education, job participation, income, and volunteering.

**Any post-secondary education or above.** To answer the first question about the individuals’ likelihood of attending a post-secondary institution, I used the educational attainment variable. I defined this variable as anyone who completed high school and who attended any post-secondary institution. I created a composite categorical variable where one group participated in any post-secondary education program, and another group did not.

**Job participation.** To answer the second research question about whether bilinguals were more likely to participate in the labor market versus non-bilinguals, I looked at the variable that specifically indicated whether Latinx individuals had ever had a job after their high school years. These data were collected at the third time point. Similar to the previous variable, I transformed the original variable into a dichotomous variable indicating whether individuals have ever had a job after high school or not.

**Income.** To answer the third question about whether Latinx bilinguals are more likely to receive a higher income compared to Latinx non-bilinguals, I used the income variable collected at the third follow-up. For ease of interpretation, I conducted the analysis using the original continuous variable for income.

**Volunteering.** To answer the fourth question about whether Latinx bilinguals are more likely to participate in volunteer activities compared to non-bilinguals, I created a composite variable labeled “volunteer” that included participation in activities that develop the social capital of individuals as adults such as volunteering for different activities at church, in youth groups, or in community organizations.
Covariates. I controlled for other variables that could affect results beyond bilingualism. I will include gender (female, male), generational status (first, second, and third), parental language (Spanish or not), parental education (college or above, or not), and household income as defined by three categories, (1) lowest family income between $10,000 and $35,000, (2) medium, $35,001 and $100,000, and (3) high for $100,000 and above. I also include fixed effects of location by using zip codes where students completed their high-school education.

Data Analysis Procedure

Coarsened Exact Matching Strategy. In this study I use “coarsened exact matching” (CEM) procedures to match Latinx balanced bilingual students with a similar comparison group of Latinx non-balanced bilinguals. CEM offers several advantages over other matching (e.g., monotonic imbalance bounding) techniques such as propensity score matching (Iacus, King & Porro, 2012). First, the user performs the balance check and the maximum imbalance between treatment and control groups ex-ante, as opposed to performing it after the analysis, increasing the amount of balance check to get the desired match in traditional matching techniques. Second, adjusting the maximum balance on one variable does not impact the maximum balance on other variables as it would happen in other traditional matching techniques (see, for example, Iacus, King & Porro, 2012). This happens because the user is not running and rerunning the matching procedure as it is required with other techniques such as the traditional propensity score matching. Finally, in addition to the efficiency of using a more straightforward method for matching, recent studies suggest that CEM outperforms traditional methods for matching by reducing imbalance, providing better error estimation, and model dependence to estimate effects (Iacus, King & Porro, 2009, 2011).
To perform a CEM, each variable is “coarsened” or recoded to group the variables by assigning them a numerical value through an algorithm. Then, the “exact matching” algorithm is applied to the recoded data to determine the matches and to drop the unmatched units. In other words, the CEM algorithm creates a set of strata with the same coarsened values and matches it to a pair (e.g., treatment and control group). Then, the algorithm eliminates the differences between the treatment and control groups. Only the matched data are retained (Iacus, King & Porro, 2011).

**Latinx sample.** For this dataset, I defined Latinx as anyone who selected Hispanic or Latino on the survey, independently of whether they specified their race. Out of the total sample in the data, 15% (\(N=542,254\)) met this requirement. After matching on the four variables mentioned above (i.e., gender, generational status, parental native language, and parental education), the matched sample represented 85% of 542,251 (i.e., \(N=460,801\)). Out of these matched individuals, 59% (\(n=271,802\)) were non-bilinguals, while the rest (\(n=188,999\)) were considered bilinguals. I define a Latinx bilingual as an individual who reports to speak Spanish at home and whose English proficiency is fluent or above as determined by the state language proficiency assessment. Table 4 provides a summary of the characteristics of participants in this study.

Given the limited availability of data on native language proficiency and type of program participation (i.e., bilingual vs. English only), I used objective and contextual variables as indicators of language proficiency. To determine participant English language proficiency, I used individual records that include scores from an objective assessment of achievement in English proficiency that was conducted as part of the ELS:2002 survey. Consistent with previous research, I decided that those who scored “fluent” or above were English proficient (Robinson-
To determine the Spanish language proficiency of the sample population, I selected students who indicated that their native language was Spanish. There was not a summative assessment of student native language proficiency.
### Table 4

**Summary of Descriptive Statistics of Matched Sample**

<table>
<thead>
<tr>
<th></th>
<th>Latinx</th>
<th>Gender</th>
<th>Parental Education</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilingual</td>
<td>41%</td>
<td>45%</td>
<td>26%</td>
</tr>
<tr>
<td>Non-bilingual</td>
<td>59%</td>
<td>45%</td>
<td>26%</td>
</tr>
<tr>
<td><strong>Number</strong></td>
<td>188,999</td>
<td>271,802</td>
<td>74%</td>
</tr>
<tr>
<td>Male</td>
<td>45%</td>
<td>45%</td>
<td>26%</td>
</tr>
<tr>
<td>Female</td>
<td>55%</td>
<td>55%</td>
<td>74%</td>
</tr>
</tbody>
</table>

|                        |        |        |                    |
| **Home Language**      |        |        |                    |
| Spanish                | 87%    | 43%    | Below $10,000      |
| Not Spanish            | 13%    | 46%    | $10,000 to $35,000 |
| 1<sup>st</sup> Gen     | 43%    | 43%    | 10%                |
| 2<sup>nd</sup> Gen     | 12%    | 45%    | 57%                |
| 3<sup>rd</sup> Gen     | 10%    | 46%    | $35,000 to $100,000|
| $100,000 and above     | 2%     | 2%     |                   |
Matching procedure. To determine the sample, for all dependent variables, I matched each Latinx balanced bilingual with a Latinx non-bilingual on gender, generational status, parental native language, and parental education. According to previous studies, these four characteristics can affect economic, educational, and social outcomes (Alarcón, Di Paolo, Heyman, & Morales, 2014b; Robinson-Cimpián, 2014). Therefore, a matching procedure on these characteristics, allows me to compare individuals with the same characteristics on these four variables and whose only difference is whether they are bilingual or not. Next, I explain the model specifications and how I conducted the statistical analyses.

Model Specification

Consistent with similar studies that evaluate the academic achievement of Latinx, I used general linear regression techniques to estimate the likelihood that Latinx bilinguals would have any post-secondary education, differentials in job participation, differentials in family income, and likelihood of participating in social activities. Moreover, I also used a second model of logistic regression for dichotomous variables (post-secondary, job participation, and social capital) for double robustness purposes. The purpose of using two models is to prove that the estimates are consistently true across different models, making the user more confident about the results. Doubly robust models provide stronger analytical methods with higher internal validity. However, it must be noted that one cannot completely discount that there might be other variables affecting outcomes that are not considered in the model. Also, in this dissertation, I analyzed the interactions between bilingualism and gender, and bilingualism and generational status.
Post-secondary education. For this categorical variable, I estimated the effects of being balanced bilingual on the likelihood of having any post-secondary education or above and job participation through a series of logistic regression models with fixed effects to account for sources of heterogeneity. I estimate the logistic regression outcome through the following general model:

\[ \text{logit}(P_i) = \alpha + \beta_1(BIL_i) + \beta_x(X_i) + \mu_i \]

Where \( P \) is a dichotomous indicator representative of whether a student had any post-secondary education or had ever participated in the job market; \( \alpha \) is a non-time varying fixed effect of location for each individual; \( X \) is an array of non-time varying characteristics for each student \( i \) (e.g., gender, generational status, parental education, parental language, household income) and \( \mu \) represents the error term. I also ran models with interactions between bilingualism and gender and generational status of the group using the following general equation:

\[ \text{logit}(P_i) = \alpha + \beta_1(BIL_i) + \beta_x(X_i) + (BIL_i \ast X_i) + \mu_i \]

Job participation. I estimated the effects of being bilingual on the likelihood of having any job participation through a series of logistic regression models with fixed effects to account for sources of heterogeneity. I estimate the logistic regression outcome through the following general model:

\[ \text{logit}(P_i) = \alpha + \beta_1(BIL_i) + \beta_x(X_i) + \mu_i \]

where \( P \) is a dichotomous indicator representative of whether a student had ever participated in the job market; \( \alpha \) is a non-time varying fixed effect of location for each individual; \( X \) is an array of non-time varying characteristics for each student \( i \) (e.g., gender, generational status, parental education, parental language, household income) and \( \mu \) represents the error term. A model for interactions of gender, generational status, and parental education was also analyzed. I
also ran models with interactions between bilingualism and gender and generational status of the group using the following general equation:

$$\logit(P_i) = \alpha_j + \beta_1(BIL_i) + \beta_x(X_i) + (BIL_i \ast X_i) + \mu_i$$

**Income.** To estimate the effects of being balanced bilingual on the level of income, I use a general linear regression on a continuous income variable through the following general model:

$$y_{inc} = \alpha_i + \beta_1(BIL_i) + \mu_i$$

Where y is the continuous outcome for income; \(\alpha\) is a non-time varying fixed effect of location for each individual \(i\), BIL is whether the student is considered a bilingual, and \(\mu\) represents the error term. As before, treatment weights are included in the models. A model for interactions of gender, generational status, and parental education was also analyzed.

$$y_{inc} = \alpha_i + \beta_1(BIL_i) + \beta_x(X_i) + \mu_i$$

**Social Capital.** I estimated the effects of being bilingual on the likelihood of participating in activities that develop social capital through a series of logistic regression models with fixed effects to account for sources of heterogeneity. For this exploratory analysis, I ran a regression on a composite variable labeled volunteer. This variable was defined by ever participating in any type of volunteer activities as adults (i.e., community services, youth groups, church groups, etc). I estimate the logistic regression outcome through the following general model:

$$\logit(P_i) = \alpha_j + \beta_1(BIL_i) + \beta_x(X_i) + \mu_i$$

where \(P\) is a dichotomous indicator representative of whether a student had ever participated in any social capital activity; \(X\) is an array of non-time varying characteristics for each student \(i\) (e.g., gender, generational status, parental education, household income) and \(\gamma\) is a non-time varying fixed effect of location for each individual.
This section reports the results of the regression estimates using bilingualism as a predictor of the four outcomes selected for this study (i.e., participation in post-secondary education, participation in the job market, income, and social capital development activities). For all the analyses, Latinx individuals were matched a priori on an array of observable variables such as gender, generational status, parental language, parental education, and household income before the analysis. The regression analyses were performed with the same variables as covariates to control for their effect. Moreover, the analyses include fixed effects for residential zip codes to ensure that students are compared to similar students in the same geographic area. The inclusion of zip codes is consistent with Robinson-Cimpián’s (2014) research, which demonstrates that there are differences in outcomes for Latinx depending on their geographic location where they live and work. As a robustness check, I estimated these effects using linear and logistic regression models. All analyses include a comparison of Latinx with very similar characteristics, who are either bilingual or non-bilingual. Moreover, analyses include comparisons between genders, generational status, and different levels of household incomes.

**Participation in Post-Secondary Education**

Table 5 reports the regression estimates on bilingualism as a predictor of participation in any post-secondary education program. Results demonstrate that on average Latinx in general are significantly less likely to participate in post-secondary education than non-Latinx. However, regression estimates from a binary logistic model indicated that bilingual Latinx have 2.68 times the odds of participating in post-secondary education than non-bilingual Latinx. Results of linear probability models provide a more straightforward interpretation of this effect, suggesting that bilingual Latinx are 54.9 percentage points more likely than non-bilingual Latinx to have any
post-secondary education. The interaction between gender and education shows a significant interaction effect where Latinx women are slightly less likely than men to attend any post-secondary education at 63 and 66 percentage points respectively. However, bilingualism increases the chances of women attending any post-secondary education (Figure 3).

![Figure 3](image_url)

**Figure 3.** Interaction of gender and bilingualism predicting the likelihood of attending any post-secondary education institution.

Regarding generational status, in general, the largest impact of bilingualism is for first generation bilingual Latinx individuals who were 55 percentage points more likely to have a post-secondary education compared to non-bilinguals (25 percentage points). This bilingual advantage seems to disappear for second generation Latinx individuals. In this comparison both, bilingual Latinx and non-bilingual Latinx had the same likelihood of attending any post-secondary education. For third generation individuals, bilingual Latinx were 9 percentage points more likely to attend any post-secondary institution than non-bilinguals. Thus, bilingualism is an advantage for first and third generation Latinx individuals, but not for second generation bilinguals as illustrated in Figure 4.
Figure 4. Interaction between generational status and bilingualism as a predictor of the likelihood of attending any post-secondary education institution.

Considering home language, individuals whose parents speak Spanish are 4.6 percentage points less likely to have any post-secondary education. Individuals of parents who attained a college degree were 18.1 percentage points more likely to have post-secondary education. Finally, in terms of income, bilingual Latinx whose parental income is in the highest category (100K and above), were the most likely to attend a post-secondary educational institution.
### Table 5

**Summary of Main Effects and Interactions in Binary and Linear Probability Models for Variables Predicting Participation of Latinx Bilinguals in any Post-Secondary Education (N = 324,233)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Main Effects</th>
<th>Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Binary Log. Model</td>
<td>Linear</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>EXP B</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.434</td>
<td>0.238</td>
</tr>
<tr>
<td>Bilingual</td>
<td>0.985</td>
<td>2.678</td>
</tr>
<tr>
<td>Gender</td>
<td>0.885</td>
<td>2.423</td>
</tr>
<tr>
<td>Bilingual*Gender</td>
<td>0.335</td>
<td>1.398</td>
</tr>
<tr>
<td>Gen. Status (3rd)</td>
<td>1.106</td>
<td>3.022</td>
</tr>
<tr>
<td>Gen. Status (2nd)</td>
<td>1.662</td>
<td>5.270</td>
</tr>
<tr>
<td>3rd Gen. * Bil.</td>
<td>-1.739</td>
<td>0.176</td>
</tr>
<tr>
<td>2nd Gen. * Bil.</td>
<td>-2.247</td>
<td>0.106</td>
</tr>
<tr>
<td>Parental Lang. Sp.</td>
<td>-0.352</td>
<td>0.703</td>
</tr>
<tr>
<td>Parental Education</td>
<td>1.964</td>
<td>7.128</td>
</tr>
<tr>
<td>Inc. 100K &amp; above</td>
<td>20.323</td>
<td>6.701E+08</td>
</tr>
<tr>
<td>Inc. 35K to 100K</td>
<td>0.216</td>
<td>1.241</td>
</tr>
<tr>
<td>Inc. 10K to 35K</td>
<td>-0.291</td>
<td>0.748</td>
</tr>
<tr>
<td>Zip Codes</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

*Note*: *p < .05. **p < .01. p<.001. Coefficients for zip codes fixed effects have been suppressed for parsimony. In a binary probability model the exponential B indicates the odds of participating in post-secondary education. In both models the B represents the percentage that a person from a specific group is likely to participate in post-secondary education. Models are estimated using robust standard errors. Variable “Inc.” refers to parental income at base year.
Job Market Participation

Table 6 reports the regression estimates on bilingualism as a predictor of ever having a job. Results of linear probability models provide a more straightforward interpretation of this effect, suggesting in general, Latinx bilinguals are significantly more likely to have a job than non-bilinguals, where bilingual Latinx are 4.3 percentage points more likely to have a job than non-bilinguals. Regarding gender, overall Latinx women, whether they are bilingual or not, are less likely to participate in the job market (92.1 percentage points). However, when considering bilingualism, bilingual Latinx women are more likely than bilingual Latinx men to participate in the job market (98.1 percentage points) (Figure 5).

![Figure 5](image.png)

**Figure 5.** Interaction between gender and bilingualism as a predictor of participating in the job market.

In terms of generational status, in general third generation Latinx are 10.1 percentage points more likely to participate in the job market than the first generation, while second generation Latinx are 5.8 percentage points more likely to ever have had a job. However, an interaction of bilingualism and generational status revealed that third generation bilingual Latinx
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

are 6.1 percentage points less likely than non-bilingual Latinx to participate in the job market.

For second generation Latinx, bilingual Latinx are 3.4 percentage points less likely to participate in the job market than non-bilinguals Latinx. These findings reveal that second and third generation bilingual Latinx are less likely to ever have had a job than the first generation Latinx individuals. Thus, bilingualism seems to be advantageous only for first generation Latinx individuals (Figure 6). Other findings from the analysis show that Latinx individuals whose parents’ native language is Spanish are slightly less likely to participate in the labor market. However, Latinx individuals whose parents have a bachelor degree or above as well as individuals whose parents are in the lowest and highest income brackets, seem to be more likely to participate in the labor market.

Figure 6. Interaction between generational status and bilingualism as a predictor of the likelihood of participating in the labor market.
Table 6
Summary of Main Effects and Interactions in Binary and Linear Probability Models for Variables Predicting the Likelihood of Latinx Bilinguals Participating in the Labor Market (N = 324,233)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Main Effects</th>
<th>Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Binary Log. Model</td>
<td>Linear</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>EXP B</td>
</tr>
<tr>
<td>Intercept</td>
<td>18.451</td>
<td>1.031E+08</td>
</tr>
<tr>
<td>Bilingual</td>
<td>3.005</td>
<td>20.186</td>
</tr>
<tr>
<td>Gender</td>
<td>-2.481</td>
<td>0.084</td>
</tr>
<tr>
<td>Bilingual*Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen. Status (3rd)</td>
<td>2.576</td>
<td>13.144</td>
</tr>
<tr>
<td>Gen. Status (2nd)</td>
<td>2.340</td>
<td>10.381</td>
</tr>
<tr>
<td>3rd Gen. * Bil.</td>
<td>-3.498</td>
<td>0.030</td>
</tr>
<tr>
<td>2nd Gen. * Bil.</td>
<td>2.477</td>
<td>11.905</td>
</tr>
<tr>
<td>Parental Lang. Sp.</td>
<td>0.638</td>
<td>1.893</td>
</tr>
<tr>
<td>Parental Education</td>
<td>2.662</td>
<td>14.325</td>
</tr>
<tr>
<td>Inc. 100K &amp; above</td>
<td>21.768</td>
<td>2.843E+09</td>
</tr>
<tr>
<td>Inc. 35K to 100K</td>
<td>2.411</td>
<td>11.145</td>
</tr>
<tr>
<td>Inc. 10K to 35K</td>
<td>2.298</td>
<td>9.954</td>
</tr>
<tr>
<td>Zip Codes</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

Note: *p < .05. **p < .01. p<.001. Coefficients for zip codes fixed effects have been suppressed for parsimony. In a binary probability model the exponential B indicates the odds of participating in post-secondary education. In both models the B represents the percentage that a person from a specific group is likely to participate in post-secondary education. Models are estimated using robust standard errors. Variable “Inc.” refers to parental income at base year.
**Income**

Table 7 shows a summary of the regression analyses using an array of variables to predict the income of individual Latinx bilinguals. For these analyses, I used two variables, individual employment income and individual employment income plus spouse or partner (if applicable). Results from the linear regression model analyses suggest that, on average, Latinx bilingual individuals earn $7,655 more annually than non-bilinguals, and bilingual households make $8,315 more than non-bilingual households.

For individual Latinx women there is a negative effect of income, where they earn $8,567 less annually than men. However, when they are bilingual, they earn $873 more annually than non-bilingual Latinx women (Figure 7).

*Figure 7. Interaction between gender and bilingualism as a predictor of income.*
As for generational status, results indicate that second generation Latinx individuals earned $11,068 more than the first generation, while the third generation Latinx earned $5,518.96 more than first. However, there was a negative effect when accounting for bilingualism, where second and third generation bilingual Latinx earned less than the first generation (Figure 8).

Figure 8. Interaction between generational status and bilingualism as a predictor of income.
### Table 7

**Summary of Main Effects and Interactions in Linear Probability Models for Variables Predicting the Income of Latinx Bilinguals (N = 324,233)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Main Effects</th>
<th></th>
<th>Interactions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>Household</td>
<td>Individual</td>
<td>Household</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>B</td>
<td>SE B</td>
</tr>
<tr>
<td>Intercept</td>
<td>9087.927</td>
<td>328.055</td>
<td>15043.468</td>
<td>605.489</td>
</tr>
<tr>
<td>Bilingual</td>
<td>3118.396</td>
<td>76.279</td>
<td>4017.213</td>
<td>119.591</td>
</tr>
<tr>
<td>Gender</td>
<td>-8257.382</td>
<td>78.838</td>
<td>-252.394</td>
<td>117.731*</td>
</tr>
<tr>
<td>Bilingual*Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen. Status (3rd)</td>
<td>1294.592</td>
<td>139.318</td>
<td>387.371</td>
<td>239.321(NS)</td>
</tr>
<tr>
<td>Gen. Status (2nd)</td>
<td>6649.101</td>
<td>89.290</td>
<td>5192.850</td>
<td>134.566</td>
</tr>
<tr>
<td>3rd Gen. * Bil.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd Gen. * Bil.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental Education</td>
<td>-2628.031</td>
<td>105.153</td>
<td>-2401.460</td>
<td>156.531</td>
</tr>
<tr>
<td>Inc. 100K &amp; above</td>
<td>6264.458</td>
<td>358.616</td>
<td>14680.932</td>
<td>775.166</td>
</tr>
<tr>
<td>Inc. 35K to 100K</td>
<td>3386.124</td>
<td>177.026</td>
<td>1442.062</td>
<td>224.537</td>
</tr>
<tr>
<td>Inc. 10K to 35K</td>
<td>2362.042</td>
<td>174.160</td>
<td>66.107</td>
<td>207.733</td>
</tr>
<tr>
<td>Zip Codes</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

*Note: *p < .05. **p < .01. p<.001. Coefficients for zip codes fixed effects have been suppressed for parsimony. In a binary probability model the exponential B indicates the odds of participating in post-secondary education. In both models the B represents the percentage that a person from a specific group is likely to participate in post-secondary education. Models are estimated using robust standard errors. Variable “Inc.” refers to parental income at base year.*
Social Capital Development Activities

Volunteer. Table 8 reports the regression estimates on bilingualism as a predictor of participation on volunteer activities. Results of linear probability models provide a more straightforward interpretation of this effect, indicating that in general Latinx individuals are more likely to participate in volunteer activities as adults. When accounting for bilingualism, bilingual Latinx are 2 percentage points less likely than non-bilinguals to participate in volunteer activities as adults. This negative effect is slightly larger for women. It is important to notice that the number of individuals who reported ever doing any volunteer work is only about 20% of the matched sample \((n=96,618)\). Regarding generational status, second generation Latinx were 2 percentage points more likely than the first generation to ever have done volunteer work. Latinx individuals who grew up in households in the $10,000- and $35,000-income bracket were 10.6 percentage points more likely to be involved in volunteer activities than individuals in the $35,000 to $100,000 (5.8 percentage points) and individuals in the $100,000 or above (3.7 percentage points).
Table 8
Summary of Main Effects and Interactions in Binary and Linear Probability Models for Variables Predicting the Likelihood of Latinx Bilinguals Participating in Volunteer Activities (N = 96,618)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Main Effects Binary Log. Model</th>
<th>Linear</th>
<th>Interactions Binary Log. Model</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>EXP B</td>
<td>SE B</td>
<td>B</td>
</tr>
<tr>
<td>Intercept</td>
<td>12.067</td>
<td>1.740E+05</td>
<td>3.682</td>
<td>0.896</td>
</tr>
<tr>
<td>Bilingual</td>
<td>8.446</td>
<td>4.656E+03</td>
<td>1.374</td>
<td>0.020</td>
</tr>
<tr>
<td>Gender</td>
<td>-11.519</td>
<td>0.000</td>
<td>2.677</td>
<td>0.007</td>
</tr>
<tr>
<td>Bilingual*Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen. Status (3rd)</td>
<td>-5.557</td>
<td>0.004</td>
<td>1.332</td>
<td>-0.022</td>
</tr>
<tr>
<td>Gen. Status (2nd)</td>
<td>-5.474</td>
<td>0.004</td>
<td>2.325*</td>
<td>-0.027</td>
</tr>
<tr>
<td>3rd Gen. * Bil.</td>
<td>-7.431</td>
<td>0.001</td>
<td>4.618(NS)</td>
<td>-0.019</td>
</tr>
<tr>
<td>2nd Gen. * Bil.</td>
<td>32.592</td>
<td>1.427E+14</td>
<td>1.945</td>
<td>0.020</td>
</tr>
<tr>
<td>Parental Lang. Sp.</td>
<td>-8.918</td>
<td>0.000</td>
<td>1.825</td>
<td>-0.015</td>
</tr>
<tr>
<td>Parental Education</td>
<td>3.389</td>
<td>29.636</td>
<td>0.329</td>
<td>0.010</td>
</tr>
<tr>
<td>Inc. 100K &amp; above</td>
<td>8.799</td>
<td>6.628E+03</td>
<td>4.465*</td>
<td>0.057</td>
</tr>
<tr>
<td>Inc. 35K to 100K</td>
<td>15.440</td>
<td>5.076E+06</td>
<td>3.790</td>
<td>0.072</td>
</tr>
<tr>
<td>Inc. 10K to 35K</td>
<td>19.417</td>
<td>2.708E+08</td>
<td>3.895</td>
<td>0.119</td>
</tr>
<tr>
<td>Zip Codes</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

Note: *p < .05. **p < .01. p<.001. Coefficients for zip codes fixed effects have been suppressed for parsimony. In a binary probability model the exponential B indicates the odds of participating in post-secondary education. In both models the B represents the percentage that a person from a specific group is likely to participate in post-secondary education. Models are estimated using robust standard errors. Variable “Inc.” refers to parental income at base year.
Summary of Results

The previous analyses provide relevant information about how bilingual Latinx individuals compare in four indicators comprising education, job participation, income, and social capital (i.e., volunteer) activities. These results help us further understand this growing segment of the population.

Regression on an array of observable variables to predict education participation indicates that bilingual Latinx are 29.6 percentage points more likely to pursue any post-secondary education. Looking at gender interactions, I found that Latinx women, in general, are more likely to attend any post-secondary education. However, when accounting for bilingualism, there was a small negative effect of bilingualism for women.

Analyses for job market participation demonstrate that bilingual Latinx are 4.3 percentage points more likely to participate in the labor market. Moreover, in general, bilingual Latinx individuals can make up to $8,315 more annually than non-bilingual Latinx. For women, there seems to be a penalty in compensation in the labor market. However, this penalty was smaller if they were bilingual.

The results of the exploratory analysis on social capital outcomes indicate that bilinguals are about 2 percentage points less likely to participate in volunteer activities. In the next section, I discuss the implications of these results, limitations and future areas of research.
Chapter 5: Discussion

The purpose of this dissertation was to explore the educational, economic, and social outcomes of the Latinx community in the United States. Specifically, this dissertation focused on the differences in outcomes between Latinx bilingual individuals who speak Spanish at home and who also speak English fluently, versus the Latinx individuals who only speak English fluently. The goal was to expand the field of research on this growing segment of the population through the use of longitudinal data that followed individuals from their high school years into adulthood. I used a novel statistical methodology, such as coarsened exact matching that reduces biases on the analyses by matching individuals with similar characteristics a priori as opposed to the more commonly used method such as propensity score matching (Iacus, King & Porro, 2012). All models had fixed effects of zip codes to control for differences between locations. Results from the analyses indicate that bilingual Latinx are more likely to participate in any post-secondary education, more likely to participate in the labor market, and more likely to have higher salaries. These results were statistically significant, and they varied for some groups (e.g., gender, generational status, parental income). For example, there was a small negative interaction effect of bilingualism and gender on post-secondary education participation, indicating that bilingual Latinx women are slightly less likely to participate in post-secondary education. Also, there was a significant interaction effect between gender and bilingualism when predicting the likelihood of participating in the labor market, where Latinx women, in general, were less likely to participate in the labor market, but this effect was much smaller when accounting for bilingualism. These results indicate that bilingualism appears to be an advantage for women entering the labor market.
As for the analysis of volunteer engagement activities, the results were not positive, indicating that bilingual Latinx are less likely to engage in volunteer activities as defined in this dissertation. This information provides insight into how bilingual Latinx individuals decide to engage in various volunteer activities. Next, I explain these findings by research question and in the context of similar research that has been conducted in the last 20 years.

Question 1: For participation in post-secondary education, results indicated that overall, bilingual Latinx individuals are significantly more likely than monolingual Latinx to attend any post-secondary education program. Thus, these findings corroborate my theoretical framework and previous studies about education success, where bilingualism is beneficial to individuals when that skill is fostered. Moreover, these results corroborate the notion that bilingualism is an asset and a human capital that has an impact on educational outcomes because bilingual Latinx are more likely to attend post-secondary education, then they are also more likely to have a higher salary compared to non-bilinguals (Callahan & Gándara, 2014). Finally, this disproves the idea that bilingualism can be detrimental to individuals under the premise that it does not allow them to engage in educational activities. Another important finding related to education is that Latinx women were slightly less likely than men ($B=62.9$ percentage points for women, $B=65.9$ percentage points for men) to attend any post-secondary programs. These results were small, but significant indicating that there might be some barriers to accessing higher education for women.

Question 2: For the labor market participation, results indicated that overall, Latinx bilinguals are more likely than Latinx non-bilinguals to participate in the labor market as defined in the dataset as “ever having a job”. This finding suggests that being bilingual is an advantage for entering the labor market, meaning that being bilingual provides more opportunities for Latinx individuals to engage in the economy. This finding also supports the theoretical
framework that being bilingual is an asset that could increase the likelihood of being able to participate in the economy. Moreover, it supports previous findings regarding bilingualism being an asset that has value in the economy and that it is a form of human capital that provides benefits to individuals who develop such skill.

The interaction results show that Latinx women, in general, are less than men to participate in the labor market, but when considering bilingualism, this effect is minimized, suggesting that bilingual women have an advantage in the labor market that reduces the gap of participation between men and women. In other words, bilingual Latinx women are more likely to participate in the labor market than bilingual Latinx men. These findings corroborate Robison-Cimián’s (2014) findings which indicate that bilingual individuals are more likely to participate in the labor market, proving that bilingual has a value in the market that benefits individuals when controlling for location (i.e., fixed effects of zip codes).

Question 3: Results from the analyses to predict income showed that bilingual Latinx saw a benefit of about $8,314 more than monolingual Latinx. Interestingly, while overall, Latinx women made significantly less than men, when including bilingualism in the analysis, women made about $5000 more than men. These findings are consistent with my hypothesis that when bilingualism is perceived as an asset, it could bring economic benefit to the individuals. Until recently, researchers have not focused on how immigrant communities differ from one another. Specifically, researchers have not looked at how bilingualism affects the Latinx community. Recent studies on this growing demographic sector have found positive results in educational and labor market outcomes. However, these results have been only for a few economic sectors, such as the service industry and client-facing jobs. (Callahan & Gándara, 2014). Previous studies have found an overall benefit for bilingual immigrant communities (including the Latinx) of
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

about $2000 higher income differential for bilingual individuals. In this dissertation, the focus was on Latinx only, and this benefit was multiplied for bilingual Latinx individuals. These findings add to the previous idea that when a language different than English is considered an asset, it is economically advantageous for immigrant communities by increasing their chances to pursue higher education degrees, participate in the economy through jobs, and having higher pay (Agirdag, 2014a; Alarcón, Di Paolo, Heyman, & Morales, 2014). Moreover, bilingualism reduced the gender gap in the labor market participation and income variables, proving to be beneficial for women.

Question 4: Regarding the outcome variable for social capital (i.e., volunteer activities), I estimated the effects of bilingualism on the Latinx individual’s involvement in volunteer activities after the college years. There was a negative effect for bilinguals on the likelihood of engaging in volunteer activities as adults, indicating that bilingual Latinx are less likely to be engaged in volunteer activities as defined in the dataset. These results do not corroborate my theory that language can be a connector in community engagement activities and a source of social capital. Moreover, these results go against previous findings related to the development of social capital within communities that speak the same language. Before drawing generalizing conclusion, it would be critical to understand how the construct of volunteering is defined for the Latinx community, as the current definitions of volunteerism may not match that of the Latinx community. In other words, the way volunteer activities are captured in the survey administered to individuals may not be consistent with how Latinx define volunteering.
Implications and Future Research

In general, results from this study are positive for the bilingual Latinx community, representing an advantage for certain activities such as education participation, job market participation, and income differentials benefiting this segment of the population. Next, I explain how these findings can have an impact on future research areas.

Post-secondary Education. The general findings of post-secondary education participation were generally positive for the bilingual Latinx community. However, there were small differentials for bilingual Latinx women who had slightly lower rates of participation in higher education. It is important to highlight that the focus of these analyses was to understand rates of participation in higher education programs, independently of what type of program or completion rates. It would be beneficial to find what types of programs Latinx individuals are pursuing and whether they are completing these programs. Moreover, it would be fundamental to find whether bilingualism plays a role in their education or whether simply general education achievement is impacting the outcomes. Positive results would indicate that bilingualism increases the participation in higher education, which could lead to the extension of bilingual education beyond elementary school and the promotion of bilingualism as an asset, particularly considering that middle-class parents are now requesting bilingual programs and given that about 25% of Latinx individuals are currently in school could become emergent bilinguals (Migration Policy Institute, 2014).

Another important angle to look at is to try to understand why second and third generations of bilingual Latinx are doing worse than the first generation when it comes to attaining post-secondary education. Finally, future studies could look at disentangling these results further for Latinx individuals, to further understand specific mechanisms that are allowing
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

individual Latinx to engage in higher education and how to continue fostering those mechanisms through policies such as the development of effective bilingual programs that increase the proficiency of students in both languages. that goes beyond primary school.

**Labor market participation.** Overall, labor market participation was positive for bilingual individuals. Moreover, bilingualism proved to be an advantage for women in the labor market, reducing the gap of participation between men and women. This is an important finding that promotes the importance of bilingualism in the Latinx community as an asset that increases their opportunities in the labor market. Future studies could look at what types of jobs bilingual Latinx individuals are taking and whether these continue to be client-facing jobs or whether these individuals are taking managerial positions as well as whether bilinguals are working in the public or private sectors or self-employed. Moreover, given the spread of the Latinx population to different areas of the country, it will be important to know where these jobs are concentrated. Regarding gender gaps, it will be important to explore the reasons for the gap in rates of participation between men and women. Exploring social issues such as childcare, access to better health systems, income inequality, and support for women who are taking care of their parents or other family members could increase women participation in the labor market. Moreover, it will be important to understand what kinds of jobs Latinx women are taking and what are some of the responsibilities in these jobs.

**Income.** The results in this dissertation suggested positive income differentials for bilingual Latinx individuals. To my knowledge these are the largest economic numbers that apply to a large immigrant community. Future studies could focus on the reasons that bilingual Latinx individuals have for taking one job over another. Findings from these analyses could
uncover information about access to the job market for bilingual Latinx individuals (e.g., educational barriers, discrimination, market demands).

Moreover, it would be interesting to know if Latinx individuals are choosing careers that are not necessarily highly paid. Instead, it could be possible that bilingual Latinx are choosing careers with other benefits such as a better quality of life and more opportunities to engage with their community (e.g., social worker, teachers). Alternatively, there could be systemic mechanisms that impede bilingual Latinx individuals from pursuing certain career opportunities.

**Social Capital.** Contrary to the hypothesis presented in this dissertation, indicating that bilingual individuals would be more likely to engage in social capital activities, results showed that Latinx individuals were less likely to be engaged in volunteer activities as defined under the available dataset. Future studies could examine how bilingualism works as a connector in the Latinx community and what barriers of engagement the Latinx community faces when it comes to becoming involved in volunteer activities. Moreover, it will be important to look at how this large demographic segment defines volunteerism and see whether these definitions match the way volunteer is traditionally defined in this dataset. This analysis is important because there may be other volunteer opportunities that are not captured in the data used for this analysis. Furthermore, perhaps people are not claiming engagement in activities that may be perceived more of a duty than a volunteer activity such as caring for a family member or participating in the education of their children. A better definition of this construct from the available dataset and its meaning to the Latinx community could give us more accurate information about this issue. After a clear definition, future studies could examine whether social engagement moderates the educational and economic outcomes of individuals.
Language proficiency. In previous studies, measures about English language proficiency in either English or native language were often self-reported and not necessarily using an objective measure. In this dissertation, I used an objective measure of English proficiency, but there were no available data regarding Spanish language proficiency. It is clear that English proficiency is vital to participate in the economy of the United States, so another way to look at outcomes is to explore Latinx monolinguals with different levels of English proficiency compared to bilinguals with the same levels of English proficiency. For example, compare monolinguals who are fluent in English to bilinguals who are also fluent in English. Future studies should look at how this effect varies across different degrees to better understand the career paths that bilingual Latinx students follow and what opportunities or barriers affect the participation in higher education programs. This comparison will allow us to understand the implications for further post-secondary engagement in other degrees, such as bachelor’s degrees, masters, and doctoral programs.

Diversity within Latinx. There is evidence that the Latinx community is not a monolith and that groups within this racial demographic segment vary by educational background, economic opportunities, and access to the labor market and social services in the United States. It will be important to disentangle these groups and look at each group specifically to determine the policies that work best in different communities depending on the concentration of Latinx groups.
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

Limitations

The results of this dissertation are novel and add to the literature about the performance of the Latinx community in the United States in four specific areas, education, labor market, income, and social capital activities. However, two limitations arise regarding the generalizability of the results and the definition of bilingualism.

**Generalizability:** The results of this study are specific to the Latinx individuals who were part of the longitudinal database used in the analyses. Results inform how the community performed on four specific variables, and they represent a large subset of the Latinx community but may not be applicable in all Latinx living in the United States as this population is continuously changing. Moreover, these results may not apply to other ethnic bilingual groups such as the Chinese or the Arabic immigrant population.

**Defining bilingualism.** Different than previous studies, in this dissertation, I aimed to define bilingualism using objective measures of language instead of a self-reported question indicating the level of proficiency in either language. This goal was only half met as the dataset only offered objective measures of language for English language proficiency, but not for Spanish language proficiency. Therefore, the level of Spanish proficiency is still not clear for Latinx individuals. While previous studies have indicated that immigrant individuals who handle the native language of the host country see better outcomes, to truly examine the impact of bilingualism it will be necessary to have an objective measure of proficiency in both languages.

Another limitation is that in this dissertation I compared an individual who is Latinx whose native language is Spanish and are fluent in English, to (a) Latinx whose native language is Spanish and are not fluent in English, (b) Latinx whose native language is not Spanish who are fluent in English, and (c) Latinx whose native language is not Spanish and who are not fluent in
BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES

English. Thus, there is the possibility that findings were inflated given that some individuals were low in both. However, in a follow-up analysis I found that only 15% of individuals fit this category. In future research I intend to examine more closely the differences in outcomes for Latinx who are both fluent in English, but who speak either Spanish or NOT Spanish at home.

Conclusion

In this dissertation, I looked at the effects of bilingualism for Latinx individuals through the use of recently released longitudinal data and through the use of novel statistical techniques that allow us to make stronger conclusions about the educational attainment, labor market participation, income, and social capital engagement. The results of educational outcomes, labor market participation, and income were significantly positive, supporting my original hypothesis, indicating that bilingual Latinx individuals are more likely to participate in post-secondary education, more likely to participate in the labor, and have higher incomes. Moreover, bilingualism appears to close the gap between Latinx men and women, benefiting women in their likelihood of having a job and closing the income gap. This study provides important information that could be used by policymakers to create more effective bilingual programs and giving bilingual Latinx individuals better chances to participate in the economy. These programs should foster the development of English to allow the individuals to participate in the U.S. economy, but also foster an inclusive environment for the native language of the students to help them enjoy future educational, economic, and potentially social benefits, such as more career choices and more opportunities to engage with their communities. This effort could result in positive outcomes for the Latinx community and for the American society as a whole, and by making it more competitive on a global scale.
References


BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES


BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES


BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES


BILINGUALISM: VISUALIZING OUTCOMES FOR LATINX STUDENTS IN THE UNITED STATES


