Distribution Agreement

Signature:

Katie Lee Cartwright Date

Investigation of a Paradox of the Latino Paradox
Social Determinants of Health and the Diabetes Disparity

By
Katie Lee Cartwright
Doctor of Philosophy
Sociology

Ellen Idler, PhD
Advisor

Irene Browne, PhD
Committee Member

Solveig Cunningham, PhD
Committee Member

Corey L.M. Keyes, PhD
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Investigation of a Paradox of the Latino Paradox

Social Determinants of Health and the Diabetes Disparity

By

Katie Lee Cartwright

M.A., English, Lehigh University, 2005

B.A., English & Spanish, Lenoir-Rhyne University, 2003

Advisor:

Ellen Idler, PhD

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Sociology

2015

Abstract

Investigation of a Paradox of the Latino Paradox

Social Determinants of Health and the Diabetes Disparity

By

Katie Lee Cartwright

Latinos living in the US report a disproportionately high prevalence of diabetes. Simultaneously, Latinos experience unexpected health advantages, including longer life expectancy and lower rates of certain cancers and heart disease. This pattern is called the "Hispanic Paradox." The Latino diabetes disparity has not been considered in the frame of this paradox. When considering both epidemiological patterns, a question emerges: how can the same social determinants of health lead to such powerful health advantages for some health outcomes, while at the same time leading to a great health disparity in regard to diabetes?

This dissertation explores this question by examining the associations between social determinants of health and self-reported diabetes *within* the US Latino population and the associations of these social determinants of health in explaining the *difference* in self-reported diabetes prevalence between the US Latino population and the non-Latino population. This project uses the National Health Interview Survey (NHIS) data from 2006-2011, accessed via the Integrated Health Interview Series (IHIS) data managed by the Minnesota Population Center (MPC). Logistic regression analyses are used to examine the patterns within the Latino population and Oaxaca-Blinder decomposition is used to examine the patterns explaining the difference in self-reported diabetes between the Latino and non-Latino population.

The logistic regression analyses show that Latinos are almost 43% more likely to report being diagnosed with diabetes than non-Latinos. Individual characteristics of age, race, and smoking behaviors are identified as suppressors of the association between Latino identity and diabetes. Conversely, measures of social inequality, social ties, acculturation, and origin of Latino heritage are all potential mediators of the association between Latino identity and diabetes.

The Oaxaca-Blinder decomposition shows that individual characteristics (particularly age, race, BMI, and smoking habits), measures of social inequality, measures of social ties, measures of acculturation, and measures of Latino ethnic origins inform the explained difference in self-reported diabetes between the Latino and non-Latino population. Social inequality measures contribute a larger part of the explained difference than social ties measures or acculturation measures.

Investigation of a Paradox of the Latino Paradox

Social Determinants of Health and the Diabetes Disparity

By

Katie Lee Cartwright

M.A., English, Lehigh University, 2005

B.A., English & Spanish, Lenoir-Rhyne University, 2003

Advisor:

Ellen Idler, PhD

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Sociology

2015

ACKNOWLEDGEMENTS

As I conclude my time as a doctoral student at Emory University, I reflect with deep gratitude on all who have supported me, inspired me, challenged me, and loved me along the way. My support system is an embarrassment of riches.

First, I want to thank Ellen Idler, my dissertation chair and my academic champion. Writing a dissertation can be a lonely enterprise, but Ellen buoyed my spirits throughout the entire process. From tangible support of reading drafts, offering feedback, the generous gift of her time, and letters of support to intangible support of encouragement, empathy, and faith in my potential, Ellen could not have been a better advisor and mentor. I am grateful for her academic support, but I am also deeply thankful for her voracious appetite for reading. Her book recommendations helped me achieve work-life balance and have brought me great joy. She is my role model and a kindred-spirit.

I am also deeply grateful for other academic mentors at Emory. I am grateful to my committee, Irene Browne, Solveig Cunningham, and Corey Keyes. Irene has supported me from day one of my time here at Emory. She guided me to literature, course work, and research experiences that helped transform me from a social science neophyte to a Sociologist. Solveig Cunningham and Corey Keyes not only supported me through my dissertation work, but they both mentored me as a scholar, a researcher, and a writer. I am deeply grateful to both of them for their mentorship. Through my work with each of them, I came to understand writing as a collaborative process, which has transformed my approach to research and the production of knowledge.

I am grateful to have learned how to be a better teacher, a better scholar, and a better colleague from so many faculty here at Emory, including but not limited to John Boli, Roberto Franzosi, Monique Hennink, Karen Hegtvedt, Cathy Johnson, Michael McQuaide, Rick Rubinson, and Tracy Scott. The support and warmth of Kim Hall, Pat Hamilton, and Neema Oliver helped me survive the iron cage of the bureaucracy of my graduate education, and I will forever be in their debt.

In addition to the department of sociology, I am grateful that my program and Emory University foster cross-disciplinary intellectual discovery. My work has been immeasurably improved by opportunities with the Rollins School of Public Health, the intensive grant writing institute, and most recently through my Mellon Graduate Teaching Fellowship mentors, Brett Gadsden and Ben Reiss, and other fellows.

I am deeply grateful to the Laney Graduate School, the Robert W. Woodruff Foundation, and the Mellon Foundation for the financial support of my graduate education. I am committed to honor their investment in me and my work as I continue the work I developed at Emory. Through the support of the Mellon Foundation, the Laney Graduate School, and Clark Atlanta University, I was able to complete my dissertation this year while simultaneously growing as a teacher and a colleague. I am indebted to my colleagues at Clark Atlanta for their support this year, and I am particularly grateful to my Mellon mentor, Dr. Obie Clayton, and Mrs. Robin Stanley-Jones. I owe my successes at Clark Atlanta to you.

Students who I have had the opportunity to teach as an instructor and as a teaching assistant at Emory University, Rollins School of Public Health, Oglethorpe

University, and Clark Atlanta University have helped me come into my strengths as a teacher and as a sociologist. These experiences have fundamentally altered the way I approach my work and my life. Thank you.

I am also deeply grateful for the gift of friendship of my fellow graduate students. I am grateful to my cohort, who have been there since the beginning. To Deena Isom, whose journey has mirrored my own, step by step, I am thankful for you. To Urmy Shukla and Roselyn Thomas, whose friendship sustained me during the difficult years of coursework and whom I have missed dearly as they have taken their many talents to impressive new positions. To Xue Li, whose late-night companionship in the lab made completing the dissertation less daunting.

I am also thankful all of my other colleagues who have bolstered me along the way. To Frau Doktor Kronberg, you have been my role model since day one, and I could not be more honored to graduate with you this May. I am inspired by your work and grateful for your friendship. To Jess Grosholz and to Lesley Watson, I am so grateful for your feedback on early stages of my dissertation work and your continued friendship. I am delighted to see you come into your own and to have had you shine as lights at the end of the PhD tunnel. To Jared Bok, thank you for literally sitting alongside me for more hours than I want to remember as we both trudged along on our respective projects. Knowing you would be there gave me strength to get through the process. To Ashlee Bledsoe, your faith in me surpasses my faith in myself. Thank you for your unceasing support and encouragement. To FikreJesus Amahazion, I am honored to call you a friend. From your guest lectures, your inspiring work ethic, and your seemingly boundless energy, I am amazed by you. I eagerly anticipate reading about the global impact of your work. To Tressie McMillan Cottom, I am in awe of your work, your courage as a public scholar, and your abilities as a teacher. I am grateful for your friendship. To Marisela, Alison, Natalie, Liz, Jennifer, Trent, Bonnie, Melissa, Dustin and Maggie, you have each boosted my spirits during times of struggle along this process. Thank you for being you.

I would not have survived this process and thrived at growing into myself if it were not for the love and support of many. My family has both encouraged me in my pursuit of my goals and endured the many sacrifices that this process demands. I am thankful to my parents, who nurtured me and empowered me through great love and sacrifice to become the woman I am today. To Kia and to Kelly, I am eternally grateful that I get to go through life as your sister. I am your witness and you are mine. You know me in ways that no one else can. Your love and support mean the world to me. To Nate and to Jimmy, I am grateful to you for your love of my sisters and for your friendship. To Haley, Jaden, Leland and Liam, thank you for making me an Aunt. I love you. To Camille and Jim, to Peter, Jody, Hank and Nick, I am proud and grateful to be part of your family. Your support has encouraged me and our adventures have given me the motivation to complete many projects. I am grateful for the memories I have of reading for my preliminary exams on the porch next to Jim. He was a great reader, and he will always be missed.

I am thankful for my vast extended family, which is comprised of both biological family and friends who have become family. As a child who moves uprooted often, I am grateful to the constancy of my family and friends for grounding me, for knowing me, and for loving me. I owe much of who I am to both Luther and Helen Bernston and Jan

and Steve Sigmon. You took me in and loved me as your own. Thank you for your unwavering support of me in times of great success but also in times of great adversity. I love you.

The friends who have supported me along the way are the very best that the world has to offer. To Kim, my dearest childhood friend, I am astounded by all that we have accomplished in over two decades of friendship. I eagerly anticipate what will come in the next twenty years. To Katie, to Erica, to Pam. You inspire me, you challenge me, you love me, you forgive me, and our friendships continue to grow in spite of the challenges of distance and time. To the great network of other friends who have carried me through this degree, you have lifted my spirits, modeled love and support, and reminded me why I chose to work toward this goal. You are what is good in this world. Thank you.

To Mari, Jason and the boys, you have fed me, housed me, and nurtured my spirit. Your family's love and support means the world to me. To Vanya, this time in Atlanta would be barren without you. You are a blessing to me. To Greg and Emily (and the menagerie), the difficulty of the past two years has been obliterated by your friendship. You have given me literal and figurative sustenance that has nourished my body and soul.

I came to study the health and well-being of Latino immigrants through a wide range of influences and experiences. My life course was altered due to a trip to Bolivia when I was sixteen years old. This trip was made possible by many who supported me financially and spiritually, but I want to honor two specific women, Juanita Imogene "Gene" Cooper and Jean "Smith" Ohman. Gene underwrote the lion's share of the financial support I needed to go on the trip and Jean drove me from Steamboat Springs to Minneapolis where I began my journey. Their support and belief in me made it possible for me to travel to Bolivia where my philosophies on inequality and justice took root. These philosophies developed further thanks to my experiences as a Border Servant Corps volunteer with Las Americas Immigrant Advocacy Center. To the people of the border, you inspire me to work for justice, to love deeply, and to live humbly. To my fellow volunteers, Sam, Josie, John, Hilary, Sarah, Amy, and Renee, I carry you in my heart. To Pam, I already addressed you as my friend, but you are also a woman I want to make as proud as you make me every day. To Rand, to Paul, to Seth, to Scott, thank you for your belief in my academic potential. Thank you for challenging me, encouraging me, and for bringing so much joy to my life in the form of literature and friendship. To Rand, thank you for the gift of Seamus Heaney, whose words and life continue to inspire.

Finally, to David. I cannot articulate the extent to which I love you. You are my true love and my true partner. You embraced the challenge of a long-distance marriage for six years, and together we have prevailed. You have been my reader, my best editor, my statistics champion, my sounding board, and my companion throughout this process. However, your emotional support of me overshadows your intellectual support. Your love, your vulnerability, your loyalty, and your faith in me have helped me evolve as a whole person. Reaching the pinnacle of my PhD together is as significant as our summit of Mt. Shasta to me. What we have accomplished together dumbfounds me. While our marriage thus far has been a whirlwind of travel and transition, I look forward to this next stage of our lives where we will focus on planting roots. I am overwhelmed by our future and cannot wait to build a life together in Albuquerque. Thank you for being my beloved.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1: DIABETES & THE US LATINO POPULATION

Latinos have been lauded as a group that has far better health outcomes than their socio-economic characteristics would predict, and this phenomenon is known as the Hispanic or Latino paradox (Afable-Munsuz et al. 2013a; Gerst et al. 2010; Lariscy, Hummer and Hayward 2015). Investigations of his paradox, particularly as it upends the predictions related to social class have proposed many possible explanatory theories. The theories have mainly focused on factors related to the large immigrant composition of the Latino population and have noted selection bias ("healthy migrant theory" and "the salmon bias"), acculturation buffers, and stronger social ties (Abraido-Lanza, Chao and Florez 2005; Palloni and Arias 2004). However, the most recent meta-review of the literature finds that even after controlling for a wide range of factors, this phenomenon is left unexplained, and while immigrants experience the largest health advantages, native born Latinos experience benefits as well (Lariscy, Hummer and Hayward 2015). However, Latinos do not experience health advantages in all areas of health, and in some areas experience disproportionate health disadvantages. Left out of the most recent assessment of the paradox is the health condition of diabetes. Unlike the mortality and cancer advantages, diabetes prevalence is much higher in the Latino population.

This pattern begs the question, how can the same social determinants of health lead to such a powerful health advantage for some health outcomes, while at the same time lead to such a disparity? While scholars have made much of the paradox, recent studies are finding that while Latinos live longer, they do not necessarily live healthier lives (Gonzalez et al. 2009; Gonzalez, Haan and Hinton 2001; Markides, Salinas and

Sheffield 2009). Therefore, as the Latino immigrant population grows, the concerns of Latino immigrant health will become increasingly more important.

Over 25 million people (over 8% of the population) have diabetes mellitus (diabetes) in the United States, and 79 million people exhibit pre-diabetic symptoms (CDC 2011). Diabetes is the most common disorder of the endocrine system, and it occurs when blood sugar levels in the body consistently stay above normal. The endocrine system of individuals with type 1 diabetes are unable to produce insulin, where the endocrine systems of individuals with type 2 diabetes cease to respond to the effects of insulin (CDC 2012). In 2010, diabetes is the seventh leading cause of death overall in the United States, but the fifth leading cause of death for Latinos (CDC 2010). The physical ramifications of diabetes are great, and the conditions associated with advanced stage diabetes include vision impairment, peripheral neuropathy (weakness, numbness and pain often in the extremities of the feet and hands), amputations, decreased mobility, and end stage renal failure.

Diabetes is particularly debilitating in the population aged 65 and older, and 10.9 million people (over 26%) in this age group have diabetes (over 90% of these cases are type 2 diabetes mellitus) (CDC 2014). The numbers of older individuals with diabetes is projected to grow rapidly as diabetes is being diagnosed at higher rates in the younger populations. Diabetes does not affect racial and ethnic groups at the same rates: 11.8% of the Latino[1] population has been diagnosed with diabetes compared to 7.1% of non-

[1] There is no consensus about terminology for Hispanics/Latinos. While the paradox is almost exclusively called the Hispanic paradox, many articles which discuss it use the word Latino to describe members of the Hispanic/Latino ethnic group throughout the article. Individuals frequently self-identify with country of origin ethnic groups. I will primarily use the term Hispanic as it is the name of the paradox, but I will also

Hispanic whites. There is additional variation within the Latino population; as of 2011, 13.3% of Mexican Americans had diabetes (CDC 2011). The costs of diagnosed diabetes in the United States in 2007 were approximately $174 billion. If the additional costs of undiagnosed diabetes, pre-diabetic symptoms, and gestational diabetes are included, the total reaches $218 billion (CDC 2011). The financial burden of disease is matched by the individual burden of disease. Not only are the symptoms of diabetes burdensome, but they exacerbate comorbid conditions of an individual.

While there is no cure for diabetes, type 2 diabetes is a largely preventable condition. While diabetes should be one of the more preventable and manageable chronic conditions, statistics show that it is one of the most pressing chronic conditions in the US and those diagnosed have a very difficult time managing the disease (Weiler and Crist 2009). Diabetes is as burdensome to the diagnosed individual as it is to society. In addition to the struggle of the symptoms of the disease, individuals with diabetes are also faced with managing diabetes-related stigma. As many (clinicians and laypersons) believe that diabetes can be managed by the diagnosed individual, those with diabetes are often blamed for it (Tak-Ying Shiu, Yee-Mei Kwan and Wong 2003; Weiler and Crist 2009). However, there are many societal structures involved, as evidenced by the patterns of who has the disease and how the disease is treated in different groups. Due to the high cost, the debilitating effects, and the preventable nature of the vast majority of diagnoses of diabetes, this condition is particularly important to consider when examining conditions facing the rapidly aging population in the Unites States.

at times use the word Latino to describe the aggregate group. I will also use countries of origin and regions when appropriate.

The United States is getting older and more diverse. By 2030, 20% of Americans will be over the age of sixty-five (Olshansky et al. 2009; Peterson 1999). Chronic health conditions tarnish the "golden years" for many across demographic groups. However, chronic conditions affect different groups at different rates and the experiences of living with certain diseases vary based on specific demographic factors. In addition to the increasing demand to understand the range of factors that affect the health of older Americans, it is also important to note that as a group older Americans are rapidly diversifying. The largest minority population of older Americans will be Latino, comprising 20% of the population aged 65 years and older, the majority of the Latino population will be Mexican-origin (Villa et al. 2012). Currently, immigrants constitute 13% of the US population (Grieco and Trevelyan 2010) and are the fastest growing segment of the population.

Diabetes disproportionately affects Latinos of all ages, and the consequences of this disease are particularly worrisome for the older population (CDC 2011). Diabetes not only negatively affects an individual's health, but the disabilities associated with advanced diabetes affect all aspects of an individual's life, including independence and well-being. Diabetes is more lethal in Mexican and Mexican-American populations than among non-Hispanic whites (Hunt et al. 2011). Latinos comprise the largest and fastest growing segment of immigrants, with immigrants of Mexican origin accounting for over half of the current immigrant population (Ruggles S et al. 2010). Based on data from the California Health Survey, older individuals of Mexican origin have poorer health outcomes as well as lower levels of income and educational attainment in comparison with their native born non-Hispanic white counterparts (Villa et al. 2012). However, these studies do not reconcile how the combination of factors that lead to a health advantage for the Latino population in one sense leads to such a disparity in regard to diabetes. This study will empirically test the association of these social determinants of health with the Latino diabetes disparity and then contextualize them within the discussion of how some of these same factors may lead to a positive mortality advantage as well.

Research Questions

- What factors predict a higher risk of diabetes for US Latinos?
- How are three of the predominant theories of social determinants of health (social inequality, social ties, and acculturation) associated with the Latino diabetes prevalence?
- How do the social determinants of health (including the three predominant theories and individual characteristics) contribute to explaining the difference between the non-Latino and Latino diabetes prevalence?

Figures 1 depicts the overarching research question and the black box that will be addressed in this project. Figure 2 depicts the specific associations that will be tested in this study.

CHAPTER 2: LITERATURE REVIEW

The most current review of the Hispanic paradox stresses the need for thorough understanding of the health issues facing the US Latino population in order to be best prepared to address population health as the US demographics shift to include a much higher percentage of individuals who identify as Latino (Lariscy, Hummer and Hayward 2015). However, a consideration of diabetes is missing from the analysis of the Hispanic paradox. Among the most prevalent chronic conditions and leading causes of death, diabetes is a disproportionately pressing problem for the US Latino population in contrast to the non-Hispanic white population (CDC 2014). If it were not for the diabetes disparity, the Hispanic health advantage would be even greater. This chapter will review the current state of research on the Hispanic paradox, how the current evidence on the Latino diabetes disparity contrasts with this evidence, and what the extant literature surmises about how social determinants of health influence the diabetes prevalence in the US Latino population.

Hispanic Paradox

The diabetes disparity facing the US Latino population is concerning on its own. However, in contrast with the overall health advantages the Latino population experiences, the diabetes disparity becomes even more of a puzzle. The Hispanic paradox was identified in Markides and Coreil's (1986) article "The health of Hispanics in the Southwestern United States: an epidemiologic paradox" (Markides and Coreil 1986).

However, after twenty-five years of subsequent research, scholars have not come to a consensus about the root causes of the phenomenon nor the extent of the trend. The Hispanic paradox notes that as a group, Latinos have disproportionately lower levels of education and income, as well as relatively poorer access to health care (Angel and Thoits 1987; Palloni and Arias 2004; Rumbaut and Weeks 1996; Singh and Hiatt 2006; Zsembik and Fennell 2005). Based on these characteristics, the Latino population should have very poor health outcomes. However, in the United States, Latinos have better than expected health outcomes and lower rates of mortality than their socioeconomic status would predict; this phenomenon is referred to as the Hispanic paradox. The Hispanic paradox is broken down into a paradox within a paradox as the foreign born Latino population, which reports even lower levels of income and education than native born Latinos, report even better health outcomes than native born Latinos. Numerous studies have noted this trend and have tried to identify the causal mechanisms behind it (Angel and Thoits 1987; Angel et al. 2010; Finch and Vega 2003; Palloni and Arias 2004; Rumbaut and Weeks 1996; Singh and Hiatt 2006; Zsembik and Fennell 2005).

The most recent assessment of the paradox reviews the theories that have been offered as lacunas to the paradox: the healthy migrant theory, the acculturation buffer advantage, the salmon bias effect, and more (Lariscy, Hummer and Hayward 2015). This study offers the most comprehensive analysis to date on the mortality advantage, which assesses not only the health advantages of Latinos as a group, but also how the health advantages are distributed among different subgroups of US Latinos. The greatest health advantage is experienced by foreign-born Latinos, but there are distinct health advantages for native born Latinos as well, particularly in relation to health outcomes related to smoking behaviors, as Latinos are much less likely to smoke than the non-Hispanic white population. However, while this study provides much needed analysis of the paradox in relation to nativity, specific countries of origin, socioeconomic status, and a range of health conditions, this study does not include an in-depth discussion of acculturation nor does it include a discussion of how the disproportionately high prevalence of diabetes fits into the overarching patterns of health and illness facing the US Latino population.

Angel and Thoits (1987) highlight the Hispanic paradox in their article theorizing how culture affects the cognitive structure of illness (Angel and Thoits 1987). The article works to uncover how illness is socially constructed, and they do this by showing the universal role culture plays in mediating individuals' experiences of illness. In particular, they emphasize how culture differences may lead to differences in perceiving illness and reporting health statuses differently. They warn against using one standardized instrument for attempting to measure the prevalence of a given illness in a diverse population, and they push for cultural sensitivity and cultural competence on the part of epidemiologists. The example they use to support their argument is the Hispanic paradox. They look at the paradox of Mexican origin individuals in the United States, and explain part of the paradox as a mismatch in the diagnostic process. Angel and Thoits (1987) argue that Mexican Americans who are not assimilated are less likely to report depressive symptoms, not because they do not have symptoms of psychological distress, but because their culture of origin does not have a means of acknowledging mental illness. They do not accept that individuals at the lower levels of the socioeconomic strata actually have as positive of health outcomes as the initial research would lead us to conclude (Angel and Thoits 1987).

While their argument about cultural sensitivity is an important one,[2] the evidence suggests that certain groups of Latinos do experience real health and longevity advantages. Rumbaut and Weeks (1996) are able to use clinical measures (blood pressure, birth weight, length of gestation, APGAR scales, and similar measures) to investigate the Hispanic paradox within the comparison context of low SES groups and finds the paradox to be largely an immigrant effect (Rumbaut and Weeks 1996). After identifying unexpected favorable health outcomes in immigrants from Mexico and Southeast Asia, they try to uncover the mechanisms behind the pattern. This study uses the Comprehensive Perinatal Program in San Diego County. This dataset only looks at health outcomes of low-income pregnant women in the treatment program, so there are some limitations of this study due to selection bias. Of the 1,464 individuals in the dataset, 1,070 are foreign-born Mexican women. Not only is this sample not nationally representative, it is also not representative of the state of California or of San Diego County. While it is important to keep these limitations in mind, this study has some unique strengths in regard to health measures.

The strengths of this dataset are the individual level data gathered (over 500 independent variables per case, a wide array of biomedical and sociocultural determinants of pregnancy outcomes). The analysis of the data looked at the maternal risk factors that would best explain ethnic and nativity differences. They find that the predominantly foreign-born Asian and Latino populations had superior outcomes in comparison with

[2] There is an extensive literature on how culture affects the social construction of illness and wellness. Quah (2010) gives a good review of the importance of culture in shaping preventative health behavior, illness behavior and sick role behavior. Immigrant cultures are identified as key sources of culture that influence health.

native born white and African American populations. They also examined nativity directly and find that outcomes were generally better for immigrants than for natives. However, for immigrants, the outcomes were poorer for women who reported greater levels of acculturation. They find that native born groups were more likely to smoke, drink alcohol, and have poorer diets. Of the 71% reporting partner information 89% of Latina women's partners were employed compared to 56% of African American partners. Immigrants had the least amount of education, but the best perinatal health outcomes. However, Rumbaut and Weeks also find that women who reported the higher amounts of knowledge about pregnancy and infant care had better health outcomes. While immigrants had better diets, 95% of women in this sample did not have access to either enough food or enough nutrients during their pregnancies. Data suggest that immigrant women are "superior health achievers," and that the Hispanic paradox is more about an immigrant effect than a Latino effect. However, they also note that "Americanization" leads to negative effects and that women who smoke, drink alcohol, and eat nutrient-poor diets are more likely to have poor perinatal health outcomes. Immigrant women who adopt these behaviors lose the protective advantage associated with nativity, which suggests that culturally-related health behaviors are an important piece of the advantage.

Nearly twenty-five years after suggesting that the Hispanic paradox is at least in part a result of misidentification, Angel et al. (2010) acknowledge that the prevalence of the unexpected favorable health patterns, specifically in regard to mortality, are well-documented and enduring (Angel et al. 2010). This study does continue to argue that cultural factors are the primary cause for the paradox, but instead of culture masking symptoms, Mexican culture is protecting immigrants from negative health outcomes.

They support this argument by showing that the effect decreases with length of time in the United States and the age at which an individual migrates. This study uses the Hispanic Established Populations for the Epidemiologic Studies of the Elderly (H-EPESE), which includes 5 waves of data from 1993-2007, and finds that older Mexican immigrants experience different health outcomes based on time of migration. Those that migrate later in life have lower risk of mortality than those who migrate in childhood or midlife. This study controls for chronic conditions and functional capacity, so the findings support both the healthy migrant hypothesis and the hypothesis that this effect wears off with increased time in the United States. The longitudinal nature of this data makes a strong case that a major part of the Hispanic paradox is directly linked to the immigrant acculturation process. However, in this study, chronic conditions (including cancer and diabetes) are grouped together, which is masking the mechanisms behind the outcomes. Other evidence suggests that where US Latinos with Mexican heritage experience lower rates of cancer, they experience higher rates of diabetes.

Finch and Vega (2003) also note a cultural component to the better than expected health outcomes of Latino populations, specifically those of Mexican ancestry (Finch and Vega 2003). They use the Mexican American Prevalences and Services Survey (MAPSS) to look at health patterns in Mexican origin adults. They note that not only are immigrants consistently in disadvantaged socioeconomic positions, but also that the acculturation process of immigrants is stressful. In fact, the acculturation process is riddled with chronic stressors including discrimination, worry over legal status, and difficulty with language. However, they find that two culturally specific aspects are protective: strong family and friend networks and reliance on religious support mechanisms. Greater numbers of peers and family members in the United States and a higher reliance on religious support mechanisms decreased the likelihood of reporting fair/poor health. However, instrumental (concrete and direct forms of assistance) and religious support moderates the effect of discrimination on physical health. This study finds that physical health is negatively associated with acculturation stressors and positively associated with social support. Discrimination is associated with poorer physical health among those who report low levels of social support. The Hispanic paradox is partially explained by culturally protective factors such as high levels of social support, but both the stressors of immigrating and acculturating are associated with declines in health. It is important to note that this study looks only at Mexican origin adults living in California.

For a more nationally representative example, Singh and Hiatt's (2006) examination of cause-specific mortality in native born and foreign born populations in the United States over time provides additional evidence that the Hispanic paradox may be more about an immigrant paradox (Singh and Hiatt 2006). They examine how health outcomes have changed for immigrants between 1979 and 2003. They use the National Health Interview Survey (NHIS) to explore immigrant health patterns at a national level. They find that the immigrant effect has become more protective over time. Immigrants experience lower risk of mortality, but this effect is not comprehensive for all causes of death. Immigrants experience lower mortality from lung and oesophageal cancer, COPD, suicide, and HIV/AIDS, but higher mortality from stomach and liver cancer. They venture that the protective effect is largely due to the changing demographics of immigrants in the United States over this period of time, and that the large increase in the percentage of immigrants from Latin America has greatly contributed to the protective effect.

From the investigations of the Hispanic paradox, the evidence suggests that nativity, ethnicity of origin, degree of acculturation, health behaviors (particularly smoking), and social support all help to explain the paradoxical health advantage in regard to mortality and cancers. However, the question regarding the existence of the diabetes disparity remains—many of these studies failed to include diabetes in their analysis, grouped diabetes in with other health conditions, or did not discuss the implications of the diabetes prevalence in contrast with their other findings. The next section will review the related literature that covers the US Latino diabetes disparity.

Paradox within Paradox: Diabetes Disparity

Low-income Latino and immigrant groups have poorer health outcomes than native-born white Americans in regard diabetes (Chaufan, Davis and Constantino 2011). One major contributing factor to the management of diabetes is the inability to adhere to recommended dietary changes due to the food environment in which the low-income immigrant groups live. Qualitative studies exploring food environments find that access to healthy food is severely restricted by cost, availability, and access (Chaufan, Davis and Constantino 2011). Food environments are examined as a main contributing factor to all low-income individual diagnoses of diabetes. This raises the question of similarity and difference. Are there systematic differences in the neighborhoods and societal experiences of low income immigrants and native born low-income populations?

Some of the factors that are considered to be systematically different for immigrants are access to health care, health literacy, and experience within the health care system. Immigrants are less likely to regularly visit the doctor, even when they have access to the same health benefits (such as Medicare). Older Mexican Americans are far more likely to be cared for in their home or the home of family members than the larger population (Markides 2000). While there are quality of life benefits to this arrangement, there may also be health detriments. Understanding how family care relates to compliance and adherence of medical advice for chronic diseases is very important as there is more room for misunderstandings and failure to comply or adhere to medical advice (Markides 2000).

Studies have found that Latino immigrants, including those with a diagnosis of diabetes, have a lack of diabetes knowledge and low health literacy. Low health literacy compounded by barriers to adequate health care may be contributing to poorer health outcomes for Latinos and immigrants (Coffman, Norton and Beene 2012). Qualitative studies suggest that Latino immigrants and their families lack sufficient understanding of diabetes as a disease and insulin as a treatment strategy (Hu et al. 2012). This is particularly troubling because without proper use of insulin the disease will progress more rapidly and result in far worse outcome (Hu et al. 2012). Additional studies have shown that Latino immigrants are unaware of their risk of having diabetes even though they have multiple risk factors (Maty and Tippens 2011). Studies have found that immigrants with diabetes who require language interpreters to communicate with their medical professionals are unlikely to receive physician ordering of diabetes care measures and even more unlikely to complete the diabetes care measures given to them

(Gany et al. 2011). Language barriers compound the already high barriers to medical care for immigrants with diabetes (Gany et al. 2011).

Understanding the correlating factors and consequences of diabetes in older Latino populations is important because it informs policies and strategies to preventing and managing this disease that disproportionately affects the Latino community. Older individuals of Mexican origin are more likely to experience health disadvantages than their white peers (Villa et al. 2012). Not only is the rate of diabetes prevalence higher in Mexican and Mexican American populations compared to the non-Hispanic white population, it is also more lethal (Hunt et al. 2011). This pattern is identified using both US and Mexican data. This pattern suggests that it may not only be disparities in resources and access to health care in the United States that contribute to the health disparities experienced by Latino populations, particularly Mexican Americans (Hunt et al. 2011).

Advanced diabetes affects individuals' independence through diminishing mobility, largely related to peripheral neuropathy and consequential amputation of extremities, and deteriorating vision, which in turn affects the ability to self-care (Ailinger 1989; Ailinger and Causey 1993). Diabetes negatively affects longevity and it affects the quality of the last years of patients afflicted with this disease (Angel et al. 2010). Immigrants with diabetes report lower levels of satisfaction with the health care they receive (Abraido-Lanza et al. 2011). Immigrants who have both diabetes and depression experience more severe diabetes-related symptoms than those who did not have depression as well as more severe depression related symptoms compared to those without diabetes (Black 1999). While these types of studies capture important patterns, it is difficult to pinpoint directionality or causality in these situations. If people are experiencing more severe types of diabetes, they may be more vulnerable to experiencing depression. In order to prevent and manage diabetes and its burden on the US Latino population a better understanding of the social determinants of health must be conducted.

Social determinants of paradox within the paradox

This study aims to discover the social factors that lead to the existence of a diabetes health disparity in contrast with the other Latino health advantages. The literature investigating the US Latino diabetes disparity can be categorized into three main groups: studies investigating the relationship between socioeconomic status and diabetes, the relationship between social ties and diabetes, and the relationship between patterns of acculturation and diabetes.

Social Inequality and Latino Diabetes Disparity

The social determinant of health that is strongly associated with the prevalence of diabetes in the Latino population is socioeconomic status. Social class has been likened to a modern day “magic bullet” (House 2002). Those who are wealthier have better health outcomes than those who are poorer. The Hispanic paradox complicates this axiom. However, in regard to the US Latino diabetes disparity, the relationship between socioeconomic status and diabetes outcomes operates in the predicted manner: those who have a higher socioeconomic status are less likely to have diabetes. This is supported by a large body of research that identifies socioeconomic status as a key factor. Bridging theories of inequality and culture, immigrants (even those who are legal permanent residents or naturalized citizens) have more limited access to health care than native born residents (Ailinger and Causey 1993). A recent study using NHIS data finds that Latinos are much less likely to have any health spending and are more likely to pay out-of-pocket for what they do pay for than their white counterparts (Bustamante and Chen 2012). In addition, Latino immigrants have even lower health expenditures than their native-born Latino or white peers. Perhaps due in part to the lack of health care access, both foreign born and native born Latinos frequently underestimate their diabetes risk and do not seek care until their health has already deteriorated (Maty and Tippens 2011).

One example of how poor doctor/patient communication exacerbates the treatment and management of diabetes in Latino immigrant communities relates to pharmacological intervention methods. Insulin is a primary part of treating and managing diabetes, and noting differences in usage and familiarity with the drug may be important. Stigma and poor understanding of how the drug works are key factors in the relationship between ethnicity and insulin use. Most of the concerns regarding insulin in the Latino population stem from lack of understanding about the progression of the disease and inadequate information about how insulin works (Hu et al. 2012). For both native born and foreign born Latinos with diabetes, health literacy needs to be improved (Coffman, Norton and Beene 2012). This again is an example of how culture and socioeconomic status compound health inequalities for this population.

There are situations where culture is completely overshadowed by structural inequalities. In regard to the relationship between being Latino and diabetes, poverty is identified as a key contributing factor (Chaufan, Davis and Constantino 2011). In this study of northern California health patterns, researchers find that Latino immigrants and service providers in this neighborhood identify diabetes as the most pressing health concern, but that access to healthy food is a major barrier to successfully treating and managing this health concern. They stress that without addressing issues of poverty and inequality in Latino immigrant populations, treatment plans for diabetes for this population will fail (Chaufan, Davis and Constantino 2011).

Immaterial of perfect understanding for why the disparity in diabetes prevalence exists, the consequences of living with diabetes for the US Latino population are great. First, as was already mentioned above, not only is the prevalence of diabetes greater, but the disease is more lethal in Latino populations than in in the non-Hispanic white population (Hunt et al. 2011). However, this is not the only relevant consequence. The quality of life for US Latinos with diabetes is compromised for a variety of reasons, including comorbid physical and mental health conditions. It should be noted that some studies find that older Latino immigrant health outcomes are worse, but their overall life satisfaction tends to be higher (Gonzalez, Giarrusso and Takagi 2007). However, this does not appear to be the case for older Latino immigrants living with diabetes. One study found that older diabetic Mexican American immigrants who also have depression have much higher rates of a range of other comorbid conditions than their native born Mexican American peers (Black 1999). This study used the Hispanic Established Population for the Epidemiologic Study of the Elderly (H-EPESE). This study has many advantages, but one limitation is that this study does not allow for additional comparisons based on ethnicity.

Low-income Latino and immigrant groups have poorer health outcomes than native-born white Americans in regard diabetes (Chaufan, Davis and Constantino 2011). One major contributing factor to the management of diabetes is the inability to adhere to recommended dietary changes due to the food environment in which the low-income immigrant groups live. Qualitative studies exploring food environments find that access to healthy food is severely restricted by cost, availability, and access (Chaufan, Davis and Constantino 2011). For these reasons, food environments are considered an important contributing factor to diagnoses of diabetes in low-income populations. This raises the question of similarity and difference. Are there systematic differences in the neighborhoods and societal experiences of low income immigrants and native born low-income populations?

The primary proxies for socioeconomic status in the current literature on the diabetes disparity are education, income and access to health care. In general, higher levels of education are associated with lower rates of diabetes (Whitaker et al. 2014). Also, for US Latinos, higher levels of education are associated with better health outcomes, leading to the conclusion that if Latinos had higher levels of education and income, the Hispanic paradox would be even greater (Lariscy, Hummer and Hayward 2015). Studies suggest that education and income have an inverse relationship with diabetes prevalence regardless of race or ethnicity (Robbins et al. 2005). Studies also find an inverse relationship between socioeconomic status or position and health outcomes within the Latino population as well (Morales et al. 2002). Some of the factors that are considered to be systematically different for immigrants are access to health care, health literacy, and experience within the health care system. Immigrants are less likely to regularly visit the doctor, even when they have access to the same health benefits (such as Medicare).

There are some studies that suggest a divergent relationship between higher socioeconomic position and diabetes for Latinos. Dinwiddie et al. (2014) find that education influences key health outcomes (including diabetes) differently, and not always in the predicted directions (Dinwiddie, Zambrana and Garza 2014). Their study finds that for Mexican-origin men residing in the US, high education levels were associated with high waist circumference and for all included in the study higher levels of education were associated with higher odds of hypertension and high waist circumference. For US-born women of Mexican heritage, the odds of diabetes increased with higher levels of education. However, many key social determinants of health are not included in these models, so the findings of this study warrant more investigation.

Social Ties and Latino Diabetes Disparity

Where social inequality focuses on the stratification and structural level influences on health, scholarship on social ties as a social determinant of health focuses on the ways in which interpersonal relationships mediate health outcomes. Overall, individuals with stronger social ties experience better health outcomes (Berkman et al. 2000). The most proximal social ties are with spouses and immediate family. Foreign-born Latinos are more likely to be married than non-Latinos, while Latinos are less likely to be married than non-Hispanic white Americans. However, both native-born and foreign-born Latinos are less likely to be divorced than non-Hispanic white Americans.

Latinos are more likely to be married and also they are more likely to live in multi-generational households than other groups. Older Mexican Americans are far more likely to be cared for in their home or the home of family members than the larger population (Markides 2000). This suggests that their social ties may be stronger, which would lead to better overall health. However, there is some evidence that suggests that children living in multigenerational households may be at a higher risk for obesity, which would set these children up for a greater risk of diabetes later in life. Social ties would then possibly contribute to solving the puzzle of both the Hispanic paradox *and* the diabetes disparity. (Almeida et al. 2009; Angel et al. 2004; Berkman et al. 2000; Finch and Vega 2003; Kawachi and Berkman 2001; Seeman 1996; Thoits 2011; Viruell-Fuentes and Schulz 2009b)

Evidence suggests that family dynamics between Latino and non-Latino populations lead to different outcomes in the management of diabetes, which suggests that social ties theories based on a mainstream white American family and social system may not be as salient or yield expected results when applied to a Latino family context (Chesla et al. 2003). Other studies suggest that social support is not strongly related to diabetes self-management for Latino individuals (Gleeson-Kreig, Bernal and Woolley 2002). However, these studies did not explore overarching dynamics between social ties and *prevalence* of diabetes in the Latino populations.

Investigations of social support and health outcomes in the US Latino population find that Latinos tend to have large social networks and that those networks are primarily made up of family. However, having a large-family based network does not necessarily indicate that Latino individuals are wholly satisfied with their support networks

(Gleeson-Kreig, Bernal and Woolley 2002). In addition, quality of social ties may be influenced by acculturation forces. Evidence suggests that nativity and ethnicity affects the perceived quality of relationships in Latino families with foreign born generations and native born generations (Almeida et al. 2009). There is evidence to suggest that the family context and dynamics of Latino patients are statistically significantly different from non-Latino patients and that the associated health outcomes are also statistically significantly different (Chesla et al. 2003). Importantly, there is evidence to suggest that in regard to diabetes management, Latinos residing in the US who have high family cohesion may actually experience poorer health outcomes and practice more detrimental health behaviors (Chesla et al. 2003).

Marriage, one of the most powerful social ties that influences health, tends to lead to positive health outcomes, with men experiencing greater benefits than women. Latinos are more likely to be married than other groups, which should be associated with better health outcomes. However, there is also evidence that suggests that for some Latinas, marriage may not be a pathway to better health. Evidence suggests that Latina women who are living in a more gender-traditional household experience poorer health outcomes due to their gender role (Pineda Olvera et al. 2007). These gender role expectations limit self-care until women have aged out of primary caretaking status. These studies suggest that although the proportion of Latinos who are married is higher (and much higher for foreign-born Latinos) than the non-Latino population, there may not be a protective benefit for married women. Protective effects for married men should still be found.

Acculturation and Latino Diabetes Disparity

There has been extensive research about the contributing factors to diabetes. Some argue that the disparity in the prevalence of diabetes in Latino immigrant populations is largely cultural. There is an established relationship between culture and health. Culture influences health behaviors and beliefs, which in turn affects health outcomes (Quah 2010). Immigrants have multiple cultural influences: cultures of origin, cultures of destination countries, as well as the process of adapting from one culture to another. Most of the studies suggest that ascribing to a traditional Latino culture is associated with a lower risk of diabetes. Latinos who adopt more mainstream American lifestyle and health behaviors experience the greatest risk of diabetes. Recent Latino immigrants have low levels of diabetes knowledge, low health literacy in general, and experience barriers to health care (Coffman, Norton and Beene 2012). These patterns would suggest that immigrants would be at higher risk of diabetes. However, the evidence suggests that new immigrants are less likely to have diabetes, but with increased length of time in the US, immigrants' risk of diabetes increases, even after controlling for age (as individuals age, the risk of diabetes increases).

Evidence suggests that Latino immigrants who reside in the US experience worse health outcomes with greater degrees of acculturation. Foreign-born Latinos who have lived in the US longer have poorer health outcomes than foreign-born Latinos who have lived in the US for shorter periods of time (Hubert, Snider and Winkleby 2005; Kaplan et al. 2004). Frequently, this is measured by the length of time Latino immigrant has spent in the US—greater lengths of time in the US is used as a proxy for high degrees of acculturation. Consistently, it is found that with greater lengths of time in country, Latino immigrants experience worse health outcomes (Abraido-Lanza, Chao and Florez 2005). Latinos who were born in the US or Latino immigrants who have resided in the US for greater than 15 years are more likely to have a high BMI score than Latino immigrants who have resided in the US for less than 15 years (Abraido-Lanza, Chao and Florez 2005). However, there are multiple limitations of these studies. For example, Abraido-Lanza et al. (2005) group native-born Latinos with Latino immigrants who have resided in the US for greater than 15 years together. They also do not consider country or region of origin as a social determinant (Abraido-Lanza, Chao and Florez 2005).

As a result of these studies, the length of time immigrants have spent in the US is frequently used as the measure for acculturation. A study using NHIS found that length of time in country correlates with higher rates of obesity for Latino immigrants. The percentage of the group who are obese for immigrants who have lived in the United States for longer than 15 years is almost 15 percentage points higher than for immigrants who have lived in the United States for less than 4 years (Kaplan et al. 2004). Similarly immigrants who immigrate later in life arguably do not have as much time to assimilate and they are more set in their ways when they do. There is some evidence to support these theories. Immigrants who immigrate later in life tend to have much lower risk of death than individuals who migrate in childhood or midlife; this pattern holds even when controlling for chronic conditions and functional capacity (Angel et al. 2010). While there are a range of more specific instruments for measuring assimilation and acculturation, this rough measure is enough to capture many divergent patterns.

Studies consistently find that while immigrants are healthier than native-born residents of the US when they first arrive, this protective advantage dissipates with time (Abraido-Lanza, Chao and Florez 2005). Research that has investigated the cause of this trend has highlighted the acculturation process. This research suggests that immigrants enter the US with many positive health behaviors, but that after immigrating to the US, these healthy behaviors decline (Abraido-Lanza, Chao and Florez 2005). Some portions of this hypothesis are supported by evidence from the NHIS. There is evidence that suggests that Latinos are less likely to smoke and to drink alcohol than non-Latino whites. However, importantly for the diabetes disparity, there is evidence that Latinos are less likely to exercise and more likely to have a higher BMI. In regard to those who have been in the US longer, evidence suggests that Latinos who have been in the US for greater than15 years are more likely to drink, smoke and gain weight than Latinos who have been in the US for a shorter amount of time. However, Latinos who are more acculturated are *more* likely to have recently exercised, which suggests that not all health behaviors decline.

Scholars believe that dietary assimilation is of particularly relevance to the prevalence of diabetes in immigrant Latino populations. The shift from more “real food” based diets to the prepared and processed diet of mainstream Americans leads to the increase in the prevalence of diabetes. One study examining diabetes in immigrant populations finds that the immigrants who have more fully adapted to mainstream US dietary habits have worse health outcomes than immigrants who maintain dietary and nutrition habits of their culture of origin (Akresh 2007). This study uses the New Immigrant Survey (NIS), which has well developed instruments to measure assimilation and acculturation. However, this dataset, and studies that implement this data are only able to look at differences within immigrant populations. In order to examine differences between foreign-born and native-born groups, which allows more inferences regarding the mechanisms, studies need to use data which have been gathered using a sampling strategy that includes a nationally representative sample (such as the NHIS). This allows for a broader understanding of how age, ethnicity, nativity, and other social determinants of health affect health outcomes of US Latinos.

However, other studies push back on the idea that acculturation or assimilation is the cause of the Latino diabetes disparity. In a study of individuals with Mexican heritage in the Sacramento area, researchers found that native born individuals with Mexican heritage were almost twice as likely to have type 2 diabetes as foreign born individuals originally from Mexico (Afable-Munsuz et al. 2013b). While in initial models there is support for the acculturation hypothesis, once the researchers controlled for socioeconomic status and other lifestyle factors, there was no statistically significant relationship between the acculturation measures and diabetes. Also, other studies consider US born Latinos as fully acculturated (Abraido-Lanza, Chao and Florez 2005). This subgroup has the highest rates of diabetes, but as a group never experienced an "acculturation" process. This would suggest that other social determinants of health are key to explaining the diabetes disadvantage.

Studies have shown that Latino immigrants are unaware of their risk of having diabetes even though they have multiple risk factors (Coffman, Norton and Beene 2012; Maty and Tippens 2011). Studies have found that immigrants with diabetes who require language interpreters to communicate with their medical professionals are unlikely to receive physician ordering of diabetes care measures and even more unlikely to complete the diabetes care measures given to them (Gany et al. 2011). Language barriers compound the already high barriers to medical care for immigrants with diabetes (Gany et al. 2011).

A major contributing factor to the treatment of Latino immigrants with diabetes is language discordance (when patients and providers are unable to communicate directly with each other) (Gany et al. 2011; Kim et al. 2011a; Kim et al. 2011b). When language discordance exists, overall rates of physician ordering diabetes care and patient compliance and adherence to the physician's orders tend to be very low (Gany et al. 2011). This trend only exacerbates the health conditions suffered by this population. However, while studies find that language proficiency affects Latino individuals' ability to *manage* diabetes, it does not explain the prevalence of diabetes.

Latino ethnicity of origin

When exploring health disparities facing the Latino immigrant community, researchers must also consider the diversity of Latino immigrants. Some factors of particular importance are birthplace, length of time in the United States, and language skills in addition to the socio-demographic factors that sociologists usually control for in medical sociology studies. Research on immigrants and acculturation finds that length of time in country correlates with poorer health outcomes and worse dietary habits. The greater the changes in dietary habits are associated with higher body mass index numbers (Akresh 2007).

One of the main critiques of studies examining Latino health disparities is the aggregation of the group into one category, when Latinos are heterogeneous in many dimensions (Rotermann 2011). One primary dimension of heterogeneity is the country or culture of origin of individual Latino individuals. The differences between Latinos who have Mexican heritage and Latinos who have Cuban heritage are notable. Studies show that disease prevalence varies, access to resources varies, and health behaviors vary between ethnic subgroups (Ai et al. 2012; Ai et al. 2013; Huang et al. 2011; Oza-Frank and Venkat Narayan 2008). What these studies show is that the Hispanic paradox may not be named particularly well. When studies look at the aggregate Latino trends, they note the immigrant effect. Many studies only look at the Mexican origin populations when studying the Hispanic paradox, which offers support for a Mexican effect, but not a holistic Latino effect. Zsembik and Fennell (2005) tease this issue apart using the NHIS data (Zsembik and Fennell 2005). The findings of this article suggest that the use of “Hispanic” or “Latino” in the Hispanic paradox is a misspecification of a social determinant of health. Instead of “Latino” or “Hispanic,” countries of birth or cultures of Latino heritage should be used. This article finds that the health patterns are clearly different between different Latino ethnic groups and countries of origin. Their overall health findings suggest that Mexicans have health advantages and Puerto Ricans experience health disparities. Cubans and Dominicans have more varied health outcomes, depending on which health outcome is studied. This study finds that worse health is associated with higher levels of SES and acculturation among Mexicans, but worse health is associated with lower levels of SES and acculturation among Caribbean-origin Latinos. Zembrik and Fennell show that researchers should be careful not to generalize about immigrant populations either (Zsembik and Fennell 2005). Immigrant heterogeneity leads to different health outcomes as well.[3]

Scholars need to do more work to clarify which Latino populations experience unexpected favorable health outcomes relative to their SES, both in regard to nativity and country and/or region of origin. Also, the name of the paradox may need to be altered in order to better match the patterns. However, what is not in question is that certain groups, even if it is a Mexican immigrant population, are experiencing health outcomes that outperform their socioeconomic class and resources.

There is some evidence to suggest a cautious approach when hypothesizing or generalizing results based on acculturation as a social determinant of health. Afable-Munsuz et al. (2013) find that after controlling for socioeconomic and lifestyle factors, the association between acculturation and diabetes risk (measured using the Acculturation Rating Scale for Mexican Americans) for Latinos of Mexican heritage disappears, but there is a lingering effect of generational status and diabetes that is not clearly linked to acculturation (Afable-Munsuz et al. 2013b). The strengths of this study are the measure of diabetes (combined measures of diabetic medication use, self-reported physician diagnosis, and clinical fasting glucose levels) and the measure of acculturation (the Acculturation Rating Scale). The limitations of this study are the sample strategy which focuses exclusively on Mexican Americans residing in the Sacramento area. However,

[3] This finding is also supported by studies of native born black Americans and black immigrant subgroups and health outcomes (Read, Emerson and Tarlov 2005). Also using NHIS data, they find that while black immigrants have better health outcomes than native born black groups, immigrants from different regions experience different health outcomes.

the findings that indicate a more thoughtful approach to the study of acculturation are well-argued.

Individual Characteristics

Other individual characteristics that emerge as important covariates in the literature include a range of achieved and ascribed statuses.

Age. An individual's risk of developing diabetes increases with age (Koopman et al. 2005). In the past few decades, the age of onset of type 2 diabetes mellitus has decreased (Koopman et al. 2005). As the Latino population is younger on average than the non-Latino population in the US, it is especially important that all models investigating the disparity control for age.

Gender/Sex. Dinwiddie et al. (2014) note that health outcomes and behaviors for Mexican heritage Latinos residing in the US are not only statistically significantly different from each other, but there is evidence to suggest that associated social determinants of health influence these outcomes differently by gender (Dinwiddie, Zambrana and Garza 2014). Recent studies investigating Latino men's and women's health find that there are key differences in patterns of health and health care by gender in the Latino population (Ai et al. 2012; Ai et al. 2013). Differences in health patterns are also found in immigrant populations (Meadows, Thurston and Melton 2001). All of these findings suggest that gender must be considered when investigating the Latino diabetes disparity.

Race. Ultimately, the literature on both the Hispanic paradox and the Latino diabetes disparity leads to the conclusion that class, while a powerful and important predictor of health outcomes, does not explain all health disparities in the United States. In order to produce better health outcomes for all, the US needs to reform more than the class system, economic inequalities, and access to health care, but the US also needs to address other forms of inequalities that lead to disparities. There is convincing evidence that there are meaningful differences in health outcomes, including diabetes, among all ethnic and racial groups residing in the US (Wang and Beydoun 2007; Williams and Sternthal 2010). Related specifically to obesity, racial and ethnic subgroups, both foreign-born and native-born, are at different risks for obesity and diabetes (Cunningham, Ruben and Venkat Narayan 2008; Cunningham, Vaquera and Long 2012).

Race and ethnicity have distinct socially constructed effects on health outcomes. The evidence gathered on black immigrants and non-immigrant groups and the ethnic diversity of Latinos in regard to region or country of origin make a strong case for the social-construction of race and ethnic variance in health outcomes. Stressors that are specifically race and ethnicity based are discrimination and prejudice which function independently of class inequality. While the Hispanic paradox is an exception to the strong inverse relationship between SES and rates of mortality, the possible causal mechanisms do not undermine the larger trend. Similarly to the historical trajectory of other examinations of race, ethnicity and health, the history of the study of race, ethnicity, and diabetes is also fraught with bad science (Tuchman 2011). When generalizations are made about the associations between race and ethnicity and diabetes, limitations of methods must be noted, and the informing theories and the existing evidence should be considered in light of these limitations.

BMI. A variable that needs to be included in all analyses if available is a measure of individual weight statuses. While occasionally studies will include a dichotomous variable capturing overweight status, far more common is the inclusion of individual's body mass index scores (BMI). BMI is a score calculated from an individual's weight and height and is used in a wide range of health studies. As there is an established association between higher levels of BMI and elevated risk of diabetes, BMI is included in studies to control for a known covariate of diabetes. Evidence consistently indicates that Latinos residing in the US have higher average BMI scores than non-Latino whites, even after controlling for a range of social determinants of health (Abraido-Lanza, Chao and Florez 2005). Latino women from Mexican, Puerto Rican and Cuban heritage residing in the United States all have high prevalence of BMI scores that fall into the categories of overweight and obese (Ai et al. 2012; Oza-Frank and Cunningham 2010). High levels of acculturation are linked to higher BMI scores (Hubert, Snider and Winkleby 2005). There is also evidence that suggests that socioeconomic factors may explain more of the difference in BMI (Latinos having higher mean BMI scores than non-Hispanic whites) for Latinos and non-Hispanic whites than between other minority groups and non-Hispanic whites (Powell et al. 2012).

Smoking. There are established relationships between smoking and diabetes, even after controlling for socioeconomic status. There are multiple studies that suggest that there is no strong tie between smoking and diabetes after controlling for socioeconomic variables. However, other studies do find that there is an elevated risk of diabetes for smokers, even after controlling for a wide range of determinants of health (Willi et al. 2007). Diabetics who smoke face more diabetes-related complications (especially related to circulation problems) than diabetics who do not smoke (Sherman 2005). Although smoking is less common in Mexican American populations, Mexican Americans who do smoke experience more severe consequences (higher smoking related cause specific mortality) (Wei et al. 1996). Smoking has emerged as the key health behavior influencing the Hispanic paradox, as a much smaller proportion of Latinos smoke than non-Latino whites (Lariscy, Hummer and Hayward 2015). However, the diabetes rate is much higher, so further investigation of this pattern is warranted.

The review of the literature produced three main theories to investigate as social determinants of health that may be driving the prevalence of diabetes in the Latino population, and in turn the diabetes disparity: the theories of social inequality, social ties and acculturation. The next chapter will discuss these theories more broadly and set up the model for the empirical analysis of this project.

CHAPTER 3: THEORETICAL FRAMEWORK

The review of the literature has highlighted three main theoretical explanations for why the diabetes disparity between Latino and non-Latino population exists: socioeconomic inequality, social ties, and acculturation. Figure 1 identifies the overarching research question and the empirical “black box” this study investigates: why is being Latino associated with a higher likelihood of reporting being diagnosed with diabetes? Figure 2 sketches out the three theoretical frameworks that emerge from the literature as social determinants related to Latino health outcomes, including diabetes. This chapter further explores these three theories and offers hypotheses based on these theories that inform the analyses conducted in this study.

Theory of Social Inequality

Social class stratification has a well-established association with health outcomes. This relationship is remarkably consistent and is supported by an extensive body of evidence. Social class inequality is linked to a wide range of health conditions and is so persistent and consistent across a wide range of different social groups and social measures that is considered a fundamental social determinant of health (Berkman and Syme 1979; Link et al. 1998; Ross and Mirowsky 2010; Wilkinson 1999). Socioeconomic factors are also consistently linked as mediators in the association between race and ethnicity and health (Kingston and Smith 1997).

A wide body of literature supports the hypothesis that social class and health status are linked. The lower an individual's position in the class system the poorer their health outcomes are (House 2002). If the United States decreased social class inequality, population-wide health disparities would decrease as well. However, reforming social inequality may not be the panacea for health inequalities that some studies would suggest. While the evidence strongly supports the claim that social inequality is a fundamental cause of health disparities, the evidence also suggests that other factors, such as race and ethnicity are also fundamental causes of health outcomes. While social class is a powerful predictor, the reality of the Hispanic paradox suggests that class is not the only fundamental cause of health disparities.

Population level health patterns are inextricably linked to socioeconomic gradients. As both behaviors and risk factors have emerged as primary factors for diseases such as diabetes, social factors (specifically socioeconomic status) are shown to affect behaviors and risk factors indirectly affecting health and are shown to have a direct effect on health outcomes as well (Phelan et al. 2004). Some of the mechanisms through which social class is thought to have an effect is through education, income, access to health care, safer jobs, healthier neighborhoods (lower risk of exposure to toxins, easily accessible fresh produce, and lower levels of crime and strain) (House 2002; Phelan et al. 2004; Sampson, Morenoff and Gannon-Rowley 2002). As an aggregate group, Latinos are at a disadvantage for most of these measures. These mechanisms and society's reproduction of inequality qualify social class as a fundamental cause of mortality and other health outcomes (Link and Phelan 1995; Phelan et al. 2004).

Phelan et al. (2004) test the theory of social inequality as a fundamental cause of mortality and find that mortality from causes of death that are preventable are substantially more strongly related to socioeconomic status than mortality from less preventable causes. This theory aligns with findings of current studies of diabetes. These findings support the theory that social class is a fundamental cause of mortality as access to greater resources is associated with lower mortality rates, where lower status is associated with higher mortality rates, even though modern science has the tools to extend their longevity. Other studies note that access to health care is not the key factor in socioeconomic health disparities. In the UK, the National Health Service was supposed to counteract social class differences in regard to health outcomes. However, while access has been made universal, health disparities have actually grown between the classes (House 2002). The explanations for the mechanisms driving these disparities return to how social class affects individuals' exposures to risk factors for both mental and physical health outcomes. House also notes that race is a fundamental cause of health. The literature review of Latino health and social class indicate that the lower social position does in fact affect health, and although Latinos as a group experience better than predicted health outcomes in many areas, their health outcomes would be even more favorable if their socioeconomic position equaled the non-Latino white population.

Social Ties Theory

Social ties influence an individual's health directly both by ramifications of isolation and connectivity. The dynamics of an individual's social network may influence an individual's health more indirectly through their health behaviors. Social isolation, the complete absence of social support, has been linked to negative health outcomes (Cacioppo and Hawkley 2003; Seeman 1996). Social support, socially cohesive networks and social contacts are largely associated with positive health outcomes (Berkman et al. 2004). Social ties may buffer individuals from stress (Bobak et al. 1998). Social support can be helpful for individuals who are managing chronic conditions (Gallant 2003).

Individuals who are the most integrated have better outcomes than individuals who are socially isolated. Social isolation leads to more sedentary lifestyle (MacDougall et al. 1997). Some studies show that the socially isolated are more at risk of being overweight for some groups (Lemeshow et al. 2008). However, other studies show that older individuals who are socially isolated are at risk for not eating enough (Donini, Savina and Cannella 2003).

There are other studies that also indicate that there are mixed health outcomes depending on quality of social network. The perception of the quality of support is associated with health outcomes. Individuals with diabetes who *perceive* that they have positive support have better success in managing their condition, while individuals with diabetes who perceive that they have negative support fair poorer (Tang et al. 2008). This indicates that just having social ties does necessarily predict better health outcomes. Building on this divergent theory of social ties, social networks are often homogenous when it comes to health behaviors related to obesity, so individuals may have networks that buffer stress and lead to other positive health outcomes that at the same time put them at greater risk for diabetes (Christakis and Fowler 2007).

Little of the social ties literature specifically investigates the association of Latino ethnicity, social ties, and health outcomes. In the literature review in Chapter 2, there was evidence that Latinos are more likely to be married than non-Latinos, but there was also evidence that marriage may be associated with more negative health outcomes for Latina women due to more rigid gender role expectations (Chesla et al. 2003). Another key factor in immigrant health outcomes is family involvement. Family integration is more important and more prevalent in immigrant households than native-born households. Older Mexican Americans are far less likely to live in nursing homes and are far more likely to live with family members than non-Hispanic white Americans (Angel et al. 2004). The number of residents in the household captures an important family dynamic that varies between immigrant households and non-immigrant households: multigenerational living arrangements. Latin American immigrants are far more likely to live in multigenerational households than the general population. In fact, 18.8% of people living in immigrant households live in multigenerational households, where 14.2% of people who are native born live in multigenerational households (Staff 2010). However, there is evidence that these strong family connections are not consistently associated with positive health outcomes (Chesla et al. 2003; Pineda Olvera et al. 2007).

Acculturation Theory

One of the reasons that Latinos have such a different family structure profile is thought to be due to differences in acculturation[4]. As Latinos have a much larger

[4] While studies of assimilation and acculturation may focus on different aspects of immigrants' adaptations to the culture of the destination country, in many cases the concepts are used interchangeably and measured identically. For this literature review, I will refer to the term specifically used in each study; for this project, I will use the word acculturation to refer to this process.

immigrant composition compared to non-Latinos, their immigrant cultures are a highlighted as a key source of difference that could be leading to differences in health outcomes. In part because socioeconomic status of Latinos fails to predict mortality and other health outcomes in the way that the evidence would suggest, studies of Latino health that solely focus on social inequality will miss crucial factors that influence health outcomes. However, in the case of diabetes, Latinos are at a disadvantage, so the relative role of acculturation in the prevalence of diabetes may be very different than the association with acculturation and mortality.

A key model that will help fill in these gaps is the acculturation and health model. Acculturation is, most generally, the process through which immigrants adapt to the culture of the destination country. However, this definition would be considered unidimensional (Lara et al. 2005). Better definitions of acculturation are bi-dimensional or multidimensional. Multidimensional models of acculturation do not assume a linear trajectory of adaptation and allow for a range of adaptation pathways into a new culture, which include assimilation, separation, integration, and marginalization (Lara et al. 2005; Thomson and Hoffman-Goetz 2009).

However, when considering the role of assimilation in the process of acculturation, there should be an acknowledgment that many studies use acculturation and assimilation as synonymous. Assimilation is also depicted as having both unidimensional and multidimensional interpretations. A unidimensional approach to assimilation is to define it as “a natural process by which diverse ethnic groups come to share a common culture and to gain equal access to the opportunity structure of society, this process consists of gradually deserting old cultural and behavior patterns in favor of new ones, and once this process begins, it moves inevitably and irreversibly toward assimilation" (Zhou 1997). This general definition of assimilation has been challenged by many researchers. While the idea of assimilation as a process where immigrants adapt to a new culture is accepted, the idea that the process is a universally linear process is not. This model also assumes that there is one mainstream culture to which one adapts. While there are some elements of American culture that are dominant, depending on region, racial or ethnic group, and a variety of other factors, theorists argue that America lacks one unified cultural identity. Segmented assimilation makes room for these different cultural identities that immigrants may encounter as they adapt to life in the United States (Portes and Rumbaut 2006; Zhou 1997). These theories set up different pathways and processes for assimilation and acknowledge that immigrants may experience very different processes of acculturation.

When considering acculturation and assimilation in a multidimensional frame, it creates a space for acculturation or assimilation to lead to a wide range of outcomes. According to a unidimensional definition of acculturation, individuals may experience better health outcomes because they gain socioeconomic status as they acculturate. However, they also may experience worse health outcomes as they adopt more sedentary and unhealthy lifestyles (Lara et al. 2005). This model faces the same limitations as the general assimilation model. The general assimilation model largely assumes that factors such as length of time in country, language ability, formal labor market success, and the adoption of American lifestyle habits lead to positive health outcomes. While some of these assumptions are valid, the process is more complicated than this model allows. Segmented assimilation theory indicates that there a range of different processes that immigrants experience, and also there is not one uniform culture to which immigrants adapt. Instead, the process is murky and the cultural standards are often muddled.

Upward assimilation does predict better health outcomes, but even this is more complicated than it appears. High levels of upward assimilation correlate with better health outcomes most of the time, but there is some evidence that suggests that high levels of assimilation correlate with worse health outcomes in a few specific contexts. Upward assimilation theories suggest that immigrants who have achieved upward assimilation, are more seamlessly incorporated into the US culture and have better health outcomes (Gonzalez et al. 2009). Social networks and support are offered as the main mechanisms for promoting health in these theories. According to this theory, immigrants who are more highly acculturated are more likely to be incorporated into their community, through churches, employment, and social organizations. These immigrants are more likely to understand what resources are available and how to access them, therefore, resulting in better health outcomes. However, assimilation theories suggest that while immigrants may adapt to the aspects of the mainstream culture associated with success (e.g. higher education, higher income, and mastery of destination language), this "achievement" comes at a cost due to "acculturation stress" (Mui and Kang 2006; Rogler, Cortes and Malgady 1991). The process of learning and adapting to a new culture is so taxing that it has a negative effect on mental and physical health.

Downward assimilation, on the other hand, is often a catch-all for all of the negative parts of American society. For example, downward assimilation would entail adopting mainstream American dietary habits, including eating fast food or processed food, ceasing to prepare meals at home, and increasing portion sizes. This frequently happens to Latino immigrants from a wide range of socioeconomic status. A common result is increasing levels of both obesity and diabetes in immigrants (Oza-Frank and Venkat Narayan 2010). If an immigrant comes from a culture where he or she is preparing most of their food at home with limited food choices, the change to the processed food culture of the United States will be detrimental to their health.

Downward assimilation also includes not prioritizing education, participating in illicit or informal labor markets, building social networks with negative role models, and engaging in criminal activities. With the exception of the adoption of some negative health behaviors, most first generation immigrants do not downward assimilate. In fact, it is frequently second generation immigrants, children of immigrants, who experience downward assimilation. While it may seem as though the upward/downward assimilation distinction is solely about socioeconomic status, this assumption is complicated by the fact that depending on region of the country, structural level policies regarding immigration, and the percentage Latino in the place of residence, immigrants tend to assimilate in upward and downward patterns that do not simply parallel socioeconomic status.

The other main process adopted by first generation immigrants is to resist assimilation all together. These immigrants live in enclaves, do not learn English, and sometimes participate in the formal labor market, but frequently participate in informal economies. Also, these immigrants are the least likely to use structural level resources and are unlikely to become full citizens, even though they are eligible. As a group, Mexican immigrants have one of the highest rates of resistance to assimilation, and thus, are some of the least likely to achieve upward assimilation (Vigdor 2008). Those who experience a protective health effect from resisting assimilation are often benefiting by protective traits from their culture of origin (Viruell-Fuentes and Schulz 2009a). Protective culture theories suggest that the source of positive health behaviors come from healthy habits from the culture of origin, which are different than the mainstream habits of the United States (or the culture of the destination country). For example, they tend to maintain dietary patterns of their home country. While this food may be simple and not as nutritionally diverse as the "best" of American diets, their diet tends to be made-from-scratch food with appropriate portions and no bad fillers such as preservatives or excess fat and sugars. This example shows how resisting immigration may lead to better health outcomes.

The main drawback of remaining separate or on the margins is that these immigrants are the most likely to have low socioeconomic status and are the least likely to be aware of and utilize health based resources. They are the least likely to go to the doctor regularly, and they are the most likely to wait too long before seeking medical care from a physician (Oakie 2007). Immigrants are more integrated with native-born individuals are more likely to know about health care resources and to seek them out, regardless of financial resources.

There are many cautions about using acculturation as a variable in health research (Hunt, Schneider and Comer 2004). Most importantly, the concept is difficult to measure in a multidimensional way in most quantitative research and the concept risks essentializing cultures of origin. In the same way that the American culture is diverse, so are the cultures of the sending countries. Measurement in studies of immigrant health patterns in comparison to native-born health patterns raises many questions. There are debates about what are the most appropriate and accurate acculturation measures, which are all dependent upon how researchers define the mainstream culture (Gelfand and Yee 1991; Yamada, Valle and Barrio 2006). However, perhaps even more contentious are the measures used to compare immigrant groups with non-immigrant groups. For instance, there is some debate as to whether mental health problems such as depression in both groups can be measured with the same scale and if there is parity in what the threshold level represents in both groups (Gerst et al. 2010). Ultimately most of these studies note the fact that the measurement tool may not be able to accurately identify patterns in both populations, but also note that in order to do a comparative study the established tools are the best option. However, most of these scales are not included in nationally representative studies.

The benefit of the NHIS data I propose to use is the facility with which I am able to make population level comparisons. However, as this study emphasized general health measures, it does not include refined acculturation measures. The best proxies for acculturation are language preference, length of time in country and citizenship status (dichotomous variable: US citizen or not). While these variables are limited, other studies using NHIS data have implemented them in order to assess effects of assimilation (Bustamante and Chen 2012; Kaplan et al. 2004).

Hypotheses

This project will test the theories of social inequality, social ties, and acculturation to gather evidence to determine the association between the social determinants of health that influence the Latino diabetes disparity. Again, Figure 2 depicts how the literature and theoretical frames suggest what the associations may be. From the review of the literature and discussion of the theories, I establish the following formal hypotheses that I will test in the empirical portion of this study.

Association between Latino identity and diabetes:

H_1: Latinos will have a higher prevalence of diabetes than non-Latinos.

Social inequality theory:

H_2: Social inequality will be associated with higher levels of diabetes in Latinos.

Social ties theory:

H_3: Higher levels of social ties will be associated with lower levels of diabetes in Latinos.

Acculturation theory:

H_4: Higher levels of acculturation will be associated with higher levels of diabetes in Latinos.

These hypotheses are based on the findings from the current literature. They will be tested formally through multivariate analysis. The empirical analytical work of this project will examine the association of these concepts with self-reported diabetes through logistic regression analysis. Then, in order to further investigate the role of these theories in explaining the diabetes disparity, I will analyze the difference in self-reported diabetes between the Latino and non-Latino populations through Oaxaca-Blinder decomposition.

CHAPTER 4: DATA AND METHODS

Data

This investigation into the Latino diabetes disparity uses the Integrated Health Interview Series (IHIS) data, which is the harmonized data and documentation of the National Health Interview Surveys (NHIS) managed by the Minnesota Population Center at the University of Minnesota (Center 2012). These data are used to investigate the diabetes health disparities of the US Latino population. This dataset is the best available dataset to answer my research questions. The National Health Interview Survey (NHIS) is one of the longest running and comprehensive sources of information on the health of the civilian community dwelling population of the United States. It is a major data collection program of the National Center for Health Statistics (NCHS) under the Centers for Disease Control and Prevention (CDC) (2011). Data have been gathered annually since 1957, although the survey questionnaire has been adapted every ten to fifteen years. The goal of the survey has always been to monitor trends in illness and health in the United States. Since 1997, the year the last major overhaul of the questionnaire was implemented, the NHIS has included an extensive range of demographic and socioeconomic characteristics, which allow for thorough analyses of social determinants of health, as well as age, ethnicity and nativity. The IHIS version of these data are publicly available, which is made possible by recoding some variables to prevent identification.

The nationally representative sample design, with strategic oversampling of select minority groups, of the NHIS makes this dataset a well-suited dataset with which to establish and analyze the general patterns and national level trends related to the population of Latinos living in the US. While some datasets focus solely on the population of specific age groups (H-EPESE), the immigrant population (NIS), or ethnicity, this dataset allows for comparative data analyses between groups, such as non-Latinos and Latinos or native born and foreign born residents. The bulk of literature on health outcomes of US Latinos focuses primarily on the dependent health variable, and the literature often does not allow for wider generalizations. One key reason for this is that the very nature of the datasets makes this kind of analysis difficult or impossible. Frequently, when the dataset contains all of the key statuses, the sample sizes of the key variables of interest are too small for valid statistical analysis, limiting the level of detail possible. There are also a series of datasets (including HEPESE, NIS, and NSHEP—National Survey of Hispanic Elderly People)) that limit comparisons by only sampling from within groups of interest. In these cases, groups such as native born residents, anyone outside of the Latino ethnic group, or younger adults are not included. The studies that stem from these data offer important, crucial insight into how social determinants of health influence health outcomes *within* the populations, but are limited from conducting cross-group analysis.

However, the NHIS allows for robust cross-group analysis. Each year, the sample includes somewhere between 50,000 to 85,000 individuals. A common strategy for data analysis is to pool years of cross-sectional data for analysis in order to statistically analyze groups that comprise relatively small proportions of the US population. I follow the widely accepted strategy of pooling years of data that results in sufficient individual level data to study the age, ethnicity, and nativity independent effects and the intersections of these statuses and how they relate to the key dependent variables of my study (Lariscy, Hummer and Hayward 2015; Whitaker et al. 2014). This study uses 2006-2011 compiled data to run all analyses. Some key variables, such as nativity and country of origin have been compressed in this dataset to ensure de-identification.

While there are many advantages to these data, one limitation is that the data are the result of a cross-sectional design, despite their apparent over-time structure. Thus, while this study identifies many associations between variables and offers support for hypotheses of future longitudinal studies, direct causal conclusions cannot be made from this study.

Concepts and Variables

In order to thoroughly investigate the association between being Latino and having a higher likelihood of diabetes, I conduct a series of univariate, bivariate, and multivariate analyses based on the empirical framework identified in Figure 2. Figure 3 includes the variables through which the three theoretical frameworks tested in this study are operationalized. A detailed discussion of how these variables are measured is offered below.

Dependent variable

Diabetes. There are a series of questions on diabetes that measure the self-reported diagnosis of diabetes, the self-reported diagnosis of pre-diabetes, and the length of time since an individual was diagnosed with diabetes. The NHIS does not differentiate between type 1 and type 2 diabetes, but based on trends in prevalence of diabetes in adults, it is reasonable to make the assumption that between 90-95% of adults who self-report diabetes have type 2 diabetes (CDC 2011; Oza-Frank and Venkat Narayan 2010). The evidence suggests that for the Latino population the proportion of individuals with type 2 diabetes is closer to the upper limit. The broadest question asked by the NHIS asks, "[Other than during pregnancy,] Have you EVER been told by a doctor or other health professional that you have diabetes or sugar diabetes?" Respondents' answers are coded: Yes (1), No (2), Borderline (3), Refused to answer (7), and Does not know (9). The variable name is DIBEV. This variable is recoded into a dichotomous variable where (1) captures all those who responded that, yes, they have been diagnosed with diabetes, and (0) captures everyone else. As the dependent variable is dichotomous, logistic regression analyses are used to examine patterns associated with this outcome. This is a conservative recoding scheme such that even more of the sample (borderline, refused, and don't know) may have diabetes than have reported having diabetes. This means the potential bias works against my hypotheses and analyses rather than for them.

Independent variables

Latino ethnicity. The accepted methodological practices used when researching racial and ethnic minority health disparities frequently essentialize the minority groups, which may result in perpetuating or exacerbating these inequalities (Gómez and López 2013; Knight, Roosa and Umaña-Taylor 2009). This limitation of quantitative methods and techniques is particularly disheartening considering that many of these projects aim to ameliorate minority health. Arguably, the structure of qualitative research allows this type of research to respond and evolve more rapidly to the critiques of racial reification. Qualitative research allows the participants to self-identify and to explain in their own words what their race or ethnicity means to them.

Quantitative research relies upon the fixed-response racial and ethnic categories. While this is useful for understanding the patterns of inequality across large groups, there are also conceptual limitations to this that most quantitative researchers do not acknowledge in the development of their methods, their analysis, nor in their discussion of limitations.

Critical race scholars have challenged researchers employing quantitative methods to be very intentional in how they measure race and ethnicity, to articulate these decisions in their research, and to exercise great caution in making generalizations and making conclusions based on their results (Knight, Roosa and Umaña-Taylor 2009; Zuberi and Bonilla-Silva 2008a). While these challenges should be implemented in all research focused on race and ethnicity, it is particularly important to heed these cautions while conducting health research. Even with overwhelming evidence that race is not biological, but a social construction, researchers continue to treat racial and ethnic identification as a biological characteristic in their research by failing to justify the use of and unpack the variables carefully (Bonilla-Silva and Zuberi 2008; Gómez and López 2013).

This study begins by measuring Latino identity with self-reported identification based off of fixed-response category. In order to prevent essentializing a “Latino effect,” a range of variables are added to the models to account for heterogeneity of the Latino population, including ethnicity of Latino origin, nativity status, and racial identification. Specific considerations of each of these variables are expanded on in this section. To this end, I use the Latino ethnicity questions from the NHIS. First, individuals are asked the question: “Do you consider yourself to be Hispanic or Latino?” Individuals are able to respond: Yes (1), No (2), refuse to answer (7), or does not know (9). The variable name is HISPYN. Only individuals who responded are included in the analysis, and the resulting variable is dichotomous.

Origin of Latino Ethnicity. In order to critically examine both methods and the social issue of Latino health disparities simultaneously. Knight, Roosa, and Umaña-Taylor challenge all who study Latino health and well-being to investigate who comprises the social group “Latino” in any study, and especially quantitative studies, where often sweeping generalizations about Latinos are made (Knight, Roosa and Umaña-Taylor 2009). By analyzing the descriptive level statistics more critically, I will be able to have a more nuanced discussion of the findings of the regression analyses.

One way in which the present study does this is to carefully consider the intragroup composition. Carefully unpacking the components of the US Latino is of particular importance as so much of the research proposes that cultural factors are the key determinants to Latino health disparities. When researchers posit that the pathways to inequality are culturally specific, there is then an onus on the researcher to establish a justification for what they mean by culture. This justification must be theoretically and empirically supported. Too frequently a “Latino culture” is discussed, but the study sample is solely comprised of Mexican Americans or Puerto Ricans.

Therefore, as much information about specific ethnicities of Latino origin as is available is included. If an individual answers “yes” to the HISPYN question, a follow-up question is asked. While exploring trends using Latino ethnicity identifies important patterns, it is extremely important to remember that the Hispanic population, regardless of nativity, is an incredibly diverse group (Borrell, Menendez and Joseph 2011; Rotermann 2011). One of the key sources of this heterogeneity is the ethnicity of an individual’s Latino origin. Once patterns have been identified, it is important to find strategies that allow for their study at the subgroup level. This survey documents the ethnicities claimed by individuals who identify as Latino. Once an individual identifies as Hispanic or Latino, the individual is then asked the question: “Please give me the number of the group that represents your Hispanic origin or ancestry.” This variable name is HISPETH. Responses are coded as: Not Hispanic/Spanish origin (10), Mexican (all Mexican heritage categories recoded as one category) (20), Puerto Rican (30), Cuban/Cuban American (40), Dominican (50), Other Hispanic (all non-specified types including multiple Hispanic recoded as one category) (60), and Central or South American (61). This variable is used as a covariate for the first Latino question in all of the models that include Latino ethnicity.

Individual Characteristics

Age. Age is a particularly important variable to include as this study investigates diabetes. The literature establishes that there is a direct positive relationship between age and being diagnosed with diabetes. As an individual ages, the likelihood of being diagnosed with diabetes increases. In the NHIS all are asked the question “What is your age?” The interview subject responds with a number. The variable name is AGEDOB_1. This variable is treated as a continous variable. As age does not influence health outcomes in a linear fashion, I also calculate a variable for age squared. This accounts for the fact that an individual’s age in years does not linearly affect being diagnosed with diabetes, since the risk increases more rapidly at advanced ages.

Sex. Sex has established relationships with the independent variables (immigrant variables, the gendered experience of ethnicity), the dependent variable (diabetes), and other covariates (health behaviors, other health outcomes, income, education, and more). Sex is an important covariate for biological and sociological factors in this study. Biologically, there are differences in the epidemiology of diabetes and sex. For example, older immigrant women experience depression and much higher rates of diabetes than older immigrant men (Black, Markides and Miller 1998). Immigrants are less likely than native born individuals to interact with the health care system; however, similar to native born patterns, immigrant women are more likely to interact with the health care system than men, even controlling for pregnancy (Read and Reynolds 2012). The NHIS measures the sex of the respondent in two ways. First, in the NHIS interview process, the interviewers have the option to assess the sex of the individual themselves.[5] If not discernable, they ask the question, “Are you male (1) or female (2)?” The variable name is SEX. Due to how this variable is ascertained, there are no missing data. The omitted category is Male throughout the analyses. There are some limitations to this approach as

[5] In the summer of 2013, I was randomly selected as a respondent for the NHIS. The individual who surveyed me was startled at my level of enthusiasm for participating and my level of familiarity with the survey questions. I was not asked my sex.

it is not entirely self-reported. By putting some onus on the individual administering the survey (which is conducted by phone), there is room for misidentification. Arguably, this misidentification is small, and as the sample size for this study is large, the likelihood that any misspecification is unduly swaying the results is also small. However, this method, while more efficient, is not the “gold standard” of assessing sex and/or gender identity in social science practice.

Race. Race is assessed separately from Hispanic/Latino ethnicity. The majority of medical research, including medical sociological research, conflate race and ethnicity into one category or recode Latinos into a “white” Latino population in order to contrast Latinos with non-Hispanic whites. However, in doing so, Latinos with other racial identifications are erased from the story. More inclusive methods which allow for racial diversity of Latinos is an act of methodological justice (Stanfield 2011; Zuberi and Bonilla-Silva 2008a). Therefore, as I examine the possible contributing factors of health disparities in the Latino population, I need to investigate *both* the differences between racial and ethnic groups *and* the differences among the Latino subgroups. The question is a self-identified race category with fixed-response options: “What race or races do you consider yourself? Please select 1 or more of these categories.” The variable name is RACE. Individuals are able to choose from: White (1), Black/African American (2), Indian (American) (3), Alaska Native (4), Native Hawaiian (5), Guamanian (6), Samoan (7), Other Pacific Islander (8), Asian Indian (9), Chinese (10), Filipino (11), Japanese (12), Korean (13), Vietnamese (14), Other Asian (15), Some other race (16), Refused (97), Don’t Know (99). In the IHIS data, these are grouped as White (1), African

American (?), Asian (?), Native American (?), and Other Race (?). This variable is treated as a categorical variable in the analyses and the omitted category is White.

Some studies investigating Latino health and well-being construct a combined race ethnicity category, in order to highlight the differences between Latinos and non-Hispanic white populations specifically. In this study, I use the dichotomous Latino ethnicity category as the primary grouping and control for race in order to highlight the role of race within both categorizations. This also allows Latinos of different racial identifications to be included in the models. While these benefits are valuable within the logistic regression models, the inclusion of race in this way is particularly illuminating in the use of the Oaxaca-Blinder decomposition method (Blinder 1973; Oaxaca 1973). The Oaxaca-Blinder decomposition “decomposes” a difference in an outcome variable by group into explained and unexplained differences by constructing the counterfactual equations (where the intercept and coefficient of the equation associated with the Latino sample are replaced with those of the non-Latino sample) (O'Donnell et al. 2008).

BMI. Body Mass Index scores are constructed by the IHIS using self-reported height and weight. The survey question assessing height asks: “How tall are you without shoes?” The survey question assessing weight asks: “How much do you weigh without shoes?” From the responses to these questions, the BMI scores are calculated. BMI calculations from 2006 forward are comparable. The variable name is BMI. This variable is treated as continuous.

Smoking. There is a set of questions assessing smoking behaviors included in the IHIS. For this study, I constructed my smoking variable from the variable SMOKESTATUS2. This variable is a recoded variable of the respondent’s current smoking status. The participants' current smoking statuses are categorized into seven categories: not in universe (for individuals under the age of 18) (00), current every day smoker (11), current some day smoker (12), former smoker (20), never smoked (30), has smoked, current smoking status unknown (40), and unknown if ever smoked (90). I recoded this variable into four categories: Current every day and some-day smokers are grouped as current smokers (1), former smokers (2), never smoked (3), and both of the unknown smoking groups are categorized together as unknown (4). This variable is treated as a categorical variable and current smoker is the omitted category. I omit current smoker in order to see the difference in diabetes outcomes between current smokers and former smokers, as well as between current smokers and non-smokers.

Acculturation variables

The benefit of the NHIS data is the ability to make population level comparisons. However, as this study emphasized general health measures, it does not include refined acculturation measures. The best proxies for acculturation are language preference, length of time in country and citizenship status (dichotomous variable: US citizen or not). While these variables are limited, other studies using NHIS data have implemented them in order to assess effects of assimilation (Bustamante and Chen 2012; Kaplan et al. 2004).

Nativity. As over one-third of the US Latino population is foreign-born and theory and evidence suggest that cultural factors influence the diabetes disparity, this study investigates the role of nativity, or place of birth. The National Health Interview Survey measures this with the question: "Were you born in the United States?" The variable name is USBORN. Individuals select from four fixed response categories: yes (1), no (2), don't know (9), or refuse to answer (7). In initial analyses I use a dichotomous variable to establish associations. However, due to multicollinearity issues with length of time in country, I only use length of time in country (with native born as the omitted variable) in the multivariate analyses presented in this dissertation.

Theory and evidence also indicate that the effects of foreign born nativity vary by the length of time an individual has spent in the US. I use the number of years an individual has spent in the US as the measure for this. This variable (YRSINUS) is an NHIS constructed variable. First, the NHIS asks individuals who have been flagged as foreign born respondents, "In what year did you come to the United States to stay?" The answers are any year from 1880 to the current year of the study. The answers are coded by the full four number year and then missing data are coded as refused to respond (9997), and the responded does not know (9999). Then this response and the variable of date of interview are used to calculate a separate variable for length of time in country. This variable is calculated by subtracting the year an individual came to the US from the year of the survey. Then responses are then categorized: Not in Universe (native born individuals) (0), Less than 1 year (1), 1 year to less than 5 years (2), 5 years to less than 10 years (3), 10 years to less than 15 years (4), 15 years or more (5), and unknown (9). I use this variable as is. This variable is treated categorically, and being native born is the omitted category.

As there is perfect collinearity between the USBORN and YRSINUS variables, I use the USBORN to test the reliability of the YRSINUS variable, and then I use the

YRSINUS variable to capture both nativity and length of time in country in the regression models and in the Oaxaca-Blinder decomposition.

Language. As discussed in the review of literature, language (particularly speaking a language other than English) is thought to have an effect on health outcomes. However, the findings in the literature are mixed. There is some evidence to suggest that speaking a language other than English has a negative effect on access to health care and quality of health care, which negatively influences health outcomes. However, there is also evidence to suggest that the less acculturated a Latino individual is, the better his or her health outcomes are. Speaking primarily in a language other than English is a common indicator that an individual is less acculturated to mainstream American culture, which should have a positive influence on health. However, both sets of findings suggest that it is important to include this variable for consideration in models where the sample includes a meaningful number of immigrants. The NHIS is given in both English and Spanish, and the data records the language in which the survey was taken. While this piece of data does not record English proficiency, it does capture one aspect of language confidence or preference. The language variable (INTERVLANG) is coded English (1), Spanish (2), English and Spanish (3), Other language (4), and Unknown (8). There are no missing data for this variable. This variable is treated categorically, and English is the omitted category.

Social Inequality variables

Education. Individuals are asked, "What is the HIGHEST level of school you have completed or the highest degree you have received?" Individuals respond with Never attended/kindergarten only (00), 1st grade (1), 2nd grade (2), 3rd grade (3), 4th grade (4), 5th grade (5), 6th grade (6), 7th grade (7), 8th grade (8), 9th grade (9), 10th grade (10), 11th grade (11), 12th grade, no diploma (12), GED or equivalent (13), High School Graduate (14), Some college, no degree (15), Associate degree: occupational, technical, or vocational program (16), Associate degree: academic program (17), Bachelor's degree (18), Master's degree (19), Professional school degree (20), Doctoral degree (21), Child under 5 years old (96), Refused (97), Don't know (99). I recoded this variable into more salient categories for this study: Less than High School (1), High School Diploma (2), GED or equivalent (3), Some College (4), Associate degrees (5), Bachelor's degree (6), Graduate degrees (7), Unknown (8). The variable name is EDUC. While many studies further conflate these categories, as the sample is large enough, I am keeping these categories distinct in order to further investigate how education operates and may operate differently by Latino ethnicity. This variable is treated categorically and the less than high school level of education is the omitted category.

Income. The measure used in this study for income is household income in relation to the federal poverty line. This is a generated variable that uses the information provided by the question (variable name FINCTOT) that asks, "What is your best estimate of your total income/income of all family members in residence from all sources, before taxes in the last calendar year?" The answer is any number between $0 and $999,994 (0-999,994) those who made over $999,995 (999,995), refused to answer (999,997), and does not know (999,999). This variable, the number of individuals in the household, and the federal level of poverty for the year of the survey are used to calculate a poverty level measure in the IHIS. The poverty level variable is coded as less than 100% of the federal poverty line (1), between 100% and 199% of the federal poverty line (2), between 200% and 399% of the federal poverty line (3), above 400% of the federal poverty line (4), and unknown (9). The generated variable name is POVERTY2. The advantages of using this variable is that this variable captures the effect of income based on the relative difference between groups of individuals. This variable is categorical, and having an income less than the federal poverty line is the omitted category.

Health care coverage. A range of other health care variables have been shown to be intervening factors in health care outcomes. One of the most powerful correlates is the type of health care coverage an individual has. One's access to health care is correlated with one's socioeconomic status as well. The NHIS assesses this information with the question in various ways. For this study, I use a broad proxy which is a dichotomous variable indicating if an individual has health coverage. The question is "Are you covered by any kind of health insurance or some other kind of health care plan." The responses are coded as Yes (1), No (2), Refused (7), and Don't Know (9). I recode this variable as Not Covered (0), Covered (1), and Unknown (9). This variable is treated as categorical and having health coverage is the omitted category.

Social Ties

Marital Status. The first proxy for social ties is marital status. Individuals were asked, “What is your current legal marital status?” The answers are coded as: Married—spouse present (11), Married--spouse not in household (12), Widowed (20), Divorced (30), Separated (40), Never married (50), and Unknown marital status (99). This variable is a categorical variable and married with a present spouse is the omitted category in the regression analyses.

Number in Household. The second proxy for social ties is the number of individuals living in the household with the respondent. This is a technical household variable that is constructed by the IHIS by compiling the reports of the number of person records that are included with each sampled unit. This is a count variable. The variable name is NUMPREC. While there is a specific family size variable (FAMSIZE), which is also a constructed variable (which is generated by the National Center for Health Statistics), the IHIS codebook recommends against using this variable to capture household dynamics and instead use the NUMPREC variable. This variable is treated as a continuous variable.

Controls

Region of country. I control for region of residence variability. The region of residence of each participant was added during the processing stage as opposed to being asked directly. The survey design was based on household information, so each region of residence was known to the data processors. The regions are categorized into four groups:

Northeast (1), North Central/Midwest (2), South (3), West (4), and Unknown (9). This variable name is REGION. Due to the way this information is gathered, there are no unknown regions of residence for participants in the 2006 through the 2011 surveys. This variable is treated as a categorical variable and living in the Northeast is the omitted category.

Year of survey. In order to control for social dynamics specific to each sample, survey year is included, as is consistent with other studies using the NHIS (Langellier et al. 2014; Whitaker et al. 2014). This variable is constructed based on the year that each individual participated. This variable is treated as a categorical variable and individuals who were surveyed in 2006 are the omitted category for the logistic regression analyses.

Data, Data Access & Analyses

Summary statistics & bivariate analyses

STATA (version 13) is used to run the statistical analyses testing each of my hypotheses. Both the complex survey design and survey weights are applied in all analyses. I use the sampling weights for pooled data by using the STRATA and PSU supplied by IHIS. All years of data included in this project are from the same sample design period, which means that they are not treated as statistically independent to prevent underestimating standard errors (Center 2012). After applying survey weights, the “svy” command is used with all analyses (univariate and multivariate) to produce estimates while still incorporating the full sample and survey design information for variance estimation.

First, summary statistics are presented (see Table 1). I also test the difference between the Latino and non-Latino means and proportions for each of the variables included in the models. Due to the complex survey design and survey weights, Student's *t*-tests and chi-square tests are not appropriate (or possible). Instead adjusted for survey Wald test and the design-based Pearson *f*-test are used to test the mean and proportional differences between the Latino and non-Latino populations and diabetes, as well as all of the variables in the analyses (Jann 2008; Koch, Freeman and Freeman 1975; Rao and Scott 1981). All of the variables are statistically significantly different from each other.

Logistic Regression models.

Three sets of logistic regression models are specified to examine the relationship between being Latino and having diabetes (see Tables 2 through 5). In the first set, the logistic regression models use the full sample (Table 2). In the second set and third set, subsamples are analyzed—the non-Latino sample is analyzed in the second set (Table 3) and the Latino sample is analyzed in the third set (Table 4). Survey procedures are used in the analysis to account for the NHIS multistage sampling design; population sample weights are used as directed by the IHIS user notes to account for each survey year included in the sample. All results with *p*-values less than 0.05 are considered statistically significant and results with *p*-values less than 0.1, but greater than 0.05, are considered marginally statistically significant. All tests are two-tailed.

In the *first* set of logistic regression models (Table 2), Model 1 analyzes the association between Latino ethnicity and diabetes. Then in order to identify variables that operate as suppressors and mediators, as well as to check for robustness and stability of results, ten additional models are specified. Model 2 includes the variables in Model 1 and adds age and age squared. Model 3 includes the variables in Model 2 and sex. Model 4 includes the variables in Model 3 and race. Model 5 includes the variables in Model 4 and BMI. Model 6 includes the variables in Model 5 and smoking status. Model 7 includes the variables in Model 6 and the proxies for social inequality. Model 8 includes the variables in Model 7 and the proxies for social ties. Model 9 includes the variables in Model 8 and the proxies for acculturation. Model 10 includes the variables in Model 9 and the origin of an individual's Latino ethnicity. Model 11 includes the variables in Model 10 and the additional control variables.

In the *second* set of logistic regression models (Table 3), the analysis focuses on how the patterns of the associations of the variables and having diabetes operate in the non-Latino subsample, with the intention of contrasting these results to the first set of logistic regression models. In order to identify variables that operate as suppressors, mediators and moderators, as well as to check for robustness and stability of results, in total, nine models are specified. In this set, the origin of Latino ethnicity is not included, because the non-Latino sample did not answer this question. Model 1 examines the relationship between age and age squared and diabetes. Model 2 includes the variables in Model 1 and sex. Model 3 includes the variables in Model 2 and race. Model 4 includes the variables in Model 3 and BMI. Model 5 includes the variables in Model 4 and smoking status. Model 6 includes the variables in Model 5 and the proxies for social inequality. Model 7 includes the variables in Model 6 and the proxies for social ties.

Model 8 includes the variables in Model 7 and the proxies for acculturation. Model 9 includes the variables in Model 8 and the additional controls.

In the *third* set of logistic regression models (Table 4), the analysis focuses on how the patterns of the associations of the variables and having diabetes operate in the Latino subsample. In order to identify variables that operate as suppressors and mediators, as well as to check for robustness and stability of results, in total, ten models are specified. Model 1 examines the relationship between age and age squared and diabetes. Model 2 includes the variables in Model 1 and sex. Model 3 includes the variables in Model 2 and race. Model 4 includes the variables in Model 3 and BMI. Model 5 includes the variables in Model 4 and smoking status. Model 6 includes the variables in Model 5 and the proxies for social inequality. Model 7 includes the variables in Model 6 and the proxies for social ties. Model 8 includes the variables in Model 7 and the proxies for acculturation. Model 9 includes the variables in Model 8 and the origin of an individual's Latino ethnicity. Model 10 includes the variables in Model 9 and the additional controls.

Oaxaca-Blinder Decomposition

After examining these factors using logistic regression analysis, I will then use a Oaxaca-Blinder decomposition to establish a more thorough understanding of what can be explained about the difference of the likelihood of reporting being diagnosed with diabetes between the Latino population and the non-Latino population. Oaxaca-Blinder decomposition is a regression-based method that determines the degree to which the any disparity in a characteristic in a sample mirrors differences in the observed characteristics of a sample, and identifies important factors associated with the disparity (Chen and Rizzo 2010a).

First, one must estimate the logistic regression analyses for the two groups in question, which has been done in the prior chapter. After estimating the logistic regressions for non-Latinos and Latinos separately, the Oaxaca-Blinder technique decomposes these regressions into observable differences (differences based on the variables included in the regressions) and differences that are caused by unobserved differences (and possibly unobservable differences) between the groups. This is achieved by constructing a *counterfactual equation* where the intercept and coefficient of the Latino logistic regression equation is replaced with those from the non-Latino logistic regression equation. I conduct the decomposition analyses by applying the method described in Jann (2008) and Sinning (2008), which adjusts for the logistic regression models (Jann 2008; Sinning, Hahn and Bauer 2008). The results are then interpreted in the following manner: if the Latino subsample had the same observable characteristics as the non-Latino subsample, they would be a given amount more or less likely to report being diagnosed with diabetes.

To show how the decomposition works, in chapter 6, I begin with a simple model that examines the difference in the proportion of individuals who report being diagnosed with diabetes by Latino ethnicity and include the observable characteristic of BMI (a mediator of self-reported diabetes) (Table 6). Then I demonstrate the analysis when it decomposes observable characteristics that are suppressors by decomposing the difference and including the observable characteristics of age and race (Tables 7 and 8 and Figures 5 through 7). Finally, the full logistic regression model is decomposed by Latino ethnicity, which includes all of the variables in the model (Table 9). These results are condensed into more readable tables by grouping the results by variable and theoretical type (Tables 10 through 15). First, I decompose the diabetes difference identified in the whole sample. Then, I decompose the difference by age group. The results from all of these decompositions are reported in the same table. After interpreting the decomposition results, I revisit the research questions and synthesize all of the evidence to determine the conclusions and implications of this study.

CHAPTER 5: SUMMARY STATISTICS AND LOGISTIC REGRESSION RESULTS

At every level of analysis, descriptive, bivariate and multivariate, the data show that there are differences in both the prevalence of self-reported diabetes and the factors that are theoretically associated with diabetes for the Latino and non-Latino populations. The univariate or descriptive distribution of the characteristics under consideration in the analysis of this paper is reported in Table 1. The distribution of these characteristics are reported in three ways. First, in column one, the summary statistics are given for the full sample of adults in the 2006-2011 NHIS data (*n*=157,228 individuals), then columns two and three report the summary statistics for the non-Latino (*n*=128,963 individuals) and Latino (*n*=28,265 individuals) subsamples. Results of the full sample are consistent with other reports on these data and are also consistent with other reports of nationally representative data. Notable descriptive patterns will be referenced in relation to the key variables. The bivariate tests (the Wald tests and the Pearson weighted *f*-tests, instead of chi-square and *t*-tests due to use of complex survey design weighting) indicate that there is a statistically significant different mean or proportion for every variable when analyzed by Latino ethnicity (non-Latino results compared to Latino results). This is in part due to the large sample size, so inferences cannot be made on these bivariate results. However, they do suggest that further investigation of these differences are warranted. All summary statistics are presented in Table 1.

Summary Statistics

Diabetes. The summary statistics in Table 1 indicate that there is a difference in self-reported diabetes between Latino and non-Latino subgroups. In the full sample, 8.5% report that they have been diagnosed with diabetes. In the non-Latino sample, 8.4% indicate that they have been diagnosed with diabetes. However, in the Latino sample, 9.2% report that they have been diagnosed with diabetes. Through the examination of the patterns of some of the demographic characteristics, it will become clear that this 0.6% difference will be much wider once other key demographic variables are adjusted for, such as age.

Latino. In Table 1, people who identify as Latino comprise 11.2% of the sample. Individuals who identify as Latino also identify with specific countries or cultures of origin. Consistent with other data, the largest sub-group of Latinos are individuals with Mexican heritage at 59.3%. The second most-represented culture of origin is Puerto Rican with 11.2%. Those with Cuban heritage comprise 4.6%, and those with Dominican heritage comprise 3.5%. Due to concerns about identification, the other specific cultures of origin have been grouped into regions; in this sample, 16.2% of Latinos have other Central American or South American heritage and 5.3% have multiple Latino origins. As the literature suggests that patterns of health and illness vary in regard to specific ethnicities of Latino origin, these summary statistics indicate that any patterns identified as distinct to the Latino subgroup must be further investigated in relation to the cultures of origin specific to individual of Latino ethnicity.

Individual characteristics

Age. These data indicate that the average age of the full sample is 48 years. However, when the data are divided by self-identification with Latino, a noteworthy difference emerges. The average age of the non-Latino subsample is slightly older than the national average, 49 years. The average age of the Latino subsample is meaningfully *younger* than the national average, 42 years.

Sex. A cursory analysis of other standard demographic variables indicate that there are other meaningful differences between the Latino and non-Latino subsamples. While the reported *sex* of the full sample and the non-Latino sample are both approximately 54% female and 46% male, the Latino sub-sample is approximately 52% female and 48% male. This difference is consistent with other reports of Latino sex and gender patterns, which are explained by differences in immigrant composition.

Race. The racial patterns of the full sample and the non-Latino sample are fairly similar—the difference is that the non-Latino subsample is comprised by approximately 1% fewer White individuals and 1% more African-American individuals than the composition of the whole sample. In these samples, approximately 13% are African American, approximately 4% are Asian, about 1% are Native American and less than 0.5% are other. However, the racial composition of the Latino subsample is very different from both the whole sample and the non-Latino subsample. Some racial groups are much greater: over 91% identify as white and 2.4% identify as Native American (compared to approximately 82% and 1% of the full sample). Other racial groups are much less common in the non-Latino subsample: only 4% identify as African-American and only

1.5% identify as Asian. These differences in racial composition are likely to have an influence on the diabetes rates in the two groups, as the prior literature suggests that non-White racial groups have higher rates of diabetes.

BMI. The average BMI was very similar for all three groups. For the whole sample and the non-Latino sample, a mean BMI score of 27.1 is reported. For the Latino sample, a slightly higher mean BMI score of 27.8 is reported. As higher BMI scores are associated with higher rates of diabetes, this difference supports a hypothesis that Latinos will experience higher rates of diabetes than non-Latinos.

Smoking. Latinos report having ever smoked less than non-Latinos. In the whole sample, 20.6% of the population reports being a current smoker and 22.8% report that they no longer smoke, but did at one time. Over half (about 56.4%) report having never smoked. These are similar to the figures for the non-Latino population: 21.4% are current smokers, 23.7% are former smokers, and 54.7% have never smoked. In comparison, only 14.8% of Latinos are current smokers and only 15.5% of Latinos are former smokers. Almost 70% of Latinos have never smoked. As smoking is associated with higher rates of diabetes, this difference in smoking behaviors should reduce the level of diabetes associated with the Latino population.

Social inequality measures

Education. The first indicator of social inequality examined is the level of *education*. In the full sample, about 14% of the population have less than high school level of education, about 24% have a high school diploma, and close to 3% have a GED.

Over half of the respondents in the full sample (about 57%) have some college or higher levels of education: 20% have some college, about 10% have an associate degree, approximately 18% have a bachelor's degree, and almost 10% have a graduate degree. The non-Latino sample echoes the full sample, except fewer of the respondents have less than a high school education (approximately 12% compared to approximately 14%), and the proportion of respondents who completed a BA degree and a graduate degree were about 1% higher each.

However, the patterns of education level in the Latino population diverge from these patterns in many categories. The most stark contrast is that over one-third (approximately 36%) of the Latino population has a less than a high school level of education, which is more than twice as many than the approximately 14% full population and 12% of the non-Latino population without high school degrees. The patterns of high school completion and those with a GED are fairly similar to the other group's patterns, 22% and 3% respectively. More than a third of the Latino population (about 39%) have some college or greater, compared to over half in the full and non-Latino samples: about 16% have some college, close to 8% have an associate degree, about 10% have a BA and nearly 4% have graduate degrees. The relatively lower level of educational attainment in the Latino population predicts that outcomes influenced by level of education (including diabetes and other health outcomes) should be poorer.

Household income. Income in relation to the federal poverty level is another indicator of social inequality, which is based on an individual's household income relative to the federal poverty line. In the full sample, about 13% have an annual household income below the federal poverty line, about 16% have a household income level between the federal poverty line to 199% of the federal poverty line, approximately one-quarter of the population has a household income between 200% and 399% of the federal poverty line, about 32% have a household income above 400%. The income is unknown for about 13% of the population. The poverty level pattern is very similar in the non-Latino population.

However, the pattern varies substantially for the Latino population. In the Latino population, over 20% have a household income below the federal poverty level, almost twice the proportion of the full sample. Approximately 23% of the Latino population have a household income level between the poverty level and 199% of the poverty level, almost 10% greater than the proportion of the full and non-Latino sample. Approximately one-quarter of the Latino population have a household income level between 200% and 399% of the federal poverty level, almost the same as the full and non-Latino sample proportions. Only about 17% of the Latino population have a household income greater than 400% of the federal poverty level—about half of the proportion of the full and non-Latino samples. Income is unknown for about 13% of the Latino sample, similar in size to the proportion of unknown income in the full and non-Latino subsample. The relatively lower levels of household income of the Latino population indicates that there should also be a difference in outcomes correlated with income, such as diabetes.

Health insurance coverage. The final measure of social inequality is health insurance coverage. In the full sample, 84% of participants are covered by some type of health insurance and 15.7% do not have health insurance coverage. In the non-Latino sample, a greater proportion have health insurance coverage (86.4%) and fewer people do not have health insurance coverage (13.3%). Again, the Latino sample has a divergent pattern: about two thirds (64.8%) of Latinos have health insurance coverage and over one third (34.8%) do not have health insurance coverage. Studies show that health insurance coverage is a reasonable proxy for health insurance coverage, but that unlike education and income (which have impressively consistent patterns with health outcomes), the relationship between health insurance coverage and health outcomes are mixed.

Social ties measures

Number in household. The first measure of social ties is the number of people residing in the same household as the participant. Overall, the sample reports a mean number of people residing in a household of 2.4 people. Slightly fewer individuals live in non-Latino households, with a mean number of 2.3 people in the household. The Latino sample reports a higher mean of 3.1 individuals living in a household.

Marital status. In all three categories, a very similar portion of individuals are married. For the full sample and non-Latino sample, approximately 44% of the sample are married with spouses present in the household. About 1.5 percentage point more married individuals with spouses present in the household are reported in the Latino subsample. A small number of individuals report being married, but that their spouse does not currently reside in their household. For the full sample and the non-Latino sample, only about 1% of the sample has this type of marital status. More than double the amount of Latinos have this marital status, which is mainly attributable to immigrant status.

The rest of the marital status categories have more variation between the three samples. While about 9% of the whole sample report being widowed, approximately 10% of the non-Latino population and 5% of the Latino population reports being widowed. About 26% of the full sample and non-Latino sample have never been married, and a higher percentage (about 29%) of the Latino sample has never been married (again this may be due to the mean age difference in categories). These two patterns are most likely associated with the difference in age between the two groups. About 15% of both the whole and non-Latino samples are divorced, while only about 12% of the Latino sample is divorced. Conversely, while about 3% of the whole and non-Latino subsample are separated, about 6% of the Latino population is currently separated from their spouse. These differences may reflect the religiosity of the groups, as Latinos are more likely to be participants in a religion that is less accepting of divorce.

Acculturation measures

Nativity. The first proxy of acculturation is *nativity*. In the full sample, about 86% of the sample is native-born and 14% is foreign born. In the non-Latino sample, approximately 92% is native-born and 8% is foreign-born. In contrast, less than half (about 42%) of the Latino sample is native born, and 58% is foreign born. Based on the literature review, as immigrants are more likely to have better health outcomes, these summary statistics would suggest that the Latino population would be experiencing an immigrant health advantage.

Length of time in country. Of those that are foreign-born, the length of time in the United States does not vary widely between the non-Latino and Latino samples. In the non-Latino sample, about 1% of the foreign born have been in the United States for less than one year, about 11% of the foreign born have lived in the US between 1 year and 5 years, about 13% of the foreign born have lived in the US between 5 and 10 years, about 13% of the foreign born have lived in the US between 10 and 15 years, and over 60% have lived in the US for more than 15 years. In the Latino sample, less than 1% of the foreign born have lived in the US for less than 1 year, about 8% have lived in the US between 1 and 5 years, about 16% have lived in the US between 5 and 10 years, 15% have lived in the US between 10 and 15 years, and about 58% have lived in the US for over 15 years. As the literature indicates that immigrants who have been in the US for more than 15 years lose the protective immigrant effect and in fact experience poorer health outcomes, the initial health advantage indicated by the considerably greater proportion of Latinos who are immigrants is tempered by the summary statistic that indicates that more than half of those Latino immigrants have been in the US long enough for that immigrant advantage to dissipate.

Language of interview. The second measure of acculturation used is the *language of interview*. In the non-Latino sample, only 0.4% of the foreign-born opted to take the survey in another language. However, in the Latino sample, 65% opted to take the survey in English, which indicates that 1 out of every 3 Latinos took the survey in another language or combination of languages. Approximately 22% took the survey in Spanish and 13.2% took the survey in a combination of English and Spanish.

Controls

Region. The *geographic region* of residence in the US also varies by subgroup. Approximately the same proportion of individuals in both subgroups live in the Northeast (about 18% of non-Latinos and 15% of Latinos) and the South (approximately 37% for both subgroups). However, where about 26% of non-Latinos live in the North Central/Midwest region, less than 10% of Latinos live in this region. This shift reverses in the West, where only about 22% of non-Latinos live, but 39% of Latinos live. Given that there are differences in diabetes disparities by region, these differences suggest that that the differences in diabetes by Latino identity may also be affected. Thus, the models control for geographic region.

Survey year. The distribution of the sample across survey years is very similar in all three groups. For the full sample, each year comprises between 15.8% and 17.6% of the total. For the non-Latino sample, each year comprises between 15.9% and 17.5% of the total. For the Latino sample, each year comprises between 15.2% and 18.6% of the total. While all three are similar, the Latino sample does have the most variation. I account for these variations by including the sample year variable in all of the full models.

Logistic Regression Results

The summary and bivariate statistics indicate that a multivariate analysis of the association between Latino ethnicity and likelihood of diabetes is warranted, as the Latino and non-Latino populations vary statistically significantly, not only in regard to the outcome variable of diabetes, but also in regard to every variable in question. The logistic regression models begin with an investigation of the full sample, and suppressors and mediators are identified (by adding variables in subsequent fashion until the full model has been specified). Then, logistic regression analyses are conducted on both the non-Latino and Latino subsamples. First, I will briefly mention the variables added in each model, and the changes in the odds ratio representing the association between Latino ethnicity and diabetes. After reporting all of the results for this association, I will discuss the association between all of the variables and diabetes separately. After establishing the association between Latino ethnicity and diabetes, I will unpack the relationship between the other variables in the complete model analyzing the full sample (the key independent variables measuring social inequality, social ties, and acculturation, as well as the controlling variables), then I will examine how these variables affect the Latino population differently by running logistic regressions of these models using subsamples of the Latino population and the non-Latino population.

Latino ethnicity and diabetes. In the analyses of the full sample (Table 2), the association between Latino ethnicity and diabetes at first seems relatively small in contrast to all of the literature on the association between Latino ethnicity and diabetes. However, in the full model, after adjusting for individual characteristics, the measures of the three theoretical frameworks, the origin of Latino identity, and controls, Latinos are about 43% more likely than non-Latinos to report having been diagnosed with diabetes. Before I unpack the full model and reflect on the associations of all of the variables included and their association with diabetes, I will comment on how the association between being Latino and reporting being diagnosed with diabetes varies in each subsequent model. In doing so, the suppressors and mediators of this association surface.

The first regression, which analyzes only the association between Latino identity and diabetes, reveals a statistically significant, but somewhat smaller than anticipated disparity with an odds ratio of 1.094. This indicates that an individual who is Latino is 9.4% more likely to report being diagnosed with diabetes than an individual who is not Latino. However, as the other variables are added to the models, the relationship between Latino identity and reporting diabetes changes. Models 2 through 6 test how individual characteristics suppress or mediate the relationship between Latino ethnicity and diabetes.

Of the individual characteristics, age and race emerge as suppressors, or factors that increase the association between being Latino and having diabetes, while BMI and smoking emerge as mediators, or factors that reduce or account for the association, and sex does not have a strong influence on the relationship between Latino ethnicity and diabetes. Model 2 adds both *age and age squared* to the model, and the results show that age is a powerful suppressor of the relationship, consistent with the prediction based from the summary statistics results of the mean age of the populations. Once age is adjusted, the odds ratio depicting the association between Latino identity and diabetes spikes up dramatically to a statistically significant 1.691. This indicates that when controlling for age, an individual who is Latino is almost 70% more likely to report being diagnosed with diabetes than an individual who is not Latino. The age squared odds ratio of 0.999 indicates that the relationship between age and diabetes is not linear, but instead the odds of Latinos reporting being diagnosed with diabetes increase with age at a decreasing rate.

This indicates that if analyses were to be divided into subgroups of age, age would not be as strongly predictive of diabetes in the oldest age group.

The association between being Latino and having diabetes is not strongly influenced by an individual's *sex*. In Model 3, in which sex is added as a control, the odds ratio associated with Latino ethnicity decreases slightly (by 0.0003) from Model 2, but it is not a meaningful change. This indicates that sex is not an important mediator of the association between Latino identity and diabetes.

However, *race*, like age, is also a meaningful suppressor of the association between Latino identity and diabetes. When race is added to the model, the odds ratio representing the relationship between Latino ethnicity and diabetes increases again to a statistically significant 1.822 from 1.69 in the previous model. This indicates that when adjusting for race (in addition to age and sex), an individual who identifies as Latino is approximately 82% more likely to report being diagnosed with diabetes than an individual who is not Latino. Race emerges as a variable that must be explored more carefully. The suppressor effect largely stems from the difference in minority race composition of the Latino and non-Latino populations. As noted in the reflection in the summary statistics, while Latinos have a higher proportion of Native Americans (a racial group with disproportionately high rates of diabetes) compared to non-Latinos, non-Latinos have a notably higher proportion of African Americans and Asians (two other groups with disproportionately high rates of diabetes).

Of the two individual characteristics that are clinically linked to higher rates of diabetes (BMI and smoking behaviors), BMI reduces, and smoking status increases the association between Latino ethnicity and diabetes. Obesity is a risk factor for diabetes.

Once BMI is adjusted for (in addition to age, sex, and race), the association between Latino ethnicity and diabetes decreases—the statistically significant odds ratio is 1.7452. This indicates that an individual who identifies as Latino is 74.5% more likely to report being diagnosed with diabetes than an individual who is not Latino. This supports the initial reading of the summary statistics, as Latinos have a statistically significantly higher mean BMI score than non-Latinos, which partially explains their higher rates.

In contrast, *smoking* is a suppressor of the association between being Latino and reporting being diagnosed with diabetes. After adjusting for smoking status, the odds ratio associated with being Latino and reporting diabetes increases by approximately 4 percentage points. This is consistent with the interpretation of the summary statistics, which indicates that Latinos are *less* likely to smoke than non-Latinos, which reduces their risk of diabetes.

As smoking is the last individual characteristic I adjust for, this model establishes the diabetes disparity after adjusting for all individual characteristics. In contrast to the initial odds ratio, where Latinos were 9.4% more likely to report being diagnosed with diabetes than non-Latinos, this model indicates that after adjusting for individual characteristics, Latinos are about 79% more likely to be report being diagnosed with diabetes than non-Latinos. Now that a clearer understanding of the disparity has been established, I add the measures of the three theoretical models to the logistic regressions.

Models 6 through 8 add the measures of each theory identified as important to the association between being Latino and diabetes subsequently; first the measures for social inequality, then the measures for social ties, and finally, the measures for acculturation are added. All are shown to mediate the association between being Latino and self-reporting diabetes. In Model 6, the proxies for social inequality (education, household income, and health insurance coverage) are added, and the statistically significant odds ratio depicting the relationship between Latino identity and diabetes drops sharply by about 0.25. In other words, after controlling for education and income (in addition to all of the individual characteristics), an individual who identifies as Latino is still over 50% more likely to self-report diabetes than an individual who is not Latino. This indicates that social inequality matters, and that the disadvantage of the Latino community partially explains their higher rates of diabetes.

The measures for social ties also account for part of the association between being Latino and diabetes. In Model 7, the proxies for social ties (number of people in the respondent's household and marital status) are added to the model. The odds ratio depicting the association between Latino identity and diabetes is 1.5, indicating that Latinos are about 50% more likely to self-report diabetes than non-Latinos after adjusting for individual characteristics and social inequality proxies.

The final theory added to the models, acculturation, also reduces the relationship between Latino identity and diabetes. In Model 8, the proxies for acculturation, length of time in country (which is also a proxy of nativity) and language are added. After adjusting for acculturation, the odds ratio representing the association between Latino ethnicity and being diagnosed with diabetes decreases to about 1.46 and remains statistically significant. This indicates that an individual who identifies as Latino is approximately 46% more likely to report being diagnosed with diabetes than an individual who is not Latino.

The final two models add the nuances of Latino identity (specific ethnicities of Latino origin) and the controls. Latino ethnicities of origin again reduce the association between Latino ethnicity and diabetes, while the controls are found to increase the association slightly. Model 9 adds the origin of Latino ethnicity, which decreases the odds ratio representing the relationship between Latino ethnicity and diabetes to a statistically significant 1.412. This indicates that when adjusting for the culture of origin of an individual's Latino ethnicity (in addition to individual characteristics and proxies for all three theoretical frameworks), an individual who identifies as Latino is approximately 41% more likely to report being diagnosed with diabetes than an individual who is not Latino. This finding supports critiques of making generalizations about Latinos as a group.

Finally, Model 11 represents the full model of this study, which adjusts for all individual characteristics, proxies of all three theoretical models, the ethnic origins of individuals' Latino identity, and the controls (region of residence and year of survey). The controls act as a slight suppressor of the association between Latino ethnicity and self-reported diabetes, which is mainly due to the fact that a smaller proportion of Latinos live in the South, which is associated with higher rates of diabetes. This model indicates that after adjusting for individual characteristics, measures of the three theories (social inequality, social ties and acculturation), origins of Latino ethnicity, and controls, an individual who is Latino is 43% more likely to report being diagnosed with diabetes than someone who is not Latino. Carefully observing how this association is suppressed and mediated through each of the models gives insight into the dynamics influencing this relationship.

Latino ethnicity and diabetes prevalence hypothesis revisited

H_1: Latinos will have a higher prevalence of diabetes than non-Latinos.

These logistic regression findings reject the null hypothesis that there is no association between Latino ethnicity and diabetes and offer support for the alternative hypothesis that there is an association between Latino ethnicity and a higher prevalence of diabetes. These results justify the next stage of analysis in this study. Now that the association between Latino ethnicity and diabetes is established, I will unpack the relationship between all of the variables in the complete model and diabetes in the whole sample, the non-Latino sample and the Latino sample. After interpreting the results of the full logistic regression model for the non-Latino and Latino samples, I will then formally analyze the differences between these models with an Oaxaca-Blinder decomposition.

Full Model Logistic Regression Results by Full Sample and Non-Latino/Latino Subsamples

Where the last section focused on determining the suppressive or mediating role of the included variables in the models, this section focuses on understanding the relationship between each of the variables and diabetes. First, I examine the odds ratios of the association of each variable in the model and the likelihood of reporting being diagnosed with diabetes in order to get a sense of how each of the independent variables operate in the full sample. Then, in the discussion of the second and third set of logistic regression models, I will compare and contrast how these factors are associated with diabetes in the Latino and non-Latino subsamples. As I am interested in how these variables help inform the association between Latino ethnicity and diabetes, I primarily focus on the results of the complete model (which adjusts for individual characteristics, proxies of the three theories, the controls, and origin of Latino ethnicity where possible) to establish the understanding of how these variables operate in the full sample and both of the subsamples. The Non-Latino sample results very closely match those of the full sample results, which is to be expected based on the relatively small size of the non-Latino sample. The logistic regression results of the Latino sample have the most divergent results. Tables 3 and 4 show all of the models sequentially for the non-Latino and Latino model respectively. Table 5 juxtaposes the results of the full model for each sample, for convenience.

Individual Characteristics

Age. In my discussion of the relationship between Latino ethnicity and diabetes, age emerged as a key suppressor. This is because the Latino population is on average much younger than the non-Latino population, and individuals' likelihood of being diagnosed with diabetes increases with age. The logistic regression results of all three samples support this. In the complete model analyzing the whole sample, the non-Latino sample, and the Latino sample, an individual's likelihood of being diagnosed with diabetes increases with age at a decreasing rate. This relationship is statistically significantly likely at the 0.001 level. This finding is consistent with other literature examining the relationship of diabetes and age. As one ages, one is more likely to be diagnosed with diabetes. Also, consistently across all three samples, the odds ratio associated with age squared indicates that the association between age and diabetes is not linear, but that the risk of being diagnosed with diabetes increases with age at a decreasing rate.

Sex. The pattern in all three samples suggest that women are less likely to self-report diabetes than men. However, this is only statistically significant in the whole and non-Latino samples. For both the whole sample and the non-Latino sample, the full model results indicate that women are about 20% less likely to report being diagnosed with diabetes than men. However, for the Latino subsample, the full model results report an odds ratio of 0.96, which suggests that Latina women may be less likely to report being diagnosed with diabetes, but the results do not allow this finding to be stated with confidence. Harkening back to the descriptive statistics, those results foreshadowed this finding. This finding indicates that for the Latino population, both sexes are equally vulnerable to being diagnosed with diabetes.

Race. Race is a key suppressor variable in this set of analyses, as has already been addressed. There are race differences within the full sample and the non-Latino sample, but not within the Latino sample. In both the analysis of the full sample and the non-Latino sample, one’s likelihood of reporting having been diagnosed with diabetes statistically significantly varies for those of different races, with the exception of other race. In both the whole and non-Latino samples, and where the comparison group is

Whites, African Americans are over 50% more likely to report having been diagnosed with diabetes than individuals who identify as White. Asian Americans are found to be 75.7% more likely to self-report diabetes in the whole sample and 83% more likely to self-report diabetes in the non-Latino sample. Native Americans are also identified as more likely to self-report diabetes in the whole sample and non-Latino subsample analyses. In the whole sample Native Americans are about 66% more likely to report diabetes, and in the non-Latino sample, Native Americans are twice as likely to report diabetes as individuals who identify as white. All of these associations are statistically significant. One's likelihood of reporting being diagnosed with diabetes is not statistically significantly different for individuals identifying with races other than African American, Asian American or Native American compared to individuals who identify as white in any of the samples.

While race is a key suppressor variable in the full sample and in the non-Latino subsample, it does not have an effect in the models analyzing the relationship between the social determinants of health and diabetes in the Latino subsample. None of the race categories has a statistically significant association with diabetes in any of these models, nor does adding race to the model affect any of the associations between the other variables and diabetes in the model. In other words, there is no internal differentiation by race within the Latino population.

BMI. In all three samples, one's likelihood of being diagnosed with diabetes increases as an individual's BMI increases. The odds ratio is very consistent across the models, which indicates that the effect of BMI has a relatively independent effect on the likelihood of being diagnosed with diabetes. This relationship is statistically significantly likely at the 0.001 level in all three models. Like age, BMI operates very similarly in the full sample models and in the Latino models.

Smoking status. Smoking does not have as clear or strong an association with diabetes in this analysis as it does in some of the studies in the literature. In the whole sample, individuals who have never smoked are about 7% less likely to report diabetes than individuals who are current smokers, and this is statistically significant at the 0.05 level. Those who have quit smoking are about 8% *more* likely to report diabetes than those who are current smokers. For non-Latinos, the only statistically significant association is this association—that former smokers are more likely to report diabetes than current smokers. For the non-Latino population, there are no statistically significant associations between smoking status and diabetes, although the odds ratios related to smoking status are in the same direction as the other models.

Social Inequality Measures

Educational attainment level. The general finding for all three groups is that there is a statistically significant protective association between education level and self-reported diabetes. However, the findings for the Latino subgroup are a little different than the findings for the whole sample and non-Latino group. The full model for both the whole sample and the non-Latino sample finds a fairly linear relationship between education and self-reported diabetes: one's likelihood of being diagnosed with diabetes decreases as an individual achieves higher levels of education compared to those with less than a high school education. For example, for both of these samples, an individual with a high school diploma is approximately 15% less likely to report diabetes than an individual who has a less than high school level of education, and for those with a bachelor's degree or graduate degree, an individual with a higher level of education is approximately 45% less likely to report diabetes than an individual with a less than high school level of education.

In the complete model analyzing the full sample and the non-Latino sample individuals who have passed *general education development tests (GED)* are not more or less likely to report diabetes than individuals with a less than high school level of education. This finding is consistent across all of the models in which education was included and indicates that for the general population, getting a high school degree is more protective than GED certification.

In the whole sample and non-Latino sample, individuals who have some college, but no college degree, and those who have an associate degree are both about 20% less likely to report diabetes than individuals with a less than high school level of education. These associations are statistically significantly different from the reference group, those with less than a high school education.

However, in both the whole sample and the non-Latino subgroup analyses, attaining a *bachelor's degree* or a *graduate degree* has a much greater, and statistically significant, effect on the likelihood of being diagnosed with diabetes. In these samples, an individual with a bachelor's degree or graduate degree is approximately 35% less likely to be diagnosed with diabetes than an individual who has a less than high school education. These associations suggest that there may be diminishing returns to education in regard to diabetes as the relationship between higher levels of education and self-reported diabetes are not linear.

For the Latino subsample, the association between education and diabetes diverged from the patterns found in the whole and non-Latino sample, although the same general premise still held true: more education reduced self-reported diabetes prevalence. For the most part, education seems to be more protective against reporting being diagnosed with diabetes for the Latino population than for the full sample. However, these findings are not statistically significant for all educational attainment categories.

Similarly to the findings for the other samples, Latino individuals with a high school diploma were about 20% less likely to report diabetes than Latinos without a high school diploma. This relationship is statistically significant. This is a slightly more protective association than the one identified in the complete model of the full sample and non-Latino samples.

The educational attainment level that has the most notably different results for Latinos is the GED. In most other studies, usually due to sample size, those who have attained GED certification are grouped with those with high school diplomas. However, the NHIS allows for separate analysis. In the complete model analyzing the Latino sample, the odds ratio for individuals who have passed *general education development tests (GED)* is 0.631. This finding is statistically significant at the 0.01 level. This indicates that for the full sample, an individual who identifies as Latino with GED certification is almost 37% less likely to have been diagnosed with diabetes than an individual who identifies as Latino with a less than high school education level. This finding is consistent across all of the models in which education was included. This finding warrants further discussion. Latinos have a markedly higher proportion of individuals who have a less than high school level of education (approximately 36% compared to 11% of non-Latinos who have a less than high school level of education). As there is such a large proportion of individuals in this category, according to human capital theory, individuals who are able to achieve a higher level of education will receive greater rewards than individuals in a group with a lower proportion of individuals with less than a high school level of education. This finding warrants further investigation in relation to a wider range of outcomes, but this finding, although unexpected, is not illogical.

Also divergent from the patterns of the whole and non-Latino samples is the finding that individuals with some college or an associate degree are not statistically significantly more or less likely to report diabetes than individuals with a less than high school education. The patterns of the association for those with associate degrees is similar to those of the other samples, but the relationship is not statistically significant. This is distinct from how these levels of education relate to self-reported diabetes in the other samples.

However, attaining a *bachelor's degree* has a much greater effect on reducing the likelihood of being diagnosed with diabetes than it does for the non-Latino group. The odds ratio associated with attaining a bachelor's degree is 0.528. An individual who identifies as Latino with a bachelor's degree is almost 50% less likely to report being diagnosed with diabetes as an individual who identifies as Latino who has a less than high school education. This relationship is statistically significant at the 0.001 level. However, individuals with *graduate degrees* do not experience the same degree of protective effect, their associated odds ratio is a less protective 0.686. Therefore, an individual with a graduate degree is over 30% less likely to have been diagnosed with diabetes than an individual with a less than high school education. This relationship is statistically significant at the 0.05 level. The implications of these findings suggest that more research investigating how education influences health outcomes for the Latino population is warranted, but the most general implication is that researchers should not assume that education operates the same way for individuals in different racial and ethnic groups.

Income in relation to the federal poverty line. The second measure of social inequality is an individual's household income in relation to the federal poverty line. In all three samples, individuals with higher levels of income are less likely to report being diagnosed with diabetes than individuals who have incomes lower than the federal poverty line, and the association is reasonably linear. The federal poverty level for a family of four in 2011, was an annual income of $22,350 (Sebelius 2011).

For individuals who have household incomes at the poverty line to just less than twice the income of the poverty line (*100% to 199%* of the federal poverty line), often classified as "near poor," individuals in both the full sample and the Latino sample report statistically significantly lower odds of self-reported diabetes than individuals who have household incomes below the poverty line. This indicates that individuals who have a household income of 100% to 199% of the federal poverty line (which in 2011 was between $22,350 to less than $44,700 for a family of four) are about 11% less likely to report diabetes in the full sample and almost 25% less likely to report diabetes in the Latino sample. Both of these associations are statistically significant. However, there is no statistically significant association between income and diabetes for individuals in the non-Latino sample who have an income between 100% and 199% of the federal poverty line. This is consistent with other studies that suggest that the “near poor” are often as vulnerable as individuals who are classified as living in poverty. This also somewhat consistent with the GED finding—as there a greater proportion of Latinos living in poverty, being able to move up to the next level of income serves as more protection.

For all three samples, individuals who are living with household incomes twice the level of the poverty line to just below four times the level of the federal poverty line (*200% to 399%* of the federal poverty line—which for a family of four in 2011 would be an annual household income of $44,700 to less than $89,400) are statistically significantly less likely to report diabetes than individuals living below the poverty line. Non-Latino individuals with a household income of 200% to 399% of the federal poverty line are about 25% less likely to report diabetes than individuals with household incomes below the federal poverty line. Latino individuals with this level of household income are 33% less likely to report diabetes than Latinos with household incomes below the federal poverty line.

Similarly, for all three samples, individuals who are living with household incomes over four times the level of the poverty line (over *400%* of the federal poverty line—which for a family of four in 2011 would have been an annual household income of $89,400 or greater) are statistically significantly less likely to report diabetes than individuals living below the poverty line. Non-Latino individuals with a household income over 400% of the federal poverty line are about 40% less likely to report diabetes than individuals with household incomes below the federal poverty line. Latino individuals with this level of household income are almost 50% less likely to report diabetes than Latinos with household incomes below the federal poverty line. The overarching pattern identified in the association of household income and diabetes for Latino and non-Latino individuals is that the protective effect of income appeared to be greater for Latinos at each increased level of income.

Health Insurance Coverage. The final measure of social inequality is health insurance coverage, which was shown to mediate the relationship between Latino ethnicity and diabetes in the previous models. Typically access to health insurance indicates a higher socioeconomic status, and higher socioeconomic status is associated with lower levels of diabetes. However, in this study, individuals from all three samples (the whole sample, the non-Latino subsample, and the Latino subsample) *with* health insurance coverage are *more* likely to report diabetes than individuals *without* health insurance coverage. All three groups show that individuals without health insurance coverage are about 30% less likely to report diabetes than individuals with health insurance coverage.

This is theoretically sound for a few reasons. First, individuals need to have access to health care in order to get a diagnosis of diabetes. Individuals with health insurance have at least one less barrier to health care than individuals without health care. Also, individuals without health insurance coverage are much less likely (both in this study and in general) to be younger. Younger individuals are less likely to have developed diabetes. Finally, another group that is less likely to have health insurance coverage are immigrants. The literature suggests that immigrants are also healthier. These are three reasons directly linked to the study sample that would lead to the outcome of individuals without health insurance coverage being less likely to have diabetes than individuals with health insurance coverage.

Social inequality hypothesis revisited

H_2: Measures of social inequality will be associated with lower levels of diabetes in Latinos.

These results reject the null hypothesis that there is no association between social inequality and self-reported diabetes for the Latino diabetes disparity. Overall, all three of the social inequality proxies are strongly associated with diabetes in the Latino population. Education and household income both have inverse associations with diabetes—higher levels of education and income are associated with lower odds of self-reported diabetes. Health insurance coverage worked against the hypothesis as individuals with health insurance coverage had a higher prevalence of diabetes. Household income and health insurance coverage had very similar patterns of association both in the Latino and non-Latino populations (although income appears slightly more protective for the Latino population). Education patterns emerged as distinct between the Latino and non-Latino populations, particularly in regard to the effect of GED certification (statistically significantly protective in the Latino population and not in the non-Latino population) and some college and associate degrees (not statistically significantly linked to diabetes in the Latino population while associated statistically significantly with lower odds of diabetes in the non-Latino population).

Social Ties Measures

Number of people in household. There is no statistically significant association between the number of people in the household and the odds of reporting diabetes for any of the samples. These odds ratios are quite close to 1.0 across all models.

Marital status. Marital status is also largely not statistically significantly associated with the odds of reporting diabetes. In the non-Latino sample, individuals who are currently separated from a spouse are statistically significantly more likely (about 15%) to report diabetes than individuals who are married and live with their spouse. For both the non-Latino and whole sample, individuals who have an unknown marital status are statistically significantly less likely (about 50% less likely) to report diabetes. There are no statistically significant associations between marital status and diabetes for the Latino population.

Social ties hypothesis revisited

> H_3: Measures of social ties will be associated with lower levels of diabetes in Latinos.

Based on the logistic regression results, there is not enough evidence to reject the null hypothesis that social ties are not associated with the diabetes prevalence of Latinos. The measures for social ties do not fully capture the social ties of an individual, so the theory needs to be tested with better measures with data that allow for it.

Acculturation Measures

Nativity and length of time in the United States. The logistic regression analyses of all three samples suggest that there is not a strong association between either nativity or length of time in country and an individual's odds of reporting diabetes. The proxies for acculturation are if an individual is foreign born and the length of time a foreign born individual has spent in the United States. In the logistic regression results for individuals in the full and the non-Latino sample, the only statistically significant difference between native born and foreign born individuals in regard to diabetes was for foreign born individuals who have lived in the United States for greater than 15 years. For both the full sample and the non-Latino sample, immigrants who have lived in the US for more than 15 years are more likely to report being diagnosed with diabetes (17% more likely and 20% more likely respectively). These findings are consistent with the theories that suggest that immigrants lose any protective advantage with greater time in the US However, surprisingly, there is no strong evidence that immigrants experience any protection from diabetes in this model.

However, for the Latino sample, nativity and length of time in country are not statistically significantly linked to diabetes at all. When specifying the full model and including measures for the theories, there does not appear to be any protective effect associated with being an immigrant and self-reporting diabetes. While studies of the paradox have put a lot of emphasis on the immigrant effect, based on recent statistics on obesity and diabetes in Mexico and for Puerto Ricans, it is not entirely unexpected that there is not a protective effect for Latino immigrants in regard to diabetes. However, it is somewhat surprising that there is no strong evidence of the association shifting over time.

Language preference. The second proxy for acculturation is an individual's language preference. This measure is not statistically significantly associated with diabetes in the analysis of the full sample nor the non-Latino sample. For the Latino sample, one's likelihood of reporting being diagnosed with diabetes is also not largely statistically significantly affected by one's language preference, with the exception of individuals who took the survey in a combination of English and Spanish. The associated odds ratio for individuals who took the survey in a combination of English and Spanish and diabetes is 1.235. This indicates that individuals who took the survey in a combination of English and Spanish were about 24% *more* likely to be diagnosed with diabetes than individuals who took the survey in English. This supports the segmented assimilation theory which suggests that those who have not assimilated or acculturated experience a protective effect as their lifestyle is still largely aligned with their culture of origin, which usually has healthier diets and other health behaviors. This theory also suggests that individuals who have successfully assimilated or acculturated have access to health resources (both in regard to health literacy and health care) that may be protective. Individuals who are in the process of moving from one group to the other may be the most vulnerable. There is some support for that theory in these findings.

Acculturation hypothesis revisited

H4: Measures of acculturation will be associated with greater levels of diabetes in Latinos.

These findings largely fail to reject the null hypothesis that there is no association between acculturation and the Latino diabetes disparity. There is no statistically significant relationship between immigrant status and diabetes prevalence in the Latino subgroup analysis. There is some evidence that individuals who took the survey in English and in Spanish (which indicates that they are in the midst of an acculturation process) are more likely to self-report diabetes.

Origin of Latino Ethnicity

In the full sample, ethnic origin of an individual's Latino ethnicity is largely not statistically significantly related to one's likelihood of reporting being diagnosed with diabetes with the exception of individuals with Cuban heritage or other Central and South American heritage. Individuals with Cuban and other Central and South American heritage are actually less likely (more than 25% less likely) to be diagnosed with diabetes. Individuals with Mexican, Puerto Rican and Dominican heritage do not experience statistically significant likelihoods of reporting diabetes, although all of the odds ratios indicate that general direction of the association is that these groups would be more likely to report diabetes. However, when this pattern is investigated more closely in the Latino subgroup analysis, while the directions of the associations are the same as in the full sample, the statistical significance disappears. This is in part due to the fact that the association between heritage of origin and Latino identity are collinear. In the non-Latino sample, the culture of origin of an individual's Latino ethnicity is dropped from the analysis because no participants reported a Latino ethnicity of origin.

Controls

Geographic region of residence. In the full sample and the non-Latino sample, the region of the US where an individual resides is statistically significantly related to the likelihood that an individual has reported having been diagnosed with diabetes. However, geographic region of residence is not statistically significantly associated with prevalence of diabetes in the Latino population. The patterns of the association between region and diabetes is very similar for both the whole sample and the non-Latino sample. First, for both groups, individuals in the North Central or Midwest region of the country are about 17% more likely to report having been diagnosed with diabetes than individuals living in the Northeastern region. This relationship is statistically significant. For individuals in the whole sample and the non-Latino subsample, compared to individuals living in the Northeastern region of the US, individuals in the South are approximately 25% more likely to report having been diagnosed with diabetes. Finally, for the whole sample and the non-Latino sample, individuals residing in the West of the US are about 9% more likely to report having been diagnosed with diabetes than individuals living in the Northeast. In the Latino subsample, the region of the US where an individual resides is *not* statistically significantly related to the likelihood that an individual has reported having been diagnosed with diabetes. This is a notable difference from the trends identified in the models analyzing the full sample.

Year of survey. In the full model for the whole sample and the non-Latino sample, individuals from the 2010 year of survey were statistically significantly more likely to report diabetes than individuals from the 2006 survey year. There were no other statistically significant associations between years of survey and diabetes. In the Latino sample, the models reports that individuals surveyed in 2007, 2010 and 2011 were all statistically significantly more likely to report diabetes than individuals surveyed in 2006. These findings support the decision to control for year of survey in the model, as it allows the model to control for secular influences by annum.

Conclusion

The logistic regression analyses established a clear Latino diabetes disparity. After adjusting for individual characteristics, measures of social inequality, measures of social ties, measures of acculturation, origin of Latino heritage, and controls, Latinos are almost 43% more likely to report being diagnosed with diabetes than non-Latinos. Individual characteristics of age, race, and smoking behaviors are identified as suppressors of the association between Latino identity and diabetes. Conversely, measures of social inequality, social ties, acculturation, and origin of Latino heritage are all potential mediators of the association between Latino identity and diabetes. The subgroup analysis identified an association between the individual characteristics of age, race, BMI and smoking status and the odds of reporting diabetes in both Latino and non-

Latino populations. The subgroup analysis also offers evidence in support of an association between the social inequality measures and the Latino diabetes disparity. There is marginal support for an association between acculturation measures and the Latino diabetes disparity. There is no support for an association between social ties and the Latino diabetes disparity. As there is evidence that all three theories should help explain the relationship between Latino ethnicity and diabetes, even though there is only strong evidence to support the association between social inequality and the Latino diabetes prevalence in the subgroup analysis, I will conduct an Oaxaca-Blinder decomposition in the next chapter to further investigate the role of these variables in the Latino diabetes disparity.

CHAPTER 6: OAXACA-BLINDER DECOMPOSITION RESULTS

In the pursuit to understand the factors that are driving the diabetes disparity between the Latino and non-Latino groups, I explored the associations between key social determinants of health and self-reported diabetes through logistic regression analyses in the whole sample, as well as in the Latino and non-Latino subgroups. These analyses indicated that there are differences between the social determinants that are associated with self-reported diabetes. The logistic regression results establish the diabetes disparity between Latino and non-Latino groups, and they also suggest that some of the key explanatory variables operate differently in relation to diabetes for the Latino and non-Latino groups. These findings warrant further investigation using the Oaxaca-Blinder decomposition analysis. Oaxaca-Blinder decomposition analysis decomposes the differences between the two subsample logistic regression analyses, identifying what portion of the difference in self-reported diabetes is explained and unexplained, as well as what portion of the explained difference are attributable to different variables (observed characteristics).

The Oaxaca-Blinder decomposition allows for a further investigation of the role of the factors that are associated with a difference in an outcome (O'Donnell et al. 2008). I decompose the differences in Latino and non-Latino diabetes outcomes for the full sample and by age group (Tables 9 through 15). The tables report both the comprehensive decomposition results and the percentage of disparities between Latino and non-Latino groups attributable to differences in each characteristic. The observable characteristics of the Latino population inform the explained difference in self-reported diabetes between the Latino and non-Latino population. More specifically, individual characteristics (particularly age, race, BMI, and smoking habits), measures of social inequality, measures of social ties, measures of acculturation, and measures of Latino ethnic origins inform the explained difference in self-reported diabetes between the Latino and non-Latino population. Social inequality measures contribute a larger part of the explained difference than social ties measures or acculturation measures.

Oaxaca-Blinder Decomposition of Latino Diabetes Difference with Age and Race

The Oaxaca-Blinder decomposition "decomposes" a difference in an outcome variable by group into explained and unexplained differences by constructing the counterfactual equations (where the intercept and coefficient of the equation associated with the Latino sample are replaced with those of the non-Latino sample). This results of this technique show what the diabetes prevalence in the Latino population would be if the Latino population had the same levels of the observable characteristics as the non-Latino population. To show how the decomposition works, I begin with a simple model that examines the difference in the proportion of Latino and non-Latino individuals who report being diagnosed with diabetes and include the observable characteristic of BMI (a mediator of self-reported diabetes that helps explain the difference by ethnicity) (Table 6). Again, as an example, this will show what the diabetes prevalence in Latinos would be if they had the same mean BMI as non-Latinos (not controlling for any other factors).

Then I demonstrate the analysis when it decomposes observable characteristics that are suppressors, by decomposing the difference and including the observable characteristics of age and race (Tables 7 and 8). Finally, the results of the decomposition of the diabetes difference including all of the variables included in the full logistic regression model are discussed (Table 9). These results are condensed into more readable table (Tables 10 through 15) by grouping the results by variable and theoretical type. In this table I first decompose the diabetes difference identified in the whole sample. Then, due to the powerful suppressor effect of age, I decompose the difference within three age groups. I constructed the three age groups by dividing the groups into tertiles by the age distribution in the full sample. The first tertile, Age group 1, includes 18 to 37 year old individuals. The second tertile, Age group 2, includes 38 to 56 year old individuals. The third tertile, Age group 3, includes individuals over the age of 56 years.

Decomposition with BMI only

As BMI is a mediator of diabetes in the regression models, I selected it as the variable to demonstrate how the decomposition functions (Table 6). For the full sample of those over the age of 18 years who participated in the NHIS from 2006 to 2011, the proportion of individuals who reported being diagnosed with diabetes is approximately 8.5% for the non-Latino population and approximately 9.2% for the Latino population, which amounts to a difference of 0.72 percentage points. This difference is statistically significant at the 0.01 level.

The decomposition reports that the *explained difference* in the diabetes outcome for Latinos and non-Latinos when including BMI in the model is 0.0051, which accounts for about 71% of the difference. There is an unexplained difference of about 0.0021, which indicates that unobserved factors account for about 29% of the difference in diabetes between Latinos and non-Latinos. This finding is also statistically significantly different at the 0.001 level. This means that if Latinos had the same mean BMI score as non-Latinos, the proportion of Latinos with diabetes would decrease by 0.005 and instead of 9.2% of Latinos, only 8.7% of Latinos would report diabetes. This makes intuitive sense—the observed mean BMI score of the Latino group is approximately 27.8, while non-Latinos have a mean BMI score of 27.1. If the Latino group was the same in all other aspects, but their mean BMI decreased to a mean score of 27.1, it would be logical to expect that their reported diabetes incidence would also decrease given the association between BMI and diabetes (seen in analyses in previous chapters). The decomposition tables also report the portion of the explained differences attributable to different distributions of each variable in Latino and non-Latino populations. However, in this decomposition, only BMI is included, therefore 100% of the explained difference is attributable to BMI (as would be true for any single variable model, no matter how much or little variance is explained); this finding is statistically significant.

This table also shows the decomposition results by age group, and the key takeaway from the decompositions of the difference in diabetes outcomes by age groups in this model is fairly straightforward. First, the difference in diabetes prevalence increases with age. In age group 1, the proportion is very low for both groups, and there there is only a small difference in diabetes prevalence between Latinos and non-Latinos

(0.0023), although even this small difference is statistically significant. In age group 2, the diabetes prevalence is much higher, and the disparity increases to a 3 percentage point difference, and this difference is statistically significant. In age group 3, diabetes prevalence is much higher still, and the disparity reaches a 9.5 percentage point gap, which is also statistically significantly different. BMI does not explain the same amount of the difference for each age group. For the youngest age group, BMI explains about 65% of the difference (.0015/.0023). For the middle age group, BMI explains about 12% of the difference (.0044/.0367). Finally, in the oldest age group, BMI also explains about 12% of the difference (.0110/.0951). When BMI is the only factor considered, it is easy to determine that BMI is more strongly associated with the diabetes disparity in the younger age group than in the older two age groups. Of course, this mini-decomposition of the Latino diabetes difference including BMI is used to demonstrate the method, not to gather evidence for the argument. BMI will be investigated more carefully in the discussion of the decomposition of the difference including all of the variables in the full model.

Decomposition of difference with age and race

However, before I discuss those results, I will walk through one additional relatively simple model to establish a baseline for how suppressors appear and influence the results of a decomposition. In this decomposition, I have added both of the key suppressors of age and race. Table 7 reports the explained difference as a percentage of the difference and the explained difference by characteristic as a percentage of the explained difference. Table 8 depicts the relative role of each characteristic as part of the explained difference, which has been calculated using absolute values of the contribution of each variable (of which the raw results can be positive or negative). These are more detailed charts and a different way of presenting the material than in the BMI example.

These percentages are reported in two ways for some strategic reasons. The raw numbers are easy to interpret, but because they are so small, it is somewhat difficult to immediately grasp the effect in regard to the total explained difference. Therefore, the percentages are calculated and reported in Tables 7 and 8. First, the percentage is reported in relation to the total, which results in both positive and negative percentages (depicted in Figure 5). Second, the absolute values associated with each characteristic are summed and this new total is used to give a more accurate depiction of the role each variable plays in the portion of the explained difference. These results are also presented in two bar charts, Figures 6 and 7. Many papers either misstate the effect of independent characteristics because they only do the former (Langellier et al. 2014), which tends to inflate individual contributions, or they neglect to discuss the results or implications of the variables that are associated with a negative raw number (which indicates that that variable is suppressing the difference not explaining the difference) (Chen and Rizzo 2010b). However, other studies also note that often factors detract from the explained gap, in other words, a variable *increases* the mean predicted gap instead of explaining it (or decreasing it) (Sen 2014). This is important to note for both the reporting and the discussion of the results of the decomposition of the difference in diabetes outcomes of Latinos and non-Latinos.

First, I will walk through the suppressor example, and then will discuss the results and implications of the full decomposition. As in the previous example, the established difference in diabetes between the two groups remains 0.007. The decomposition indicates that the characteristics of age and race are acting as suppressors of the ethnicity difference in the proportion of individuals who report having been diagnosed with diabetes. For the full sample of adults, the decomposition reports that the explained difference in the diabetes outcome for Latinos and non-Latinos is a statistically significant -0.0288. When a negative number is reported for the "explained difference," it indicates that the true difference based on the included factors is *larger* than the initially reported difference. This brings the total gap in this model to 3.6 percentage points, none of which is explained by the population characteristics included in this model. Therefore, adjusting the Latino populations' observable characteristics (in this decomposition that would be age and race distributions) to the levels of non-Latinos would *increase* the proportion of Latinos who would report being diagnosed with diabetes by 2.9 percentage points; i.e. if Latinos had the same mean age and the same racial composition as non-Latinos, they would report a 12% prevalence of self-reported diabetes.

These findings are consistent with the initial reading of the summary statistics. Latinos report a much younger mean age (which would suppress the diabetes rate), and the Latino population has a smaller proportion of racial minorities (as a whole—they do have a greater proportion of Native Americans) than the non-Latino subgroup, which also suppresses the reported prevalence of diabetes as both age and race are associated with higher rates of diabetes**.** Therefore, in relation to the reported difference, in this decomposition, the explained difference (the difference based on observable characteristics) is -400%, which brings the total unexplained difference to 500% of the reported difference. This number is not easily interpreted compared to factors that explain a portion of the difference. However, what this means is that differences in age and race *suppress* the difference in diabetes between Latino identity and diabetes. Again, in other words, if Latinos had the same mean age and the same racial composition as non-Latinos, the prevalence of diabetes in the Latino subgroup would be much greater.

The other key evidence that a decomposition offers is the relative role of the different observable characteristics in explaining the difference. Thus, I turn back to the direct interpretation of the findings of the portion of explained difference attributable to the individual characteristics included in the decomposition. For the full sample of adults, age comprises a statistically significant -0.0248 of the -0.0288 explained difference. This means that if Latinos had the same mean age as non-Latinos (if they were about 7 years older on average), the proportion of Latinos with diabetes would increase to about 11.7%. Table 7 shows what portion of the explained difference age comprises. Many studies report the explained difference attributable to individual characteristics as a portion of the total explained difference. If I do this with age, it looks like age is responsible for 86% of the difference (Table 8). However, that is overstating the role of age in the explained difference. As there are both factors that contribute and detract from the total explained difference, while not technically incorrect to make the preceding statement about the role of age, it is more precise to report the role in regard to all of the contributions. A better way of considering these figures is by taking an absolute value of each, summing these, and then recalculating the percentage with this new denominator. After these calculations, the table reports that the decomposition attributes approximately 81% of the explained difference in the diabetes outcome of Latinos and non-Latinos to age (Table 9). In this specific decomposition, this may not seem very important, but this will become more important in the decomposition of the difference including all of the variables in the full model.

Race is the other suppressor, and the decomposition of the difference indicates that the difference in distribution of African Americans, as opposed to the distribution of Asians and Native Americans, accounts for most of the explained difference attributable to race. In regard to the raw numbers, if Latinos had the same distribution of African-Americans as non-Latinos, the proportion of Latinos reporting being diagnosed with diabetes would *increase* from 9.2% to about 9.6%. However, if Latinos had the same distribution of Native Americans as non-Latinos, the proportion of Latinos reporting being diagnosed with diabetes would *decrease* from 9.2% to about 9.1%. Both of these findings are statistically significant at the 0.001 level. Again, this is consistent with the intuitive reading of the summary statistics. All three groups have higher rates of diabetes than non-Latinos. The difference is due to the fact that African Americans comprise almost 13.7% of the non-Latino population, but only 3.9% of the Latino population. If African Americans also comprised 13.7% of the Latino population, based on this decomposition including age and race, the proportion of Latinos with diabetes would be much higher. Conversely, Native Americans, another group that is disproportionately likely to be diagnosed with diabetes, comprise just 1.0% of the non-Latino population, but 2.4% of the Latino population. Therefore, if the Latino population's Native American population decreased to 1.0%, I would also see a lower proportion of Latinos with diabetes. The difference in the African American composition of the two groups is accountable for about 15% of the explained difference, and the difference in the Native American composition of the two groups is accountable for about 3% of the explained difference.

For this example, the key takeaway from the decomposition of the difference in diabetes outcomes including age and race by age groups, is that the individual characteristics contribute differently across the age groups. Figure 5 depicts the difference in diabetes outcome and the proportion of that difference that is explained and unexplained by the decomposition. Figure 6, the stacked bar charts, depicts the role of each characteristic in the explained difference by the raw percentages. The bar chart for the whole sample shows that about 86% of the difference is explained by age and the rest is explained by race. However, for the youngest age category (which has the smallest difference), the role of race is reversed. Older age, as well as identifying as Native American or Asian American suppress the explained difference. Only the difference in the proportion of people identifying as African American contributes positively to the explained difference. However, this difference is not statistically significant, so this explanation should just be used to understand how I present results, and should not be used to make inferences about the evidence. For the older two age groups, age contributes less to the explained difference in the older age categories than in the younger age categories, and the role of race statistically significantly accounts for more of the explained difference, which is a reversal from the pattern identified in the decomposition of the full sample. For the middle age category, age accounts for about two-thirds of the explained difference and race explains about one-third of the difference. In the oldest age category, age accounts for less than a quarter of the explained difference and race explains over three-quarters.

For one last example, the adjusted bar charts depict the relative role of each variable in the total explained difference, which allows for a more straightforward comparison across the models. Where in Table 7 the youngest age category is difficult to compare to the other categories, in the adjusted percentages (Table 8 and Figure 7), the relative role of the characteristics is more comparable. As noted before, if there are suppressors of an explained difference among the characteristics, it can cause other characteristics to appear to contribute more to the explanation than they do (such as age). However, by reporting them in relation to their relative role, a more accurate role of their contribution is revealed. For example, if I used the raw percentages of the role of age in relation to the explained total, it would be inflated for all three categories. By reporting the relative role, age is shown to still be important, but the contribution is no longer overstated. These example decompositions establish the foundation for the interpretation of the main decompositions—the decompositions of the difference in diabetes outcomes between the Latino and non-Latino groups including all of the variables included in the full model.

Decomposition of difference with all variables included in full model

Overview

The decomposition of the difference in the diabetes outcome of the Latino and non-Latino groups including all of the variables in the full logistic regression model highlights a few key findings that were not obvious from the logistic regression results on their own. This section will discuss the findings of the decomposition of the whole sample, and as age is a major suppressor of the difference in self-reported diabetes, I will use additional decompositions by age group to further illustrate the role of each of the measures in the explained difference. Table 9 shows the full results of the Oaxaca Blinder decomposition, from which I calculated the condensed versions, Tables 10 through 15. I will mainly refer to these tables as I discuss the decomposition findings.

The diabetes difference between the two groups remains the same for the whole sample and by age group, but the characteristics included in the full logistic regression model result in a greater *explained* proportion of these differences, compared with the previous examples. Figure 8 visually depicts these differences. Again, for the full sample, the proportion of Latinos with diabetes is about 9.2% and the proportion of non-Latinos with diabetes is about 8.5%— a difference of 0.7 percentage points. Based on the full model's logistic regression results, the difference explained is a statistically significant 0.0057, or about 80% of the difference, and the unexplained difference is 0.0015, or about 20% of the difference. This indicates that if Latinos had the same levels of the characteristics included in this decomposition as non-Latinos, the proportion of Latinos with diabetes would be 8.6% instead of 9.2%.

The difference in self-reported diabetes for individuals between the ages of 18 and 37 years is a statistically insignificant 0.24 percentage points. Both groups report less than 2% prevalence of diabetes in this age range. Of this difference, the model identifies 95.83% as explained and 4.17% as unexplained. However, as the difference is not large enough to be statistically significant, findings for this age group are not particularly informative. By the next age bracket (38 to 56 years old), a disparity clearly emerges: Latinos of this age range report a 10.2% prevalence of diabetes, while non-Latinos only report a 6.5% prevalence, which is a 3.7 percentage point difference. Of this difference, the decomposition indicates that 97.8% of the difference is explained and only 2.17% is unexplained. These findings are statistically significant. By the oldest age range (57 years and older), the diabetes difference has widened further, and over 25% of Latinos of this age report diabetes, compared to 16.8% of non-Latinos. This is a 9.5 percentage point gap. Differences in observable characteristics explain 93.9% of this difference, and 6.1% of the difference remains unexplained.

This pattern of how the observable characteristics contribute to the explained differences identified in each sample varies. The decompositions of the logistic regression equations by age group highlights how these characteristics wax and wane in influence over the lifespan. These findings are presented by percentage (both in direct relation to the explained difference and as a relative contributor to the net difference) in Tables 10 through 12. These findings are visually depicted in Figures 9 and 10.

Individual characteristics

Age. Age is the greatest suppressor of the diabetes disparity in the full model. In the decomposition of the full sample of adults, age is the major suppressor of the explained difference. Including all of the variables from the full logistic model, if Latinos had the same mean age as the non-Latino population, the proportion of Latinos with diabetes would increase by a statistically significant 2.3 percentage points. As the explained difference for this model is a net 0.0057, age accounts for -473.68% of this difference. This number, while unwieldy, simply captures how powerful of a suppressor age is. When the absolute values are taken of the characteristics contributing to the explained difference and summed, age accounts for about 32% of the breakdown of the explained difference by individual characteristics.

However, while Age Group 1 (18-37 years old) does not have a statistically different diabetes outcome (the difference of 0.0024 and the explained difference of 0.0024 are not statistically significant), there are individual characteristics that make a statistically significant contribution to the explained difference, including age. However, in this decomposition, age is not a suppressor. If Latinos between the ages of 18 and 37 had the identical age distribution as non-Latinos between the ages of 18 and 37, the proportion of Latinos reporting diabetes would decrease by 0.0010. This is a statistically significant difference at the 0.01 level. Age accounts for about 5.5% of the explained difference. By Age Group 2 (28-56 years old), the suppressor effect of age reemerges. Including the all of the variables from the full logistic model, if Latinos had the same mean age as the non-Latino population, the proportion of Latinos with diabetes would *increase* by 0.0071. This finding is statistically significant at the 0.001 level. Age accounts for about 8.5% of the explained difference. Across the decompositions of three age groups, age matters much less, which is what the subgroup analysis aims to accomplish. By the oldest age group, age only accounts for 2.61% of the relative explained difference. This indicates that the vast majority of the explained difference is accounted for by different characteristics, as well as that stratifying the analysis by age reduces the effect of the variable to the narrower ranges within the groups.

The younger mean age of the Latino population matters greatly in relation to the reported diabetes disparity. However, by conducting subgroup decompositions by age, we are able to more clearly see the other role of the other social determinants in the explained difference.

Sex. The difference in the distribution of sexes in the Latino and non-Latino groups does not statistically significantly contribute to the difference in reported diabetes in any of the models. The difference in distribution of the sexes accounts for less than 1% of the explained difference across all of the models. While not significant, this finding indicates that the role of sex is stable across all of the age groups.

Race. Like the role of sex, the contribution of race to the decomposition is fairly consistent across all of the samples. However, unlike the role of sex, the racial composition is statistically significant contributor to the explained difference in all of the decompositions. It is a suppressor in all of the models, which indicates that if Latinos had the same racial composition as non-Latinos, the reported prevalence of diabetes would increase. Race contributes about 5% to the relative explained difference in all of the models for the youngest group, where it contributes slightly more.

BMI. The effect of BMI is persistent and statistically significant across all of the models. It accounts for 6 to 10 percent of the explained difference in reported diabetes prevalence between Latinos and non-Latinos. In every model, if Latinos had the same mean BMI score as non-Latinos (which would be a lower score), they would report lower levels of diabetes. As the NHIS does not have measures for diet and exercise for the whole survey, we are not able to further decompose factors which contribute to the BMI difference. However, as BMI has emerged as an important factor, further investigation of the BMI difference is warranted.

Smoking. Differences in smoking matter (statistically speaking) in all of the models with the exception of the youngest age group. However, differences in smoking patterns contribute less than 2% to the total of the explained difference in each of the models.

In sum, differences in these individual characteristics contribute meaningfully to the Latino diabetes disparity. However, after accounting for age with the subgroup analysis, the net effect of the individual characteristics decline over age as well, and the net contribution of individual characteristics reaches a low of about 19% in Age Group 3 (and of that BMI accounts for the largest contributor).

Social Inequality

Factors of social inequality make meaningful and statistically significant contributions to the explained difference across all models. Education, income, and having health insurance coverage all influence the prevalence of diabetes. The net contribution of these characteristics to the explained difference ranges from 18.62% in the whole sample to 30.38% in the middle age group. If Latinos had the same levels of education and income as non-Latinos, they would have much lower rates of diabetes. That said, if Latinos had the same proportion with health insurance coverage as non-Latinos, they would actually have *higher* reported diabetes (due to selection bias—healthy, young--related to health insurance coverage).

Education. If Latinos had the same levels of education as non-Latinos, the proportion of Latinos with diabetes would decrease across all models. The portion of the explained difference that is related to education is statistically significant and fairly stable across the models: education contributes about 6% in the full model, 10% in the young age group, about 9% for both the middle and oldest age groups. In all models, the proportion of the effect that is attributable to education is on par with the proportion of the effect that is attributable to BMI. Latinos without a high school level of education comprise over a third, by far the largest segment, of the Latino population. This education disparity plays a very meaningful role in the Latino diabetes disparity.

Income. Similarly, almost double the proportion of Latinos have a household income lower than the federal poverty line compared to non-Latinos. This matters. If Latinos had the same levels of income as non-Latinos, the proportion of Latinos with diabetes would decrease. The contribution of income to the explained difference is statistically significant in all of the decompositions but the youngest age category. However, in the other three decompositions, income has a sizable contribution to the explained difference, very similar to the effect of BMI and education for the whole sample and the oldest age group. Noteworthy, for the middle age group (38-56), income comprises 15.67% of the explained difference in diabetes, twice the contribution of BMI.

Health Insurance Coverage. Health insurance coverage contributes a statistically significantly to the explained difference in all of the models. However, while Latinos are less likely to have health insurance coverage, the findings of the decomposition analysis indicate that if Latinos had the same rates of health insurance coverage as non-Latinos, they would report higher rates of diabetes. This is in part due to selection bias for who opts out of health insurance coverage: the healthy and the young, and also due to the possibility that those without insurance may have diabetes, but have not been diagnosed. As the Affordable Care Act continues to be implemented, the role of health insurance coverage may shift.

Social Ties

The measures for social ties did not have as sizeable a contribution to the explained difference. Also, the decomposition of social ties was only statistically significant for the full model and the oldest age group decompositions.

Number of individuals in household. The number of individuals in the household was only statistically significant for the oldest age group. For the older Latinos, the number of individuals in the household did help to explain the Latino diabetes disparity. If older Latinos were to have the same number of individuals in their households as non-Latinos (fewer people residing in each house), Latinos would report a *lower* prevalence of diabetes.

Marital status. For the whole sample, Latinos' proportions of being widowed or being never married statistically significantly influenced the diabetes disparity. If Latinos had the same proportion of widowed individuals as non-Latinos (which would be a *higher* percentage of widows and widowers), Latinos would report *lower* levels of diabetes. However, if Latinos had the same proportion of people who had never married (which would result in a lower proportion of individuals who had never married), Latinos would report *higher* levels of diabetes.

Acculturation

The measures of acculturation only contribute a statistically significant amount to the whole sample and the oldest age group. While there are effects reported for the other age groups, they are not statistically significant. For the full model, differences in acculturation measures account for about 10% of the sum of the absolute values of the contributors to the explained difference. These findings are not as consistent or robust as the findings for age, race, BMI, or social inequality measures.

Length of time in country. The only statistically significant contributors are the differences in distribution of those who have lived in the United States for more than 15 years, and this finding only holds true for the full sample and the oldest age group (who also happen to have had the greatest opportunity to have lived in the United States for greater than 15 years). In the full sample, if Latinos had the same distribution of foreign born individuals who have lived in the United States for more than 15 years as non-Latinos have, the proportion of Latinos with diabetes would decrease by 0.0045, or from

9.2% to about 8.7%. For Latino individuals over the age of 57, if they had the same proportion of individuals who had lived in the United States for greater than 15 years as non-Latinos, the reported diabetes prevalence would decrease from 26.3% to about 25%. There is no statistically significant portion of the explained difference that is attributable to the acculturation variables of nativity/length of time in country and language preference in the decomposition exploring the difference in diabetes outcomes in the younger and middle age group.

Language. Supporting the results identified in the logistic regression models, language does not have a meaningful effect on the Latino diabetes disparity. Language only makes a statistically significant contribution to the decomposition of the full sample. This finding indicates that if the portion of Latinos who took the survey in Spanish had taken the survey in English, the diabetes rate would be slightly lower (less than a 1% decrease). However, when the subgroup analysis by age is conducted, there is no statistically significant contribution of language to the explained difference in the diabetes disparity. This finding indicates that the diabetes prevalence is not strongly associated with language preference.

Latino Origin

A major contributor to the diabetes disparity is the Latino culture of origin. The inclusion of this variable in the decomposition brings into focus the heterogeneity of the Latino population. The different ethnic origins are associated with varying amounts of the explained difference. For the whole sample, the Age Group 2 (38-56 years) sample, and the Age Group 3 sample (57 years and older), the decomposition indicates that the diabetes disparity is meaningfully and significantly linked to Mexican and Puerto Rican heritage. For the whole sample, Latino origin accounts for about 23% of the relative contribution to the explained difference. For Age Group 2, Latino origin accounts for about 32% of the explained difference, and for Age Group 3, Latino origin accounts for about 34% of the explained difference. These findings indicate that the Latino diabetes disparity is largely driven by the prevalence of diabetes in the Mexican and Puerto Rican subpopulations. This finding is supported by other investigations of diabetes within the Latino community. However, what the decomposition does particularly well is it highlights the relative influence of range of Latino ethnic origins. This supports the case that generalizations made about Latino health must be made with significant caveats about the heterogeneity of the Latino population. However, the explained difference is not entirely attributable to differences in Latino origin, which also indicates that it is not only a Mexican or Puerto Rican diabetes problem, and that there are social determinants of health that are identified in this study on which actions can be based. Nonetheless, further investigations of the determinants of the diabetes disparity in these Latino subgroups are warranted by these findings.

Controls

Region. Region remains fairly stable across all four models, and is statistically significant in all but the youngest age group. Region contributes between 3.5 and 5.6% of the total contributions to the explained difference. This indicates that region of residence is not highly variable across these subsamples, and that region has a consistent association with diabetes. Region is a suppressor (mainly due to the fact that fewer

Latinos live in the North Central/Midwest region of the country). If Latinos lived in the same geographic regions as non-Latinos, they would report *higher* rates of diabetes. However, the size of the effect of region is considerably smaller, although statistically significant. The effect of region would widen the gap by no more than 0.5 percentage points in each sample. This suggests that while region targeted interventions may be appropriate, region would not be the most effective variable to address in efforts to decrease the Latino diabetes disparity.

Survey year. None of the explained difference in diabetes outcomes between Latinos and non-Latinos is significantly attributable to the survey year of the participants for any of the samples.

Oaxaca-Blinder Decomposition Hypotheses Revisited/Discussion

By analyzing the difference in diabetes outcomes for Latinos using Oaxaca-Blinder decomposition, I find additional support for the logistic regression analysis findings, as well as support for the claim that the disparity between Latino and non-Latino groups is largely due to social differences.

The findings analyzing the full models suggest that differences in the observable characteristics between the groups explain over 80% of the reported difference in diabetes outcomes, and after stratifying by age, this model explains over 90% of the difference in diabetes outcomes between the Latino and non-Latino groups.

The Oaxaca-Blinder decomposition analyses results support the rejection of the null hypotheses. The observable characteristics of the Latino population *do* inform the explained difference in the prevalence of self-reported diabetes between the Latino and the non-Latino population. In other words, if the Latino population had the same measurable characteristics as the non-Latino population, they would have a much lower rate of diabetes. Even after controlling for age, which is the largest suppressor of the difference, it is clear that measurable differences in BMI, social inequality, social ties, acculturation, and Latino ethnic origins all contribute, more or less, to the explained difference. Of the three theories, the greatest contributor to the explained difference is the differences in social inequality, but all three contribute to the explained difference in one or more age groups. However, a crucial finding is that the source of Latino origin contributes a sizeable portion of the explained difference in self-reported diabetes, which suggests that all generalizations based on the group of Latinos must be expressed cautiously.

The results of this decomposition indicate that while individual characteristics, especially age, race, and BMI, but to a lesser extent smoking as well, explain a sizeable portion of the explained difference, or the diabetes disparity, the role of the fundamental social determinant of health, social inequality, is as important. To a lesser extent, acculturation explains the difference, while social ties do not account for a meaningful amount of the difference in the whole model. However, social ties have their greatest importance in the older age group. Finally, the decomposition offers persuasive evidence that Latino ethnicity of origin is also a key contributor to the diabetes disparity. Many of the findings were very consistent across the decompositions, even though the complete model accounts for a sizable amount of the explained difference in each decomposition. However, just because this model has accounted for all of the *explained* difference does not mean that the disparity is solved. Unexplained differences remain, and as different factors have been emphasized, new black boxes have emerged. In the final chapter of the dissertation, I will discuss the implications of these findings, the limitations, and the next steps.

CHAPTER 7: CONCLUSION

This project set out to identify factors that contribute to the paradox within the paradox: what factors drive the diabetes disparity while simultaneously promoting unexpected health advantages in regard to mortality and other conditions for the Latino population? Both the logistic regression analyses and the Oaxaca-Blinder decomposition point to a few key characteristics as possible answers.

The statistical analyses of the association between identifying as Latino and self-reported diabetes accomplished two main goals. First, the logistic regression analyses identified the variables that are associated with the prevalence of diabetes in the Latino population. Second, the Oaxaca-Blinder decomposition analyses established which variables (and to what extent) contribute to the explained difference in self-reported diabetes between the two groups.

The logistic regression results indicate that variables that statistically significantly influence self-reported diabetes in the Latino population include some individual characteristics (age, BMI and origin of Latino ethnicity), all of the factors associated with social inequality (education, household income and health insurance coverage) and only one measure of acculturation (language). The evidence of this study suggests that acculturation and immigrant status is not as strongly associated with diabetes outcomes as they have been shown to be associated with other health outcomes (such as mortality and cancers), especially compared to other studies that measure nativity and acculturation in similarly cursory ways.

The Oaxaca-Blinder decomposition statistically depicts the counterfactual—if Latinos had the same measurable characteristics as non-Latinos, the diabetes disparity would be much smaller. The characteristic contributing the most to the difference between the populations in the whole sample is age, but it is acting as a suppressor. After adjusting for age through the subgroup analyses, the Oaxaca-Blinder decomposition results indicate that more than 90% of the difference for all three age groups is explained by the characteristics included in the decompositions. More factors emerge from the decomposition as explanatory pieces of the diabetes disparity than in the logistic regression models. Differences in almost all of the individual characteristics (age, race, BMI, smoking status, region of country, and origin of Latino ethnicity), all of the measures of social inequality (education, income and health insurance coverage), two measures of acculturation (being an immigrant and length of time in country), and one measure of social ties (marital status) statistically significantly contribute to the diabetes disparity between the Latino and non-Latino populations.

Ultimately, there is evidence to suggest that the differences in the acculturation effect and patterns of smoking behaviors may *not* be protective against diabetes, where the studies of the Latino paradox suggest that these factors *are* protective in regard to mortality, heart disease and certain types of cancer. Without this protective buffer, the disparities in education and income function the way that the theory predicts, those with fewer economic resources also have poorer health outcomes. However, socioeconomic factors do not explain the entire disparity.

One key pattern that emerged from these data and analyses is the diversity within the Latino population as well as between Latino and non-Latino populations, specifically, the statistically significant associations between the origin of Latino ethnicity and self-reported diabetes in both the logistic regressions and the Oaxaca-Blinder decomposition. While the origin of Latino ethnicity does not wholly explain the diabetes difference, the contribution of the various ethnic origins (particularly the role of Mexican or Puerto Rican heritage) suggests that in studies where a more detailed origin of Latino ethnicity can be included it should be. In studies investigating patterns related to Latino identity in datasets that do not include more detail about ethnicity of origin, conclusions about the Latino population must be made cautiously and should always include caveats about the heterogeneity of the Latino population. Of course, there is heterogeneity in the group of Latinos with Mexican origin that is also not explored, frequently because the data are limited, but also because studies of Latino health disparities often essentialize Latinos as a phenotype as opposed to a socially constructed category.

This leads to another limitation of this study, which is possible bias in the sampling strategy used to recruit Latino participants to this study. A nuanced conversation about decisions made regarding race and ethnicity in sampling strategies are often missing, particularly in secondary data analyses where the sampling and administration of the survey was done by others. While the researchers cannot change how the data were gathered and the sample was selected, they miss an opportunity to discuss both strengths and key limitations of their study by not making the story of their data transparent. In their 2009 book *Studying Ethnic Minority and Economically Disadvantaged Populations*, Knight, Roosa, and Umaña-Taylor discuss how the failure to sample populations representatively biases the analyses and limits generalizations and applications of the study (Knight, Roosa and Umaña-Taylor 2009).

While most large data sets, like the NHIS, employ a representative sampling strategy, the sampling strategy is representative at a preliminary level. The designs are based on the larger groups of identification, the fixed race categories. Which means that although they NHIS oversamples Latinos, they do not have a sampling design that is representative of the diversity of the Latino population. This sampling strategy may result in inaccurate findings particularly for the Latino population. The Latino population is so diverse that there is a lot of room for inaccurate representation to occur. The data may capture the right proportion of the Latino population in its schema, but after this level of representation is established, researchers need to investigate exactly who this group embodies.

Other health researchers of minority populations caution that these rich, nationally representative data sets are often not designed with a critical race and ethnicity mindset, but with assumptions that racial categories are enough to capture key differences between groups at best and that reify racial and ethnic categories as phenotypes at worst (Knight, Roosa and Umaña-Taylor 2009; Zuberi and Bonilla-Silva 2008b). When adding these arguments to the robust findings from my empirical investigations that Latino ethnicity of origin matters in regard to patterns of self-reported diabetes and in explaining the Latino diabetes disparity, the generalizability of these results to the Latino population must be made cautiously with the understanding that the Latino population is a compilation of many different groups of people from a wide range of cultural backgrounds.

Other limitations of this study include limitations of other measures, the need to further investigate how these variables influence each other, and the previously discussed cross-sectional design. The measures for acculturation and for social ties are both cursory, although many studies use these measures. For acculturation, the inclusion of immigrant status, length of time in country and language *do* capture patterns associated with being an immigrant. However, these measures do not capture to what degree an immigrant is acculturated (like the acculturation indexes would), nor do they capture a multidimensional understanding of acculturation (such as the theories of segmented assimilation call for). In order to gather evidence about why the association between acculturation and diabetes prevalence in the Latino population is not strong in the empirical investigations of this study, additional quantitative and qualitative work is warranted. First, a study investigating these patterns with data that include more detailed acculturation measures (such as the New Immigrant Survey) would give more insight into how acculturation specifically influences or does not influence the diabetes prevalence. Second, additional qualitative work exploring how acculturation processes (assuming that most immigrants experience assimilation and acculturation in segmented ways) influence health behaviors related to diabetes would also be warranted.

Also, additional work needs to be done investigating the role of "acculturation" and second and third generation immigrants. While there is no monolithic Latino culture, as there is evidence to suggest that specific origins of Latino ethnicity have different associations with diabetes prevalence, regardless of immigrant status, additional exploration of culture should be done.

The measures of social ties are also cursory. Only marital status and immediate household ties are accounted for, while other relationships (e.g. friends or family living separately from the individual) and other sources of social support (church, work relationships, and other social groups) cannot be measured using these data. Also, there are no measures of the quality of the social ties available in these data. Social ties may emerge as more influential with better measures.

One aspect of culture that transcends generations that is not investigated in this study is the cultural influence of food and diet. While the findings associated with immigrant status in these studies suggest that there is not a strong association between *immigrant* culture and diet, the findings associated with different origins of Latino heritage suggest that there may be diet differences between these groups that influence the diabetes disparity. The empirical results of this study set up future work that delves more deeply into these topics.

In a similar way that the origin of Latino heritage and immigrant status may be interrelated, many of the variables in this study are also interrelated. While the variables are not endogenous enough to pose multicollinearity problems statistically, the associations between income, education, acculturation, health insurance access, and BMI are not completely distinct from each other. Awareness of possible endogenous relationships will allow for further investigation of these dynamics, which can be done through additional empirical work of the interactions among these variables and also through empirical study of longitudinal data, which these data were not. The fact that all findings in this study come from cross-sectional data means that causality cannot be drawn from this project, but that these findings help crystalize the direction of future longitudinal work that can do more to establish causality.

Ultimately, however, by using widely accepted measures of concepts to conduct logistic regression analyses and Oaxaca-Blinder decompositions of the Latino diabetes disparity identified by previous literature, this project has identified the social determinants of health that are influencing both the likelihood of self-reported diabetes within the Latino population and the characteristics that contribute to the difference in diabetes outcomes between the Latino and non-Latino populations. These patterns associated with diabetes have not been included in discussions of the Latino paradox, possibly because the pattern of the paradox is not replicated in the patterns of diabetes (Lariscy, Hummer and Hayward 2015). Incorporating diabetes into the paradox complicates the story. However, if studies of the Latino paradox accounted for the diabetes disparity, the mortality health advantages may be even wider.

At the end of these analyses, these results set up the justification for additional analyses. First, replicating these analyses with a study that includes measures of diet, would allow for further insight into the association between acculturation and diabetes. This project was only able to measure acculturation with nativity, duration, and language, but the inclusion of culturally-salient diet questions may reveal more about its role. Also, as diabetes as a cause of death is increasing, an investigation of the social determinants of health associated with diabetes-specific mortality rates in the Latino population will demonstrate how this disparity is reflected in the final stages of life. Also, longitudinal studies are needed to gather evidence from which more causal conclusions can be drawn. The puzzle of the diabetes disparity needs to be answered as the consequences are so severe. However, as type 2 diabetes has great potential to be prevented and managed well, there is great promise that investigations of the social determinants related to the diabetes disparity can lead to improvements in longevity and quality of life.

WORKS CITED

2011. "Data file documentation, National Health Interview Survey, 2010." edited by National Center for Health Statistics. Hyattsville, MD: National Center for Health Statistics.

Abraido-Lanza, AF, A Cespedes, S Daya, and KR Florez. 2011. "Satisfaction with health care among Latinas." *J Health Care Poor Underserved* 22(2):491-505.

Abraido-Lanza, Ana F., Maria T. Chao, and Karen R. Florez. 2005. "Do healthy behaviors decline with greater acculturation?: Implications for the Latino mortality paradox." *Social Science & Medicine* 61(6):1243-55.

Afable-Munsuz, A., S.E. Gregorich, Kyriakos S Markides, and E.J. Perez-Stable. 2013a. "Diabetes risk in older Mexican Americans: effects of language acculturation, generation and socioeconomic status." *J Cross Cult Gerontol* 28(3):359-73.

Afable-Munsuz, Aimee, Elizabeth Rose Mayeda, Eliseo J. Perez-Stable, and Mary N. Haan. 2013b. "Immigrant Generation and Diabetes Risk Among Mexican Americans: The Sacramento Area Latino Study on Aging." *American Journal of Public Health* 103(5):e45-e52.

Ai, A.L., H.B. Appel, B. Huang, and K. Lee. 2012. "Overall health and healthcare utilization among Latino women in the United States." *J Womens Health* 21(8):878-85.

Ai, A.L., L.T. Noel, H.B. Appel, B. Huang, and W.E. Hefley. 2013. "Overall health and health care utilization among Latino American men in the United States." *Am J Mens Health* 7(1):6-17.

Ailinger, Rita L. 1989. "Functional Capacity of Hispanic Elderly Immigrants." *Journal of Applied Gerontology* 8(1):97-109.

Ailinger, Rita L, and Maria Recio Causey. 1993. "Home Health Service Utilization by Hispanic Elderly Immigrants: A Longitudinal Study." *Home Health Care Services Quarterly* 14(2/3):85-96.

Akresh, Ilana Redstone. 2007. "Dietary assimilation and health among Hispanic immigrants to the United States." *Journal of Health and Social Behavior* 48(4):404-17.

Almeida, Joanna, Beth E. Molnar, Ichiro Kawachi, and S.V. Subramanian. 2009. "Ethnicity and nativity status as determinants of perceived social support: Testing the concept of familism." *Social Science & Medicine* 68(10):1852-58.

Angel, Jacqueline L, RJ Angel, MP Aranda, and TP Miles. 2004. "Can the Family Still Cope?: Social Support and Health as Determinants of Nursing Home Use in the Older Mexican-Origin Population." *Journal of Aging and Health* 16:338-54.

Angel, R., and Peggy A. Thoits. 1987. "The impact of culture on the cognitive structure of illness." *Cult Med Psychiatry* 11(4):465-94.

Angel, RJ, JL Angel, C Diaz Venegas, and C Bonazzo. 2010. "Shorter stay, longer life: age at migration and mortality among the older Mexican-origin population." *J Aging Health* 22(7):914-31.

Berkman, Lisa F., Maria Melchior, Jean-Francois Chastang, Isabelle Niedhammer, Annette Leclerc, and Marcel Goldberg. 2004. "Social Integration and Mortality: A Prospective of French Employees of Electricity of France-Gas of France: The GAZEL Cohort." *American Journal of Epidemiology* 159(2):167-74.

Berkman, Lisa, Thomas Glass, Ian Brissette, and Teresa E. Seeman. 2000. "From social integration to health: Durkheim in the new millennium." *Social Science & Medicine* 51(6):843-57.

Berkman, Lisa, and S. Syme. 1979. "Social networks, host resistance, and mortality: A nine-year follow-up of Alameda County residents." *AJE* 109:186-204.

Black, SA. 1999. "Increased health burden associated with comorbid depression in older diabetic Mexican Americans. Results from the Hispanic Established Population for the Epidemiologic Study of the Elderly survey." *Diabetes Care* 22(1):56-64.

Black, SA, KS Markides, and TQ Miller. 1998. "Correlates of depressive symptomatology among older community-dwelling Mexican Americans: the Hispanic EPESE." *J Gerontol B Psychol Sci Soc Sci* 53(4):S198-208.

Blinder, A.S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8:436-55.

Bobak, Martin, Hynek Pikhart, Clyde Hertzman, Richard Rose, and Michael Marmot. 1998. "Socioeconomic factors, perceived control and self-reported health in Russia. A Cross-sectional survey." *Social Science & Medicine* 47(2):269-79.

Bonilla-Silva, Eduardo, and Tukufu Zuberi. 2008. "Toward a Definition of White Logic and White Methods." in *White Logic, White Methods: Racism and Methodology*, edited by Tukufu Zuberi and Eduardo Bonilla-Silva. Lanham: Rowman & Littlefield Publishers.

Borrell, LN, BS Menendez, and SP Joseph. 2011. "Racial/ethnic disparities on self-reported hypertension in New York City: examining disparities among Hispanic subgroups." *Ethn Dis* 21(4):429-36.

Bustamante, AV, and J Chen. 2012. "Health expenditures dynamics years of U.S. residence: analyzing spending disparities among Latinos by citizenship/nativity status." *Health Serv Res* 47(2):794-818.

Cacioppo, John T., and Louise C. Hawkley. 2003. "Social Isolation and Health, with an Emphasis on Underlying Mechanisms." *Perspectives in Biology and Medicine* 46(3):S39-S52.

CDC. 2011. "National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011.". Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

—. 2014. "National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014." Atlanta, GA: U.S. Department of Health and Human Services.

Center, Minnesota Population Center and State Health Access Data Assistance. 2012. "Integrated Health Interview Series: Version 5.0." Minneapolis: University of Minnesota.

Chaufan, C, M Davis, and S Constantino. 2011. "The twin epidemics of poverty and diabetes: understanding diabetes disparities in a low-income Latino and immigrant neighborhood." *J Community Health* 36(6):1032-43.

Chen, J., and J.A. Rizzo. 2010a. "Racial and ethnic disparities in use of psychotherapy: evidence from US national survey data." *Psychiatric Services* 61(4):364-72.

Chen, Jie, and John Rizzo. 2010b. "Racial and Ethnic Disparities in Use of Psychotherapy: Evidence from U.S. National Survey Data." *Psychiatric Services* 61(4):364-72.

Chesla, Catherine A., Lawrence Fisher, Marilyn M. Skaff, Joseph T. Mullan, Catherine L. Gillis, and Richard Kanter. 2003. "Family predictors of disease management over one year in Latino and European American patients with Type 2 Diabetes." *Family Process* 42(3):375-90.

Christakis, N.A., and J.H. Fowler. 2007. "The Spread of Obesity in a Large Social Network over 32 Years." *NEJM* 357:370-79.

Coffman, Maren J, Cynthia K Norton, and Lara Beene. 2012. "Diabetes symptoms, health literacy, and health care use in adult Latinos with diabetes risk factors." *Journal of Cultural Diversity* 19(1):4-9.

Cunningham, S. A., Julia D. Ruben, and K. M. Venkat Narayan. 2008. "Health of foreign-born people in the United States: A review." *Health & Place* 14(4):623-35.

Cunningham, S. A., Elizabeth Vaquera, and Jeanne L. Long. 2012. "Race, Ethnicity, and the Relevance of Obesity for Social Integration." *Ethn Dis* 22(3):317-23.

Dinwiddie, Gniesha Y., Ruth E. Zambrana, and Mary A. Garza. 2014. "Exploring Risk Factors in Latino Cardiovascular Disease: The Role of Education, Nativity, and Gender." *Am J Public Health* 104:1742-50.

Donini, Lorenzo M., Claudia Savina, and Carlo Cannella. 2003. "Eating Habits and Appetite Control in the Elderly: The Anorexia of Aging." *International Psychogeriatrics* 15(1):73-87.

Finch, Brian K, and WA Vega. 2003. "Acculturation, Stress, Social Support, and Self-Rated Health among Latinos in California." *Journal of Immigrant Health* 5(109-117).

Gallant, Mary P. 2003. "The Influence of Social Support on Chronic Illness Self-Management: A Review and Direction for Research." *Health Education and Behavior* 30(2):170-95.

Gany, F, JC Leng, G Winston, CH Tseng, A Qin, and J Changrani. 2011. "Diabetes care and language discordance among Chinese and Latino primary care patients." *Ethn Dis* 21(4):473-9.

Gelfand, DE, and BWK Yee. 1991. "Trends and Forces: Influence of Immigration, Migration, and Acculturation on the Fabric of Aging in America." *Generations* 15(4):7-10.

Gerst, Kerstin, Majd Al-Ghatrif, Holly A Beard, Rafael Samper-Ternent, and Kyriakos S Markides. 2010. "High depressive symptomalogy among older community-dwelling Mexican Americans: the impact of immigration." *Aging & Mental Health* 14(3):347-54.

Gleeson-Kreig, JoAnn, Henrietta Bernal, and Steve Woolley. 2002. "The Role of Social Support in the Self-Management of Diabetes Mellitus Among a Hispanic Population." *Public Health Nursing* 19(3):215-22.

Gómez, Laura E., and Nancy López (Eds.). 2013. *Mapping "Race": Critical Approaches to Health Disparities Research*. New Brunswick: Rutgers University Press.

Gonzalez, HM, M Ceballos, W Tarraf, BT West, ME Bowen, and WA Vega. 2009. "The Health of Older Mexican Americans in the Long Run." *American Journal of Public Health* 99(10):1879-85.

Gonzalez, HM, MN Haan, and L Hinton. 2001. "Acculturation and the prevalence of depression in older Mexican Americans: baseline results of the Sacramento Area Latino Study on Aging." *J Am Geriatr Soc* 49(7):948-53.

Gonzalez, Martha, Roseann Giarrusso, and Emiko Takagi. 2007. "Mexican Immigrant and Mexican American Elders: Life Satisfaction, Health, and Acculturation." Pp. 20 in *American Sociological Association*.

Grieco, E.M., and E.N. Trevelyan. 2010. "Place of Birth of the Foreign-Born Population: 2009. American Community Survey Reports." Department of Commerce Economics and Statistics Administration. U.S. CENSUS BUREAU.

House, James T. 2002. "Understanding Social Factors and Inequalities in Health: 20th Century Progress and 21st Centure Prospects." *Journal of Health and Social Behavior* 43(2):125-42.

Hu, J, KA Amirehsani, DC Wallace, and S Letvak. 2012. "The meaning of insulin to Hispanic immigrants with type 2 diabetes and their families." *Diabetes Educ* 38(2):263-70.

Huang, Cheng, Neil K. Mehta, Irma T. Elo, S. A. Cunningham, Rob Stephenson, David F. Williamson, and K. M. Venkat Narayan. 2011. "Region of Birth and Disability Among Recent U.S. Immigrants: Evidence from the 2000 Census." *Popul Res Policy Rev* 30:399-418.

Hubert, Helen B., John Snider, and Marilyn A. Winkleby. 2005. "Health status, health behaviors, and acculturation factors associated with overweight and obesity in Latinos from a community and agricultural labor camp survey." *Preventive Medicine* 40:642-51.

Hunt, KJ, ME Gonzalez, R Lopez, SM Haffner, MP Stern, and C Gonzalez-Villalpando. 2011. "Diabetes is more lethal in Mexicans and Mexican-Americans compared to Non-Hispanic whites." *Ann Epidemiol* 21(12):899-906.

Hunt, Linda M., Suzanne Schneider, and Brendon Comer. 2004. "Should "acculturation" be a variable in health research? A critical review of research on US Hispanics." *Social Science & Medicine* 59(5):973-86.

Jann, Ben. 2008. "The Blinder-Oaxaca decomposition for linear regression models." *The Stata Journal* 8(4):453-79.

Kaplan, Mark S, Nathalie Huguet, Jason T Newsom, and Bentson H McFarland. 2004. "The association between length of residence and obesity among Hispanic immigrants." *American Journal of Preventative Medicine* 27(4):323-26.

Kawachi, Ichiro, and Lisa Berkman. 2001. "Social ties and mental health." *Journal of Urban Health* 78(3):458-67.

Kim, G., CX Aguado Loi, DA Chiriboga, Y. Jang, P Parmelee, and RS Allen. 2011a. "Limited English proficiency as a barrier to mental health service use: a study of Latino and Asian immigrants with psychiatric disorders." *J Psychiatr Res* 45(1):104-10.

Kim, G., CB Worley, RS Allen, L Vinson, MR Crowther, P Parmelee, and DA Chiriboga. 2011b. "Vulnerability of older Latino and Asian immigrants with limited English proficiency." *J Am Geriatr Soc* 59(7):1246-52.

Kingston, R.S., and J.P. Smith. 1997. "Socioeconomic status and racial and ethnic differences in functional status associated with chronic diseases." *American Journal of Public Health* 87(5):805-10.

Knight, George P., Mark W. Roosa, and Adriana J. Umaña-Taylor. 2009. *Studying Ethnic Minority and Economically Disadvantaged Populations*. Washington, D.C.: American Psychological Association.

Koch, G.G., D.H. Freeman, and J.L. Freeman. 1975. "Strategies in the Multivariate Analysis of Data from Complex Surveys." *International Statistical Review* 43:59-78.

Koopman, Richelle J., Arch G. Mainous, III, Vanessa A. Diaz, and Mark E. Geesey. 2005. "Changes in Age at Diagnosis of Type 2 Diabetes Mellitus in the United States, 1988-2000." *Annals of Family Medicine* 3(1):60-63.

Langellier, Brent A., Jie Chen, A. Vargas Bustamante, Moira Inkelas, and Alexander N. Ortega. 2014. "Understanding health-care access and utilization disparities among Latino children in the United States." *Journal of Child Health Care*.

Lara, M., C. Gamboa, M. Iya Kahramanian, L.S. Morales, and David E. Hayes Bautista. 2005. "Acculturation and Latino Health in the United States: A Review of the Literature and its Sociopolitical Context." *Annual Review of Public Health* 26:367-97.

Lariscy, Joseph T., Robert A. Hummer, and Mark D. Hayward. 2015. "Hispanic Older Adult Mortality in the United States: New Estimates and an Assessment of Factors Shaping the Hispanic Paradox." *Demography* 52(1):1-14.

Lemeshow, Adina R., Laurie Fisher, Elizabeth Goodman, Ichiro Kawachi, Catherine S. Berkey, and G.A. Colditz. 2008. "Subjective Social Status in the School and Change in Adiposity in Female Adolescents: Findings from a Prospective Cohort Study." *Archives of Pediatric and Adolescent Medicine* 162(1):23-28.

Link, Bruce G., M. Northridge, J. Phelan, and M. Ganz. 1998. "Social Epidemiology and the Fundamental Cause Concept: On the Structuring of Effective Cancer Screens by Socioeconomic Status." *The Millbank Quarterly* 76:375-402.

Link, Bruce G., and J. Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* 35(Extra Issue):80-94.

MacDougall, Colin, Richard Cooke, Neville Owen, Kristyn Willson, and Adrian Bauman. 1997. "Relating Physical Activity to Health Status, Social Conections and Community Facilities." *Australian and New Zealand Journal of Public Health* 21(6):631-37.

Markides, K.S., and J. Coreil. 1986. "The health of Hispanics in the southwestern United States: an epidemiologic paradox." *Public Health Reports* 101(3):253-65.

Markides, KS. 2000. "Longitudinal Study of Mexican American Elderly Health." National Institute on Aging.

Markides, KS, JJ Salinas, and K Sheffield. 2009. "The Health of Older Immigrants." *Generations* 32(4):46-52.

Maty, SC, and KM Tippens. 2011. "Perceived and actual diabetes risk in the Chinese and Hispanic/Latino communities in Portland, OR, USA." *Diabet Med* 28(6):658-67.

Meadows, LM, WE Thurston, and C Melton. 2001. "Immigrant Women's Health." *Social Science & Medicine* 52(9):1451-58.

Morales, LS, M Lara, RS Kington, RO Valdez, and Jose J. Escarce. 2002. "Socioeconomic, cultural, and behavioral factors affecting Hispanic health outcomes." *J Health Care Poor Underserved* 13(4):477-503.

Mui, AC, and SY Kang. 2006. "Acculturation Stress and Depression among Asian Immigrant Elders." *Social Work* 51(3):243-55.

O'Donnell, Owen, Eddy van Doorslaer, Adam Wagstaff, and Magnus LIndelow (Eds.). 2008. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation*. Washington, D.C.: The World Bank.

Oakie, S. 2007. "Immigrants and Health Care: At the Intersection of Two Broken Systems." *New England Journal of Medicine* 357:525-29.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14(3):693-709.

Olshansky, SJ, DP Goldman, Y. Zheng, and JW Rowe. 2009. "Aging in America in the Twenty-First Century: Demographic Forecasts from the MacArthur Foundation Research Network on an Aging Society." *Millbank Quarterly* 87(4):842-62.

Oza-Frank, R., and S. A. Cunningham. 2010. "The weight of US residence among immigrants: a systematic review." *Obesity Reviews* 11(4):271-80.

Oza-Frank, R., and K. M. Venkat Narayan. 2008. "Masked heterogeneity in obesity between immigrant subgroups." *American Journal of Public Health* 98(6):967-68.

—. 2010. "Overweight and Diabetes Prevalence Among US Immigrants." *American Journal of Public Health* 100(4):661-68.

Palloni, A., and E. Arias. 2004. "Paradox lost: explaining the Hispanic adult mortality advantage." *Demography* 41(3):385-415.

Peterson, PG. 1999. "Gray Dawn: The Global Aging Crisis." *Foreign Affairs* 78(1):42-55.

Phelan, J., Bruce G. Link, Ana Diez-Roux, Ichiro Kawachi, and Bruce Levin. 2004. "Fundamental Causes of Social Inequalities in Mortality: A Test of the Theory." *Journal of Health and Social Behavior* 45:265-85.

Pineda Olvera, Anna E., Sunita M. Stewart, Linda Galindo, and Stephens Jacqualene. 2007. "Diabetes, Depression, and Metabolic Control in Latinas." *Cultural Diversity and Ethnic Minority Psychology* 13(3):225-31.

Portes, Alejandro, and Ruben G. Rumbaut. 2006. *Immigrant America: A Portrait*. Berkeley: University of California Press.

Powell, L.M., R. Wada, R.C. Krauss, and Y. Wang. 2012. "Ethnic disparities in adolescent body mass index in the United States: the role of parental socioeconomic status and economic contextual factors." *Social Science & Medicine* 75(3):469-76.

Quah, S.R. 2010. "Health and Culture." Pp. 27-46 in *The New Blackwell Companion to Medical Sociology*, edited by William C. Cockerham. London: Blackwell Publishing.

Rao, J.N.K., and A.J. Scott. 1981. "The Analysis of Categorical Data from Complex Surveys: Chi-Squared Tests for Goodness of Fit and Independence in Two-Way Tables." *Journal of the American Statistical Association* 76:221-30.

Read, JG, and MM Reynolds. 2012. "Gender differences in immigrant health: the case of Mexican and Middle Eastern immigrants." *J Health Soc Behav* 53(1):99-123.

Robbins, Jessica M., Viola Vaccarino, Heping Zhang, and Stanislav V. Kasl. 2005. "Socioeconomic status and diagnosed diabetes incidence." *Diabetes Research and Clinical Practice* 68(3):230-36.

Rogler, LH, DE Cortes, and RG Malgady. 1991. "Acculturation and mental health status among Hispanics: convergence and new directions for research." *American Psychology* 46(6):585-97.

Ross, C.E., and J. Mirowsky. 2010. "Gender and the Health Benefits of Education." *The Sociological Quarterly* 51(1):1-19.

Rotermann, M. 2011. "The impact of considering birthplace in analyses of immigrant health." *Health Rep* 22(4):37-43.

Ruggles S, JT Alexander, K Genadek, R Goeken, MB Schroeder, and M Sobek. 2010. "Integrated Public Use Microdata Series [Machine-readable database]." edited by University of Minnesota. Minneapolis.

Rumbaut, Ruben G., and John R. Weeks. 1996. "Unraveling a Public Health Enigma: Why do immigrants experience superior perinatal health outcomes?" *Research in the Sociology of Health Care* 13:335-88.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing "Neighborhood Effects": Social Processes and New Directions in Research." *Annual Review of Sociology* 28:443-78.

Sebelius, Kathleen. 2011. "Annual Update of the HHS Poverty Guidelines." edited by Office of the Secretary. Washington, D.C.: U.S. Deaprtment of Health & Human Services.

Seeman, Teresa E. 1996. "Social ties and health: The benefits of social integration." *Annals of Epidemiology* 6(5):442-51.

Sen, B. 2014. "Using the Oaxaca-Blinder decomposition as an empirical tool to analyze racial disparities in obesity." *Obesity* 22(7):1750-55.

Sherman, Justin J. 2005. "The Impact of Smoking and Quitting Smoking on Patients with Diabetes." *Diabetes Spectrum* 18(4):202-08.

Singh, G. K., and R.A. Hiatt. 2006. "Trends and disparities in socioeconomic and behavioral characteristics, life expectancy, and cause-specific mortality of native-born and foreign-born populations in the United States, 1979-2003." *Int J Epidemiology* 35(4):903-19.

Sinning, Mathias, Markus Hahn, and Thomas K. Bauer. 2008. "The Blinder-Oaxaca decomposition for nonlinear regression models." *The Stata Journal* 8(4):480-92.

Staff, Pew Social Trends. 2010. "The Return of the Multi-Generational Family Household." in *Pew Research Center*.

Stanfield, John H., III (Ed.). 2011. *Rethinking Race and Ethnicity in Research Methods*. Walnut Creek, CA: Left Coast Press, Inc.

Tak-Ying Shiu, Ann, Jo Yee-Mei Kwan, and Yee-Man Wong. 2003. "Social Stigma as a barrier to diabetes self-management: implications for multi-level interventions." *Journal of Clinical Nursing* 12(1):149-50.

Tang, Tricia S., Morton B. Brown, Martha M. Funnell, and Robert M. Anderson. 2008. "Social Support, Quality of Life, and Self-Care Behaviors among African Americans with Type 2 Diabetes." *Diabetes Educator* 34(2):266-76.

Thoits, Peggy A. 2011. "Mechanisms Linking Social Ties and Support to Physical and Mental Health." *Journal of Health and Social Behavior* 52(2):145-61.

Thomson, Maria D., and Laurie Hoffman-Goetz. 2009. "Defining and measuring acculturation: A systematic review of public health studies with Hispanic populations in the United States." *Social Science & Medicine* 69(7):983-91.

Tuchman, Arleen Marcia. 2011. "Diabetes and Race: A Historical Perspective." *Am J Public Health* 101(1):24-33.

Vigdor, J. 2008. *Measuring Immigrant Assimilation in the United States: Civic Report 53*. New York: Manhattan Institute for Policy Research.

Villa, VM, SP Wallace, S Bagdasaryan, and MP Aranda. 2012. "Hispanic Baby Boomers: health inequities likely to persist in old age." *Gerontologist* 52(2):166-76.

Viruell-Fuentes, EA, and AJ Schulz. 2009a. "Toward a Dynamic Conceptualization of Social Ties and Context: Implications for Understanding Immigrant and Latino Health." *American Journal of Public Health* 99(12):2167-75.

—. 2009b. "Toward a dynamic conceptualization of social ties and context: implictaions for understanding immigrant and Latino health." *American Journal of Public Health* 99(12):2167-75.

Wang, Youfa, and May A. Beydoun. 2007. "The Obesity Epidemic in the United States--Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis." *Epidemiologic Reviews* 29(1):6-28.

Wei, Ming, Braxton D. Mitchell, Steven M. Haffner, and Michael P. Stern. 1996. "Effects of Cigarette Smoking, Diabetes, High Cholesterol, and Hypertension on All-Cause Mortality and Cardiovascular Disease Mortality in Mexican Americans." *American Journal of Epidemiology* 144(11):1058-65.

Weiler, Dawn M, and Janice D Crist. 2009. "Diabetes Self-management in a Latino Social Environment." *The Diabetes Educator* 35(2):285-92.

Whitaker, Shanta M., Janice V. Bowie, Rachael McCleary, Darrell J. Gaskin, Thomas A. LaVeist, and Roland J. Thorpe, Jr. 2014. "The association between educational attainment and diabetes among men in the United States." *American Journal of Men's Health* 8(4):349-56.

Wilkinson, R.G. 1999. "Income inequality, social cohesion, and health: clarifying the theory. A reply to Muntaner and Lynch." *International Journal of Health Services* 29(3):525-43.

Willi, Carole, Patrick Bodenmann, William A. Ghali, Peter D. Faris, and Jacques Cornuz. 2007. "Active Smoking and the Risk of Type 2 Diabetes." *JAMA* 298(22):2654-64.

Williams, David R., and Michelle Sternthal. 2010. "Understanding Racial-Ethnic Disparities in Health: Sociological Contributions " *Journal of Health and Social Behavior* 51:S15-S27.

Yamada, AM, R Valle, and C Barrio. 2006. "Selecting an Acculturation Measure for Use with Latino Older Adults." *Research on Aging* 28(5):519-61.

Zhou, M. 1997. "Segmented Assimilation: Issues, Controversies, and Recent Research on the New Second Generation." *International Migration Review* 31(4):975-1008.

Zsembik, B.A., and D. Fennell. 2005. "Ethnic variation in health and the determinants of health among Latinos." *Social Science & Medicine* 61(1):53-63.

Zuberi, Tukufu, and Eduardo Bonilla-Silva. 2008a. "Telling the Real Tale of the Hunt: Toward a Race Conscious Sociology of Racial Stratification." Pp. 329-41 in *White Logic, White Methods: Racism and Methodology*, edited by Tukufu Zuberi and Eduardo Bonilla-Silva. Lanham: Rowman & Littlefield Publishers.

— (Eds.). 2008b. *White Logic, White Methods: Racism and Methodology*. Lanham: Rowman & Littlefield Publishers.

Table 1.1: Summary Statistics for adults over 18 years old, 2006-11 National Health Interview Survey (IHIS)

Variable %	Full Sample	Non-Latinos	Latinos
Latino	11.2%	0.0%	100.0%
Diabetes	8.5%	8.4%	9.2%
Age (years)	47.9	48.7	41.6
Sex (Female)	53.7%	54.0%	51.8%
BMI (kg/m^2)	27.1	27.1	27.8
People in Household (count)	2.4	2.3	3.1
Race			
White	82.2%	81.0%	91.6%
African-American	12.6%	13.7%	3.9%
Asian	3.9%	4.2%	1.5%
Native American	1.0%	0.8%	2.4%
Race other	0.4%	0.3%	0.6%
Smoking			
Smokes	20.6%	21.4%	14.8%
Quit	22.8%	23.7%	15.5%
Never smoked	56.4%	54.7%	69.5%
Unknown	0.2%	0.2%	0.2%
Education			
<HS	14.3%	11.6%	35.6%
High School Diploma	24.1%	24.3%	22.3%
GED	2.8%	2.8%	3.1%
Some College	20.1%	20.6%	16.0%
AA	10.3%	10.6%	7.9%
BA	18.3%	19.3%	10.0%
Grad Degree	9.7%	10.4%	4.1%
Unknown	0.5%	0.4%	0.9%
Household Income			
<FPL	13.1%	11.9%	22.3%
100%-199% FPL	15.6%	14.6%	23.6%
200%-399% FPL	25.7%	25.8%	24.6%
>400% FPL	32.3%	34.2%	16.9%
Unknown	13.4%	13.5%	12.6%
Health Insurance			
Covered	84.0%	86.4%	64.8%
Not Covered	15.7%	13.3%	34.8%
Unknown	0.3%	0.3%	0.4%
Marital Status			
Married-present	44.2%	44.1%	45.5%
Married-absent	1.4%	1.2%	2.9%
Widowed	9.7%	10.2%	5.2%
Divorced	15.0%	15.4%	11.9%
Separated	3.1%	2.8%	5.7%
Never married	26.2%	25.9%	28.5%
Unknown	0.3%	0.3%	0.2%

Table 1.2: Summary Statistics for adults over 18 years old, 2006-11 National Health Interview Survey (IHIS)

	Full Sample	Non-Latinos	Latinos
Variable %			
Years in United States			
Always in US	86.0%	91.5%	42.2%
<1 year	0.2%	0.1%	0.5%
1 to 5 years	1.4%	0.9%	4.7%
5 to 10 years	2.0%	1.1%	9.2%
10 to 15 years	1.9%	1.1%	8.9%
> 15 years	8.4%	5.2%	33.6%
Unknown	0.2%	0.1%	0.9%
Language of Interview			
English	95.7%	99.5%	65.3%
Spanish	2.4%	0.0%	21.6%
English & Spanish	1.5%	0.0%	13.0%
Other Language	0.4%	0.4%	0.1%
Unknown	0.0%	0.0%	0.0%
Latino Origin			
Non-Latino	88.8%	100.0%	0.0%
Mexican	6.6%	0.0%	59.3%
Puerto Rican	1.3%	0.0%	11.2%
Cuban	0.5%	0.0%	4.6%
Dominican	0.4%	0.0%	3.5%
Other Central/S. American	1.8%	0.0%	16.2%
Multiple Latino Origins	0.6%	0.0%	5.3%
Region			
Northeast	17.3%	17.7%	14.8%
North Central/Midwest	24.5%	26.4%	9.3%
South	36.5%	36.5%	37.0%
West	21.6%	19.4%	39.0%
Year			
2006	16.8%	16.9%	16.2%
2007	16.4%	16.4%	15.8%
2008	15.8%	15.9%	15.2%
2009	17.0%	17.0%	17.3%
2010	16.4%	16.3%	16.9%
2011	17.6%	17.5%	18.6%
Observations (n)	141,681	116,556	25,125

Table 2.1: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Latino Diabetes	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation. Diabetes	Latino Origin Diabetes	Controls Diabetes
Latino	1.0943**	1.6906***	1.6903***	1.8220***	1.7452***	1.7862***	1.5225***	1.5001***	1.4562***	1.4117**	1.4279***
	(0.0339)	(0.0535)	(0.0534)	(0.0581)	(0.0574)	(0.0591)	(0.0515)	(0.0521)	(0.0605)	(0.1489)	(0.1496)
Age		1.2251***	1.2237***	1.2248***	1.1838***	1.1795***	1.1964***	1.1990***	1.1970***	1.1972***	1.1973***
		(0.0072)	(0.0072)	(0.0072)	(0.0069)	(0.0068)	(0.0069)	(0.0070)	(0.0071)	(0.0071)	(0.0071)
Age^2		0.9987***	0.9987***	0.9987***	0.9990***	0.9991***	0.9989***	0.9989***	0.9989***	0.9989***	0.9989***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Female			0.8776***	0.8636***	0.8309***	0.8564***	0.8084***	0.8094***	0.8090***	0.8099***	0.8108***
			(0.0188)	(0.0185)	(0.0181)	(0.0188)	(0.0179)	(0.0184)	(0.0184)	(0.0185)	(0.0184)
Race (white)											
African-American				2.0355***	1.7563***	1.7738***	1.5688***	1.5645***	1.5628***	1.5680***	1.5264***
				(0.0518)	(0.0475)	(0.0482)	(0.0431)	(0.0443)	(0.0443)	(0.0447)	(0.0440)
Asian				1.2294***	1.7432***	1.8118***	1.8442***	1.8231***	1.7660***	1.7384***	1.7568***
				(0.0656)	(0.0949)	(0.0984)	(0.1008)	(0.1001)	(0.1040)	(0.1033)	(0.1070)
Native American				2.0075***	1.8487***	1.8210***	1.6626***	1.6581***	1.6584***	1.6459***	1.6620***
				(0.2464)	(0.2276)	(0.2276)	(0.2016)	(0.2000)	(0.2002)	(0.2011)	(0.2065)
Race other				1.0133	0.9907	0.9955	0.9372	0.9371	0.9294	0.9342	0.9579
				(0.1719)	(0.1769)	(0.1776)	(0.1689)	(0.1687)	(0.1670)	(0.1680)	(0.1720)
BMI					1.1270***	1.1274***	1.1216***	1.1215***	1.1215***	1.1214***	1.1212***
					(0.0023)	(0.0023)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
Smoking Status (Smokes)											
Quit Smoking						0.9659	1.0799*	1.0766*	1.0772*	1.0765*	1.0794*
						(0.0341)	(0.0388)	(0.0389)	(0.0391)	(0.0391)	(0.0393)
Never Smoked						0.8155***	0.9348*	0.9310*	0.9306*	0.9295*	0.9285*
						(0.0265)	(0.0307)	(0.0307)	(0.0309)	(0.0309)	(0.0309)
Unknown						1.1972	1.3528	1.4169	1.4184	1.4172	1.4173
						(0.2639)	(0.3054)	(0.3253)	(0.3243)	(0.3236)	(0.3248)

Table 2.2 - continued: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Latino Diabetes	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation Diabetes	Latino Origin Diabetes	Controls Diabetes
Education (<HS)											
High School Diploma							0.8242***	0.8282***	0.8307***	0.8383***	0.8450***
							(0.0279)	(0.0280)	(0.0283)	(0.0285)	(0.0289)
GED							0.9885	0.9920	0.9948	1.0011	0.9969
							(0.0645)	(0.0648)	(0.0652)	(0.0656)	(0.0653)
Some College							0.8054***	0.8114***	0.8135***	0.8212***	0.8234***
							(0.0304)	(0.0308)	(0.0312)	(0.0317)	(0.0319)
AA							0.7965***	0.8019***	0.8035***	0.8119***	0.8139***
							(0.0358)	(0.0362)	(0.0367)	(0.0373)	(0.0376)
BA							0.6270***	0.6314***	0.6333***	0.6410***	0.6431***
							(0.0262)	(0.0264)	(0.0267)	(0.0273)	(0.0275)
Graduate Degree							0.6396***	0.6447***	0.6469***	0.6536***	0.6574***
							(0.0316)	(0.0319)	(0.0323)	(0.0328)	(0.0331)
Unknown							0.7897	0.8110	0.8190	0.8277	0.8322
							(0.1121)	(0.1162)	(0.1170)	(0.1174)	(0.1177)
Income-FPL											
100%-199% FPL							0.8868**	0.8846**	0.8830**	0.8834**	0.8857**
							(0.0373)	(0.0372)	(0.0372)	(0.0373)	(0.0373)
200%-399% FPL							0.7316***	0.7298***	0.7283***	0.7289***	0.7325***
							(0.0298)	(0.0303)	(0.0303)	(0.0304)	(0.0307)
>400% FPL							0.5724***	0.5718***	0.5708***	0.5707***	0.5779***
							(0.0268)	(0.0276)	(0.0277)	(0.0278)	(0.0283)
Unknown							0.6855***	0.6867***	0.6865***	0.6864***	0.6979***
							(0.0319)	(0.0327)	(0.0328)	(0.0328)	(0.0334)
Health Insurance (covered)											
Not Covered							0.7092***	0.7092***	0.7149***	0.7131***	0.7011***
							(0.0276)	(0.0275)	(0.0285)	(0.0286)	(0.0284)
Unknown							0.6033	0.6064	0.6068	0.6038	0.6104
							(0.2162)	(0.2190)	(0.2203)	(0.2196)	(0.2236)

Table 2.3 - continued: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Latino Diabetes	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation. Diabetes	Latino Origin Diabetes	Controls Diabetes
People in Household								1.0220	1.0217	1.0199	1.0223
								(0.0116)	(0.0117)	(0.0117)	(0.0117)
Marital Status (married-present)											
Married-absent								1.0483	1.0591	1.0558	1.0654
								(0.0979)	(0.0999)	(0.0991)	(0.0998)
Widowed								1.0119	1.0132	1.0115	1.0175
								(0.0406)	(0.0407)	(0.0405)	(0.0407)
Divorced								0.9853	0.9852	0.9853	0.9926
								(0.0352)	(0.0352)	(0.0352)	(0.0356)
Separated								1.0949	1.0952	1.0965	1.1133
								(0.0691)	(0.0692)	(0.0695)	(0.0706)
Never married								1.0019	1.0017	1.0023	1.0253
								(0.0410)	(0.0410)	(0.0410)	(0.0422)
Unknown								0.4597**	0.4619**	0.4613**	0.4794**
								(0.1150)	(0.1156)	(0.1156)	(0.1199)
Years in US (Always)											
<1 year									0.6884	0.7139	0.7249
									(0.4217)	(0.4507)	(0.4557)
1 to 5 years									0.8948	0.9354	0.9548
									(0.1735)	(0.1811)	(0.1853)
5 to 10 years									0.9040	0.9429	0.9703
									(0.1112)	(0.1180)	(0.1215)
10 to 15 years									0.8269	0.8533	0.8739
									(0.0833)	(0.0864)	(0.0890)
> 15 years									1.1067*	1.1387**	1.1711***
									(0.0444)	(0.0467)	(0.0491)
Unknown									0.7727	0.7728	0.8028
									(0.2033)	(0.2024)	(0.2097)

Table 2.4 - continued: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Latino Diabetes	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation Diabetes	Latino Origin Diabetes	Controls Diabetes
Language of Interview (English)											
Spanish									0.9330	0.9906	0.9913
									(0.0686)	(0.0731)	(0.0729)
English & Spanish									1.1366	1.1496	1.1398
									(0.0950)	(0.0970)	(0.0962)
Other Language									1.0525	1.0364	1.0716
									(0.1622)	(0.1602)	(0.1640)
Unknown									0.3106	0.3032	0.2943
									(0.3250)	(0.3168)	(0.3131)
Latino Origin (Non-Latino)											
Mexican										1.0998	1.0774
										(0.1225)	(0.1189)
Puerto Rican										1.0808	1.1177
										(0.1423)	(0.1460)
Cuban										0.7754	0.7102*
										(0.1245)	(0.1137)
Dominican										0.8692	0.9283
										(0.1533)	(0.1687)
Other Central/S. American										0.7249*	0.7141*
										(0.1052)	(0.1035)
Region (Northeast)											
North Central/Midwest											1.1722***
											(0.0483)
South											1.2508***
											(0.0479)
West											1.0902*
											(0.0473)

Table 2.5 - continued: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Latino Diabetes	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation Diabetes	Latino Origin Diabetes	Controls Diabetes
Year (2006)											
2007											1.0020
											(0.0436)
2008											1.0245
											(0.0455)
2009											1.0865
											(0.0504)
2010											1.1153**
											(0.0463)
2011											1.0586
											(0.0414)
Constant	0.0922***	0.0001***	0.0001***	0.0001***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
	(0.0013)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	141,681	141,681	141,681	141,681	141,681	141,681	141,681	141,681	141,681	141,681	141,681

Table 3.1: Multivariate logistic regression models predicting diabetes, non-Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation Diabetes	Controls Diabetes
Age	1.2168***	1.2151***	1.2169***	1.1742***	1.1699***	1.1872***	1.1887***	1.1879***	1.1880***
	(0.0079)	(0.0079)	(0.0080)	(0.0076)	(0.0075)	(0.0075)	(0.0078)	(0.0078)	(0.0079)
Age^2	0.9987***	0.9987***	0.9987***	0.9991***	0.9991***	0.9990***	0.9989***	0.9990***	0.9990***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Female		0.8619***	0.8459***	0.8133***	0.8371***	0.7921***	0.7928***	0.7921***	0.7927***
		(0.0196)	(0.0192)	(0.0189)	(0.0195)	(0.0186)	(0.0191)	(0.0191)	(0.0190)
Race (white)									
African-American			2.0875***	1.7839***	1.8026***	1.5983***	1.5927***	1.5884***	1.5418***
			(0.0554)	(0.0506)	(0.0514)	(0.0469)	(0.0481)	(0.0482)	(0.0471)
Asian			1.2735***	1.8535***	1.9341***	1.9677***	1.9410***	1.8169***	1.8255***
			(0.0684)	(0.1015)	(0.1061)	(0.1079)	(0.1069)	(0.1160)	(0.1200)
Native American			2.5655***	2.3244***	2.2767***	2.0555***	2.0463***	2.0466***	2.0584***
			(0.3187)	(0.2993)	(0.2992)	(0.2629)	(0.2603)	(0.2603)	(0.2685)
Race other			0.9101	0.8862	0.8952	0.8231	0.8161	0.7997	0.8226
			(0.1906)	(0.1909)	(0.1927)	(0.1770)	(0.1758)	(0.1730)	(0.1769)
BMI				1.1289***	1.1294***	1.1235***	1.1234***	1.1236***	1.1236***
				(0.0025)	(0.0025)	(0.0026)	(0.0026)	(0.0026)	(0.0026)
Smoking Status (Smokes)									
Quit Smoking					0.9630	1.0801*	1.0777	1.0764	1.0795*
					(0.0361)	(0.0413)	(0.0414)	(0.0414)	(0.0417)
Never Smoked					0.8124***	0.9392	0.9366	0.9346	0.9329
					(0.0294)	(0.0344)	(0.0344)	(0.0344)	(0.0344)
Unknown					1.2495	1.3942	1.4597	1.4563	1.4521
					(0.2915)	(0.3379)	(0.3600)	(0.3575)	(0.3593)

Table 3.2 - continued: Multivariate logistic regression models predicting diabetes, non-Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation Diabetes	Controls Diabetes
Education (<HS)									
High School Diploma						0.8402***	0.8430***	0.8438***	0.8530***
						(0.0319)	(0.0321)	(0.0322)	(0.0327)
GED						1.0525	1.0539	1.0567	1.0546
						(0.0768)	(0.0768)	(0.0770)	(0.0767)
Some College						0.8037***	0.8084***	0.8090***	0.8123***
						(0.0337)	(0.0340)	(0.0341)	(0.0344)
AA						0.8053***	0.8092***	0.8079***	0.8115***
						(0.0394)	(0.0398)	(0.0398)	(0.0403)
BA						0.6448***	0.6489***	0.6482***	0.6521***
						(0.0297)	(0.0299)	(0.0299)	(0.0303)
Graduate Degree						0.6481***	0.6534***	0.6523***	0.6573***
						(0.0357)	(0.0360)	(0.0362)	(0.0365)
Unknown						0.7694	0.7910	0.7902	0.7948
						(0.1197)	(0.1241)	(0.1249)	(0.1251)
Income-FPL									
100%-199% FPL						0.9204	0.9172	0.9165	0.9189
						(0.0452)	(0.0451)	(0.0451)	(0.0452)
200%-399% FPL						0.7510***	0.7475***	0.7472***	0.7516***
						(0.0339)	(0.0343)	(0.0343)	(0.0346)
>400% FPL						0.5889***	0.5867***	0.5862***	0.5946***
						(0.0311)	(0.0320)	(0.0321)	(0.0327)
Unknown						0.6990***	0.6991***	0.6991***	0.7108***
						(0.0373)	(0.0381)	(0.0382)	(0.0389)
Health Insurance (covered)									
Not Covered						0.7208***	0.7231***	0.7234***	0.7119***
						(0.0342)	(0.0343)	(0.0345)	(0.0342)
Unknown						0.6535	0.6599	0.6667	0.6726
						(0.2578)	(0.2620)	(0.2649)	(0.2695)

Table 3.3 - continued: Multivariate logistic regression models predicting diabetes, non-Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation Diabetes	Controls Diabetes
People in Household							1.0217	1.0208	1.0235
							(0.0141)	(0.0142)	(0.0142)
Marital Status (married-present)									
Married-absent							1.0308	1.0282	1.0400
							(0.1120)	(0.1121)	(0.1129)
Widowed							1.0101	1.0093	1.0167
							(0.0443)	(0.0443)	(0.0446)
Divorced							0.9845	0.9834	0.9917
							(0.0392)	(0.0392)	(0.0396)
Separated							1.1360	1.1346	1.1532*
							(0.0810)	(0.0811)	(0.0826)
Never married							0.9777	0.9777	1.0025
							(0.0459)	(0.0459)	(0.0471)
Unknown							0.4674**	0.4672**	0.4866**
							(0.1273)	(0.1274)	(0.1323)
Years in US (<1 year)									
1 to 5 years								0.9654	0.9905
								(0.2666)	(0.2731)
5 to 10 years								1.0999	1.1472
								(0.1679)	(0.1753)
10 to 15 years								0.8731	0.9065
								(0.1303)	(0.1356)
> 15 years								1.1602**	1.2018***
								(0.0566)	(0.0597)
Unknown								0.8507	0.8956
								(0.3239)	(0.3402)

Table 3.4 - continued: Multivariate logistic regression models predicting diabetes, non-Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation Diabetes	Controls Diabetes
Language of Interview (English)									
Spanish								0.3262	0.3289
								(0.2752)	(0.2741)
English & Spanish								0.5203	0.4932
								(0.3992)	(0.3790)
Other Language								0.9804	1.0118
								(0.1567)	(0.1602)
Region (Northeast)									
North Central/Midwest									1.1744***
									(0.0506)
South									1.2656***
									(0.0513)
West									1.0986*
Year (2006)									
2007									0.9713
									(0.0464)
2008									1.0128
									(0.0488)
2009									1.0774
									(0.0541)
2010									1.1014*
									(0.0499)
2011									1.0414
									(0.0442)
Constant	0.0001***	0.0001***	0.0001***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	116,556	116,556	116,556	116,556	116,556	116,556	116,556	116,357	116,357

Table 4.1: Multivariate logistic regression models predicting diabetes, Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Accultur-ation Diabetes	Latino Origin Diabetes	Controls Diabetes
Age	1.2610***	1.2611***	1.2612***	1.2274***	1.2235***	1.2403***	1.2480***	1.2463***	1.2447***	1.2451***
	(0.0168)	(0.0168)	(0.0168)	(0.0167)	(0.0165)	(0.0167)	(0.0171)	(0.0173)	(0.0172)	(0.0171)
Age^2	0.9985***	0.9985***	0.9985***	0.9987***	0.9988***	0.9986***	0.9986***	0.9986***	0.9986***	0.9986***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Female		1.0041	1.0019	0.9633	1.0118	0.9437	0.9475	0.9462	0.9579	0.9613
		(0.0520)	(0.0518)	(0.0506)	(0.0526)	(0.0493)	(0.0514)	(0.0520)	(0.0529)	(0.0533)
Race (white)										
African-American			1.0945	1.1250	1.1248	1.0716	1.0637	1.0749	1.1307	1.1328
			(0.1497)	(0.1587)	(0.1587)	(0.1522)	(0.1519)	(0.1528)	(0.1601)	(0.1625)
Asian			0.6312	0.6523	0.6512	0.6464	0.6445	0.6448	0.6574	0.6723
			(0.1535)	(0.1585)	(0.1583)	(0.1605)	(0.1608)	(0.1615)	(0.1646)	(0.1682)
Native American			0.8622	0.8283	0.8272	0.8102	0.8091	0.8084	0.7866	0.7884
			(0.1449)	(0.1368)	(0.1359)	(0.1354)	(0.1347)	(0.1342)	(0.1337)	(0.1327)
Race other			1.2775	1.2641	1.2457	1.2851	1.3036	1.3157	1.3691	1.3512
			(0.3877)	(0.4245)	(0.4210)	(0.4494)	(0.4530)	(0.4545)	(0.4701)	(0.4723)
BMI				1.1131***	1.1125***	1.1068***	1.1066***	1.1046***	1.1033***	1.1032***
				(0.0054)	(0.0054)	(0.0055)	(0.0055)	(0.0054)	(0.0054)	(0.0054)
Smoking Status (Smokes)										
Quit Smoking					1.0079	1.0749	1.0777	1.0948	1.0961	1.0951
					(0.0911)	(0.0981)	(0.0974)	(0.1000)	(0.1001)	(0.1000)
Never Smoked					0.8203*	0.8606	0.8604	0.8803	0.8773	0.8764
					(0.0656)	(0.0719)	(0.0712)	(0.0763)	(0.0757)	(0.0753)
Unknown					0.9271	1.0880	1.1569	1.1727	1.1627	1.1875
					(0.4843)	(0.5718)	(0.6167)	(0.6201)	(0.6211)	(0.6166)

Table 4.2 - continued: Multivariate logistic regression models predicting diabetes, Latino adult sample, IHIS 2006-2011

Model	Age	Sex	Race	BMI	Smoking	SES	Social Ties	Accultur-ation	Latino Origin	Controls
Variable (omitted)	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes
Education (<HS)										
High School Diploma						0.7624***	0.7694***	0.7654***	0.8096**	0.8045**
						(0.0528)	(0.0527)	(0.0560)	(0.0607)	(0.0596)
GED						0.6220**	0.6252**	0.6066**	0.6246**	0.6159**
						(0.1007)	(0.1010)	(0.0979)	(0.1003)	(0.0988)
Some College						0.9373	0.9494	0.9361	1.0012	0.9970
						(0.0837)	(0.0843)	(0.0857)	(0.0928)	(0.0924)
AA						0.7857	0.7965	0.7937	0.8554	0.8507
						(0.0963)	(0.0979)	(0.0993)	(0.1068)	(0.1059)
BA						0.4695***	0.4736***	0.4832***	0.5374***	0.5278***
						(0.0554)	(0.0563)	(0.0585)	(0.0676)	(0.0663)
Graduate Degree						0.6121**	0.6125**	0.6281**	0.6942*	0.6864*
						(0.1019)	(0.1019)	(0.1061)	(0.1193)	(0.1179)
Unknown						0.8578	0.8827	0.9342	0.9852	0.9908
						(0.2514)	(0.2611)	(0.2750)	(0.2822)	(0.2802)
Income-FPL										
100%-199% FPL						0.7713***	0.7675***	0.7594***	0.7604***	0.7571***
						(0.0581)	(0.0582)	(0.0589)	(0.0598)	(0.0593)
200%-399% FPL						0.6730***	0.6719***	0.6597***	0.6639***	0.6627***
						(0.0518)	(0.0533)	(0.0536)	(0.0541)	(0.0550)
>400% FPL						0.5202***	0.5222***	0.5074***	0.5088***	0.5095***
						(0.0494)	(0.0522)	(0.0535)	(0.0542)	(0.0546)
Unknown						0.6724***	0.6716***	0.6655***	0.6652***	0.6852***
						(0.0613)	(0.0616)	(0.0617)	(0.0625)	(0.0645)
Health Insurance (covered)										
Not Covered						0.6710***	0.6654***	0.6937***	0.6887***	0.6805***
						(0.0500)	(0.0503)	(0.0543)	(0.0543)	(0.0542)
Unknown						0.3660	0.3572	0.3403	0.3255	0.3341
						(0.2285)	(0.2250)	(0.2188)	(0.2088)	(0.2183)

Table 4.3 - continued: Multivariate logistic regression models predicting diabetes, Latino adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation Diabetes	Latino Origin Diabetes	Controls Diabetes
People in Household							1.0179	1.0197	1.0137	1.0149
							(0.0192)	(0.0197)	(0.0197)	(0.0199)
Marital Status (married-present)										
Married-absent							1.1055	1.1626	1.1578	1.1403
							(0.2290)	(0.2396)	(0.2358)	(0.2317)
Widowed							1.0217	1.0332	1.0224	1.0203
							(0.1046)	(0.1052)	(0.1048)	(0.1046)
Divorced							0.9667	0.9601	0.9690	0.9704
							(0.0905)	(0.0900)	(0.0912)	(0.0913)
Separated							0.9010	0.9126	0.9217	0.9136
							(0.0959)	(0.0982)	(0.1000)	(0.0996)
Never married							1.1340	1.1330	1.1460	1.1442
							(0.1031)	(0.1032)	(0.1023)	(0.1019)
Unknown							0.3529*	0.3686	0.3642	0.3620
							(0.1866)	(0.1963)	(0.2008)	(0.2009)
Years in US (Always)										
<1 year								1.0555	1.1819	1.1573
								(0.6437)	(0.7483)	(0.7263)
1 to 5 years								0.7319	0.8429	0.8470
								(0.1754)	(0.1994)	(0.1997)
5 to 10 years								0.6702*	0.7619	0.7610
								(0.1330)	(0.1560)	(0.1550)
10 to 15 years								0.7187*	0.8066	0.7906
								(0.1064)	(0.1246)	(0.1219)
> 15 years								0.9137	1.0117	1.0113
								(0.0652)	(0.0774)	(0.0774)
Unknown								0.5918	0.6277	0.6483
								(0.2283)	(0.2391)	(0.2449)

Table 4.4 - continued: Multivariate logistic regression models predicting diabetes, Latino adult sample, IHIS 2006-2011

Model	Age	Sex	Race	BMI	Smoking	SES	Social Ties	Acculturation	Latino Origin	Controls
Variable (omitted)	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes	Diabetes
Language of Interview (English)										
Spanish								1.0027	1.0449	1.0640
								(0.0820)	(0.0844)	(0.0856)
English & Spanish								1.2305*	1.2319*	1.2353*
								(0.1134)	(0.1140)	(0.1145)
Other Language								2.5880	2.7372	2.6952
								(1.9030)	(2.2507)	(2.2764)
Unknown								0.9932	0.9555	0.8745
								(1.1207)	(1.1006)	(1.0622)
Latino Origin (Non-Latino)										
Mexican									1.1769	1.1530
									(0.1371)	(0.1311)
Puerto Rican									1.1335	1.1111
									(0.1530)	(0.1515)
Cuban									0.7977	0.7539
									(0.1407)	(0.1307)
Dominican									0.9085	0.9020
									(0.1646)	(0.1883)
Other Central/S. American									0.7803	0.7669
									(0.1200)	(0.1176)
Region (Northeast)										
North Central/Midwest										1.2509
										(0.2027)
South										1.0804
										(0.1319)
West										0.9775
										(0.1227)

Table 4.5 - continued: Multivariate logistic regression models predicting diabetes, full adult sample, IHIS 2006-2011

Model Variable (omitted)	Age Diabetes	Sex Diabetes	Race Diabetes	BMI Diabetes	Smoking Diabetes	SES Diabetes	Social Ties Diabetes	Acculturation Diabetes	Latino Origin Diabetes	Controls Diabetes
Year (2006)										
2007										1.2865*
										(0.1333)
2008										1.1446
										(0.1254)
2009										1.2079
										(0.1165)
2010										1.2815**
										(0.1167)
2011										1.2530*
										(0.1138)
Constant	0.0001***	0.0001***	0.0001***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	25,125	25,125	25,125	25,125	25,125	25,125	25,125	25,125	25,125	25,125

Table 5.1: Full Logistic Regression Model Results, Full, Non-Latino, & Latino Samples

Model Variable (omitted)	Full Sample Diabetes	Non-Latinos Diabetes	Latinos Diabetes
Latino	1.4279***		
	(0.1496)		
Age	1.1973***	1.1880***	1.2451***
	(0.0071)	(0.0079)	(0.0171)
Age^2	0.9989***	0.9990***	0.9986***
	(0.0000)	(0.0001)	(0.0001)
Female	0.8108***	0.7927***	0.9613
	(0.0184)	(0.0190)	(0.0533)
Race (white)			
African-American	1.5264***	1.5418***	1.1328
	(0.0440)	(0.0471)	(0.1625)
Asian	1.7568***	1.8255***	0.6723
	(0.1070)	(0.1200)	(0.1682)
Native American	1.6620***	2.0584***	0.7884
	(0.2065)	(0.2685)	(0.1327)
Race other	0.9579	0.8226	1.3512
	(0.1720)	(0.1769)	(0.4723)
BMI	1.1212***	1.1236***	1.1032***
	(0.0024)	(0.0026)	(0.0054)
Smoking Status (Smokes)			
Quit Smoking	1.0794*	1.0795*	1.0951
	(0.0393)	(0.0417)	(0.1000)
Never Smoked	0.9285*	0.9329	0.8764
	(0.0309)	(0.0344)	(0.0753)
Unknown	1.4173	1.4521	1.1875
	(0.3248)	(0.3593)	(0.6166)
Education (<HS)			
High School Diploma	0.8450***	0.8530***	0.8045**
	(0.0289)	(0.0327)	(0.0596)
GED	0.9969	1.0546	0.6159**
	(0.0653)	(0.0767)	(0.0988)
Some College	0.8234***	0.8123***	0.9970
	(0.0319)	(0.0344)	(0.0924)
AA	0.8139***	0.8115***	0.8507
	(0.0376)	(0.0403)	(0.1059)
BA	0.6431***	0.6521***	0.5278***
	(0.0275)	(0.0303)	(0.0663)
Graduate Degree	0.6574***	0.6573***	0.6864*
	(0.0331)	(0.0365)	(0.1179)
Unknown	0.8322	0.7948	0.9908
	(0.1177)	(0.1251)	(0.2802)

Table 5.2: Full, Non-Latino, & Latino Sample Results

Model Variable (omitted)	Full Sample Diabetes	Non-Latinos Diabetes	Latinos Diabetes
Income-FPL			
100%-199% FPL	0.8857**	0.9189	0.7571***
	(0.0373)	(0.0452)	(0.0593)
200%-399% FPL	0.7325***	0.7516***	0.6627***
	(0.0307)	(0.0346)	(0.0550)
>400% FPL	0.5779***	0.5946***	0.5095***
	(0.0283)	(0.0327)	(0.0546)
Unknown	0.6979***	0.7108***	0.6852***
	(0.0334)	(0.0389)	(0.0645)
Health Insurance (covered)			
Not Covered	0.7011***	0.7119***	0.6805***
	(0.0284)	(0.0342)	(0.0542)
Unknown	0.6104	0.6726	0.3341
	(0.2236)	(0.2695)	(0.2183)
People in Household	1.0223	1.0235	1.0149
	(0.0117)	(0.0142)	(0.0199)
Marital Status (married-present)			
Married-absent	1.0654	1.0400	1.1403
	(0.0998)	(0.1129)	(0.2317)
Widowed	1.0175	1.0167	1.0203
	(0.0407)	(0.0446)	(0.1046)
Divorced	0.9926	0.9917	0.9704
	(0.0356)	(0.0396)	(0.0913)
Separated	1.1133	1.1532*	0.9136
	(0.0706)	(0.0826)	(0.0996)
Never married	1.0253	1.0025	1.1442
	(0.0422)	(0.0471)	(0.1019)
Unknown	0.4794**	0.4866**	0.3620
	(0.1199)	(0.1323)	(0.2009)
Years in US (Always)			
<1 year	0.7249		1.1573
	(0.4557)		(0.7263)
1 to 5 years	0.9548	0.9905	0.8470
	(0.1853)	(0.2731)	(0.1997)
5 to 10 years	0.9703	1.1472	0.7610
	(0.1215)	(0.1753)	(0.1550)
10 to 15 years	0.8739	0.9065	0.7906
	(0.0890)	(0.1356)	(0.1219)
> 15 years	1.1711***	1.2018***	1.0113
	(0.0491)	(0.0597)	(0.0774)
Unknown	0.8028	0.8956	0.6483
	(0.2097)	(0.3402)	(0.2449)

Table 5.3: Full, Non-Latino, & Latino Sample Results

Model Variable (omitted)	Full Sample Diabetes	Non-Latinos Diabetes	Latinos Diabetes
Language of Interview (English)			
Spanish	0.9913	0.3289	1.0640
	(0.0729)	(0.2741)	(0.0856)
English & Spanish	1.1398	0.4932	1.2353*
	(0.0962)	(0.3790)	(0.1145)
Other Language	1.0716	1.0118	2.6952
	(0.1640)	(0.1602)	(2.2764)
Unknown	0.2943		0.8745
	(0.3131)		(1.0622)
Latino Origin (Non-Latino)			
Mexican	1.0774		1.1530
	(0.1189)		(0.1311)
Puerto Rican	1.1177		1.1111
	(0.1460)		(0.1515)
Cuban	0.7102*		0.7539
	(0.1137)		(0.1307)
Dominican	0.9283		0.9020
	(0.1687)		(0.1883)
Other Central/S. American	0.7141*		0.7669
	(0.1035)		(0.1176)
Region (Northeast)			
North Central/Midwest	1.1722***	1.1744***	1.2509
	(0.0483)	(0.0506)	(0.2027)
South	1.2508***	1.2656***	1.0804
	(0.0479)	(0.0513)	(0.1319)
West	1.0902*	1.0986*	0.9775
	(0.0473)	(0.0512)	(0.1227)
Year (2006)			
2007	1.0020	0.9713	1.2865*
	(0.0436)	(0.0464)	(0.1333)
2008	1.0245	1.0128	1.1446
	(0.0455)	(0.0488)	(0.1254)
2009	1.0865	1.0774	1.2079
	(0.0504)	(0.0541)	(0.1165)
2010	1.1153**	1.1014*	1.2815**
	(0.0463)	(0.0499)	(0.1167)
2011	1.0586	1.0414	1.2530*
	(0.0414)	(0.0442)	(0.1138)
Constant	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)
Observations	141,681	116,357	25,125

Table 6: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, BMI Example

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	0.0917***	0.0174***	0.1012***	0.2633***
Non-Latino	0.0845***	0.0150***	0.0645***	0.1682***
Difference	0.0072**	0.0023	0.0367***	0.0951***
Explained by BMI	0.0051***	0.0015***	0.0044***	0.0110***
Unexplained by BMI	0.0021	0.0008	0.0324***	0.0841***
Observations	141,681	48,200	48,267	45,214

Table 7: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Age & Race Example

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	0.0917***	0.0174***	0.1012***	0.2633***
Non-Latino	0.0845***	0.0150***	0.0645***	0.1682***
Difference	0.0072**	0.0023	0.0367***	0.0951***
Explained Difference (%)	-400***	-13.04	-25.89***	-9.67***
Unexplained Diff. (%)	500***	113.04	125.89***	109.67***
	%	%	%	%
Age	86.11***	-1533.33	65.26***	22.83***
Race				
African American	16.32***	2133.33	41.05***	85.87***
Asian	1.04**	-66.67	2.11	5.43**
Native American	-3.13***	-466.67	-9.47***	-14.13***
Race (Other)	0.00	0.00	0.00	-1.09
Observations (n)	141,681	48,200	48,267	45,214

Table 8: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Age & Race, (Absolute Value %)

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	0.0917***	0.0174***	0.1012***	0.2633***
Non-Latino	0.0845***	0.0150***	0.0645***	0.1682***
Difference	0.0072**	0.0023	0.0367***	0.0951***
Explained Difference (%)	-400***	-13.04	-25.89***	-9.67***
Unexplained Diff. (%)	500***	113.04	125.89***	109.67***
	%ABS	%ABS	%ABS	%ABS
Age	80.78***	36.51	55.36***	17.65***
Race				
African American	15.31***	50.79	34.82***	66.39***
Asian	0.98**	1.59	1.79	4.20**
Native American	2.93***	11.11	8.04***	10.92***
Race (Other)	0.00	0.00	0.00	0.84***
Observations (n)	141,681	48,200	48,267	45,214

Table 9.1: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model

VARIABLES	Whole Sample		Age Group 1		Age Group 2		Age Group 3	
Latino	0.0917***		0.0176***		0.1015***		0.2633***	
Non-Latino	0.0846***		0.0152***		0.0646***		0.1682***	
Difference	0.0071**		0.0024		0.0368***		0.0951***	
Explained €	0.0057*		0.0023		0.0360***		0.0893***	
Unexplained (U)	0.0015*		0.0001		0.0008		0.0058*	
Individual Characteristics	E	U	E	U	E	U	E	U
Age	-0.0270***	-0.0106	0.0010**	-0.0001	-0.0071***	-0.0691	-0.0036***	-0.0239
Female	0.0003***	-0.0022	-0.0001	0.0000	0.0001	-0.0030	0.0004	0.0144*
African American	-0.0032***	0.0003	-0.0006*	0.0001	-0.0026***	0.0005	-0.0051***	-0.0004
Asian	-0.0011***	0.0003	-0.0004*	0.0000	-0.0011***	0.0011	-0.0011***	-0.0016
Native American	0.0006**	0.0004	0.0003	0.0000	0.0004	0.0010	0.0009**	-0.0014*
Race (Other)	0.0000	-0.0001	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0003
BMI	0.0061***	0.0092	0.0019**	-0.0002	0.0062***	0.0256	0.0134***	-0.0707**
Quit Smoking	-0.0001	0.0000	0.0000	0.0000	0.0000	-0.0031	-0.0017**	-0.0087
Never Smoked	-0.0017***	0.0009	0.0000	0.0002	-0.0014**	-0.0064	0.0000	-0.0183
Unknown Smoking	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000
Subtotal	-0.0261	-0.0018	0.0021	0.0000	-0.0055	-0.0534	0.0033	-0.1103

Table 9.2: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model

VARIABLES	Whole Sample		Age Group 1		Age Group 2		Age Group 3	
Social Inequality	E	U	E	U	E	U	E	U
HS Diploma	0.0002**	0.0004	-0.0001	0.0000	0.0003*	0.0013	0.0019***	-0.0013
GED	0.0000	0.0003	0.0000	0.0000	0.0000	0.0013	0.0000	-0.0014
Some College	0.0006***	-0.0006	0.0001	0.0000	0.0006**	-0.0015	0.0021***	0.0017
AA	0.0003**	-0.0001	0.0000	0.0000	0.0008**	-0.0008	0.0006*	0.0001
BA	0.0028***	0.0005	0.0012*	0.0000	0.0033***	0.0011	0.0041***	-0.0024
Graduate Degree	0.0017***	0.0000	0.0005	0.0000	0.0023***	-0.0003	0.0036***	0.0000
Unknown Education	-0.0001	0.0000	0.0000	0.0000	-0.0001	-0.0001	-0.0002	0.0005
100-199% FPL	-0.0008**	0.0007	-0.0004	0.0001	-0.0013*	0.0028	-0.0004	-0.0018
200-399% FPL	0.0002**	0.0004	0.0003	0.0001	-0.0003	0.0009	0.0008*	-0.0030
>400% FPL	0.0056***	0.0004	0.0009	0.0001	0.0112***	0.0018	0.0101***	-0.0018
Unknown FPL	0.0002*	0.0000	-0.0001	0.0000	-0.0003	0.0000	0.0010**	0.0003
Not Covered	-0.0050***	-0.0001	-0.0013*	0.0000	-0.0049***	-0.0005	-0.0044***	-0.0034
Unknown Ins Coverage	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	-0.0004	-0.0002
Subtotal	0.0057	0.0019	0.0011	0.0003	0.0116	0.0061	0.0188	-0.0127
Social Ties	E	U	E	U	E	U	E	U
Number in Household	-0.0005	-0.0006	-0.0007	0.0002	-0.0003	-0.0097	0.0069***	-0.0197*
Married-absent	0.0000	-0.0001	0.0001	0.0000	-0.0001	-0.0006	0.0002	0.0000
Widowed	0.0008***	0.0001	0.0001	0.0000	0.0000	0.0003	-0.0007	-0.0013
Divorced	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0005	0.0000	-0.0020
Separated	0.0002	0.0002	0.0001	0.0000	-0.0001	-0.0006	0.0015**	-0.0043*
Never Married	-0.0003**	-0.0008	-0.0001	0.0001	-0.0001	-0.0017	0.0000	0.0015
Unknown	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000
Subtotal	0.0002	-0.0012	-0.0005	0.0003	-0.0004	-0.0128	0.0079	-0.0258

Table 9.3: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model

VARIABLES	Whole Sample		Age Group 1		Age Group 2		Age Group 3	
Acculturation	E	U	E	U	E	U	E	U
<1 year	-0.0002	-0.0001	0.0000	0.0000	-0.0001	0.0000	0.0003	0.0000
1 to 5 years	-0.0005	0.0000	-0.0002	0.0000	-0.0004	0.0008	0.0005	0.0006
5 to 10 years	-0.0006	0.0003	-0.0002	0.0000	-0.0001	0.0010	0.0001	-0.0009
10 to 15 years	-0.0009	0.0001	-0.0008	0.0000	-0.0009	-0.0002	0.0007	-0.0011
> 15 years	0.0045***	0.0006	0.0004	0.0000	0.0037	0.0010	0.0134***	-0.0145
Unknown years	-0.0001	0.0000	-0.0002	0.0000	0.0000	-0.0002	-0.0002	-0.0012
Spanish	0.0008*	-0.0001	-0.0001	0.0000	0.0008	-0.0006	-0.0013	0.0034
English & Spanish	0.0018	-0.0001	0.0012	0.0000	0.0007	-0.0002	0.0020	0.0017*
Other Language	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	-0.0001	0.0000
Unknown Language	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Subtotal	0.0048	0.0007	0.0001	0.0000	0.0037	0.0015	0.0154	-0.0120
Latino Origin	E	U	E	U	E	U	E	U
Mexican	0.0171***	0.0030	0.0022	0.0001	0.0198***	0.0078	0.0335***	-0.0195
Puerto Rican	0.0037***	0.0008	0.0000	0.0000	0.0043***	0.0019	0.0106***	-0.0062
Cuban	-0.0002	0.0003	0.0001	0.0000	0.0004	0.0007	-0.0009	-0.0026
Dominican	0.0007	0.0003	0.0001	0.0000	0.0007	0.0007	0.0014	-0.0020
Other Central/S. American	0.0002	0.0009	-0.0016	0.0000	0.0016	0.0028	-0.0001	-0.0045
Subtotal	0.0215	0.0053	0.0008	0.0001	0.0268	0.0139	0.0445	-0.0348

Table 9.4: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model

VARIABLES	Whole Sample		Age Group 1		Age Group 2		Age Group 3	
Controls	E	U	E	U	E	U	E	U
North Central/Midwest	-0.0019***	-0.0001	-0.0001	0.0000	-0.0025*	-0.0021	-0.0042**	-0.0029
South	0.0001	0.0015	0.0000	-0.0001	0.0000	0.0004	0.0008	-0.0133
West	0.0013*	0.0011	-0.0008	0.0000	0.0022*	0.0034	0.0022	-0.0044
2007	0.0000	-0.0009	0.0000	0.0000	0.0000	-0.0002	0.0000	0.0081
2008	0.0000	-0.0003	0.0000	0.0000	-0.0001	0.0016	0.0000	0.0053
2009	0.0000	-0.0004	0.0000	0.0000	0.0001	0.0001	0.0000	0.0048
2010	0.0000	-0.0005	0.0000	0.0000	0.0001	0.0008	0.0002	0.0062
2011	0.0000	-0.0006	0.0000	-0.0001	0.0000	0.0009	0.0002	0.0079
Subtotal	-0.0005	-0.0002	-0.0009	-0.0002	-0.0002	0.0049	-0.0008	0.0117
Constant	-0.0034		-0.0003		0.0405		0.1898	
Observations (n)	141,681	141,681	48,198	48,198	48,266	48,266	45,213	45,213

Table 10: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model, by Groups

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	**0.0917**	**0.0176**	**0.1015**	**0.2633**
Non-Latino	**0.0846**	**0.0152**	**0.0646**	**0.1682**
Difference	**0.0071**	0.0024	**0.0368**	**0.0951**
Explained difference	0.0057*	0.0023	0.0360***	0.0893***
Unexplained difference	0.0015*	0.0001	0.0008	0.0058*
Explained by Variable Group	E	E	E	E
Individual Characteristics				
Age	**-0.0270**	**0.0010**	**-0.0071**	**-0.0036**
Sex	0.0003	-0.0001	0.0001	0.0004
Race	**-0.0037**	**-0.0007**	**-0.0033**	**-0.0053**
BMI	**0.0061**	**0.0019**	**0.0062**	**0.0134**
Smoking	**-0.0018**	0.0000	**-0.0014**	**-0.0017**
Subtotal	**-0.0261**	**0.0021**	**-0.0055**	**0.0033**
Social Inequality				
Education	**0.0055**	**0.0017**	**0.0072**	**0.0121**
Income	**0.0052**	0.0007	**0.0093**	**0.0115**
Health Ins. Coverage	**-0.0050**	**-0.0013**	**-0.0049**	**-0.0048**
Subtotal	**0.0057**	**0.0011**	**0.0116**	**0.0188**
Social Ties				
Number in Household	-0.0005	-0.0007	-0.0003	**0.0069**
Marital Status	**0.0007**	0.0002	-0.0001	**0.001**
Subtotal	**0.0002**	-0.0005	-0.0004	**0.0079**
Acculturation (subtotal)				
Length of time in country	**0.0022**	-0.001	0.0025	**0.0148**
Language	**0.0026**	0.0011	0.0015	0.0006
Subtotal	**0.0048**	0.0001	0.0037	**0.0154**
Latino Origin (subtotal)				
Mexican	**0.0171**	0.0022	**0.0198**	**0.0335**
Puerto Rican	**0.0037**	0.0000	**0.0043**	**0.0106**
Cuban	-0.0002	0.0001	0.0004	-0.0009
Dominican	0.0007	0.0001	0.0007	0.0014
Other C./S. American	0.0002	-0.0016	0.0016	-0.0001
Subtotal	**0.0215**	0.0008	**0.0268**	**0.0445**
Controls (subtotal)				
Region	**-0.0005**	-0.0009	**-0.0003**	**-0.0012**
Survey Year	0.0000	0.0000	0.0001	0.0004
Subtotal	**-0.0005**	-0.0009	**-0.0002**	**-0.0008**
Observations	141,681	48,198	48,266	45,213

Results in bold are statistically significant at least at the 0.05 level.

Table 11: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model, by Groups
Raw results as %

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	**0.0917**	**0.0176**	**0.1015**	**0.2633**
Non-Latino	**0.0846**	**0.0152**	**0.0646**	**0.1682**
Difference	**0.0071**	0.0024	**0.0368**	**0.0951**
Explained difference (%)	**80.28**	95.83	**97.83**	**93.90**
Unexplained difference (%)	**21.13**	4.17	2.17	**6.10**
Explained by Variable Group	E	E	E	E
Individual Characteristics				
Age	**-473.68**	**43.48**	**-19.72**	**-4.03**
Sex	5.26	-4.35	0.28	0.45
Race	**-64.91**	**-30.43**	**-9.17**	**-5.94**
BMI	**107.02**	**82.61**	**17.22**	**15.01**
Smoking	**-31.58**	0.00	**-3.89**	**-1.90**
Subtotal	**-457.89**	**91.30**	**-15.28**	**3.70**
Social Inequality				
Education	**96.49**	**73.91**	**20**	**13.55**
Income	**91.23**	30.43	**25.83**	**12.88**
Health Ins. Coverage	**-87.72**	**-56.52**	**-13.61**	**-5.38**
Subtotal	**100.00**	**47.83**	**32.22**	**21.05**
Social Ties				
Number in Household	-8.77	-30.43	-0.83	**7.73**
Marital Status	**12.28**	8.70	-0.28	**1.12**
Subtotal	**3.51**	-21.74	-1.11	**8.85**
Acculturation (subtotal)				
Length of time in country	**38.60**	-43.48	6.94	**16.57**
Language	**45.61**	47.82	4.17	0.67
Subtotal	**84.21**	4.35	10.28	**17.25**
Latino Origin (subtotal)				
Mexican	**300.00**	95.65	**55.00**	**37.51**
Puerto Rican	**64.91**	0.00	**11.94**	**11.87**
Cuban	-3.51	4.35	1.11	-1.01
Dominican	12.28	4.35	1.94	1.56
Other C./S. American	3.51	-69.57	4.44	-0.11
Subtotal	**377.19**	34.78	**74.44**	**49.83**
Controls (subtotal)				
Region	**-8.77**	-39.13	**-0.83**	**-1.34**
Survey Year	0.00	0.0000	0.28	0.45
Subtotal	**-8.77**	-39.13	**-0.56**	**-0.90**
Observations	141,681	48,198	48,266	45,213

Results in **bold** are statistically significant at least at the 0.05 level.

Table 12: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model, by Groups
Adjusted % (ABS) (semi-condensed version)

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	**0.0917**	**0.0176**	**0.1015**	**0.2633**
Non-Latino	**0.0846**	**0.0152**	**0.0646**	**0.1682**
Difference	**0.0071**	0.0024	**0.0368**	**0.0951**
Explained difference (%)	**80.28**	95.83	**97.83**	**93.90**
Unexplained difference (%)	**21.13**	4.17	2.17	**6.10**
Explained by Variable Group	E (%)	E (%)	E (%)	E (%)
Individual Characteristics				
Age	**28.72**	**5.46**	**8.49**	**2.61**
Sex	0.32	0.55	0.12	0.29
Race	**5.21**	**7.10**	**4.90**	**5.15**
BMI	**6.49**	**10.38**	**7.42**	**9.72**
Smoking	**1.91**	0.00	**1.67**	**1.31**
Subtotal	**42.65**	**23.49**	**22.60**	**19.08**
Social Inequality				
Education	**6.06**	**10.38**	**8.85**	**9.06**
Income	**7.23**	9.29	**15.67**	**8.92**
Health Ins. Coverage	**5.32**	**7.10**	**5.86**	**3.48**
Subtotal	**18.61**	**26.77**	**30.38**	**21.46**
Social Ties				
Number in Household	0.53	3.83	0.36	**5.00**
Marital Status	**1.38**	2.19	0.60	**1.74**
Subtotal	**1.91**	6.02	0.96	**6.74**
Acculturation (subtotal)				
Length of time in country	**7.23**	9.84	6.22	**11.02**
Language	**2.77**	7.10	1.79	2.47
Subtotal	**10.00**	16.94	8.01	**13.49**
Latino Origin (subtotal)				
Mexican	**18.19**	12.02	**23.68**	**24.29**
Puerto Rican	**3.93**	0.00	**5.14**	**7.67**
Cuban	0.21	0.55	0.48	0.65
Dominican	0.74	0.55	0.84	1.02
Other C./S. American	0.21	8.74	1.91	0.07
Subtotal	**23.30**	21.86	**32.06**	**33.72**
Controls (subtotal)				
Region	**3.51**	4.92	**5.62**	**5.22**
Survey Year	0.00	0.00	0.36	0.29
Subtotal	**3.51**	4.92	**5.98**	**5.51**
Observations	141,681	48,198	48,266	45,213

Results in bold are statistically significant at least at the 0.05 level.

Table 13: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model (condensed version)

VARIABLES	Whole Sample	Age Group 1 18-37	Age Group 2 38-56	Age Group 3 57+
Latino	0.0917	0.0176	0.1015	0.2633
Non-Latino	0.0846	0.0152	0.0646	0.1682
Difference	0.0071	0.0024	0.0368	0.0951
Explained difference (%)	80.28	95.83	97.83	93.90
Unexplained difference (%)	21.13	4.17	2.17	6.10
% Explained by Variable Group	%	%	%	%
Individual Characteristics	**-0.0261**	**0.0021**	**-0.0055**	**0.0033**
Social Inequality	**0.0057**	**0.0011**	**0.0116**	**0.0188**
Social Ties	**0.0002**	-0.0005	-0.0004	**0.0079**
Acculturation	**0.0048**	0.0001	0.0037	**0.0154**
Latino Origin	0.0215	0.0008	**0.0268**	**0.0445**
Controls	**-0.0005**	-0.0009	**-0.0002**	**-0.0008**
Observations	141,681	48,198	48,266	45,213

Results in **bold** are statistically significant at least at the 0.05 level.

Table 14: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model, Raw % (Condensed)

VARIABLES	Whole Sample	Age Group 1 18-37	Age Group 2 38-56	Age Group 3 57+
Latino	0.0917	0.0176	0.1015	0.2633
Non-Latino	0.0846	0.0152	0.0646	0.1682
Difference	0.0071	0.0024	0.0368	0.0951
Explained difference (%)	80.28	95.83	97.83	93.90
Unexplained difference (%)	21.13	4.17	2.17	6.10
% Explained by Variable Group	%	%	%	%
Individual Characteristics	**-457.89**	**91.30**	**-15.28**	**3.70**
Social Inequality	**100.00**	**47.83**	**32.22**	**21.05**
Social Ties	**3.51**	-21.74	-1.11	**8.85**
Acculturation	**84.21**	4.35	10.28	**17.25**
Latino Origin	377.19	34.78	74.44	**49.83**
Controls	**-8.77**	-39.13	**-0.56**	**-0.90**
Observations	141,681	48,198	48,266	45,213

Results in bold are statistically significant at least at the 0.05 level.

Table 15: Oaxaca-Blinder Decomposition of Latino Diabetes Disparity, Full Model Absolute Value % (Condensed)

VARIABLES	Whole Sample	Age Group 1	Age Group 2	Age Group 3
Latino	0.0917	0.0176	0.1015	0.2633
Non-Latino	0.0846	0.0152	0.0646	0.1682
Difference	0.0071	0.0024	0.0368	0.0951
Explained difference (%)	80.28	95.83	97.83	93.90
Unexplained difference (%)	21.13	4.17	2.17	6.10
% Explained by Variable Group	%ABS	%ABS	%ABS	%ABS
Individual Characteristics	42.65	23.50	22.61	19.07
Social Inequality	18.62	26.78	30.38	21.46
Social Ties	1.91	6.01	0.96	6.74
Acculturation	10.00	16.94	8.01	13.49
Latino Origin	23.30	21.86	32.06	33.72
Controls	3.51	4.92	5.98	5.51
Observations (n)	141,681	48,198	48,266	45,213

Results in bold are statistically significant at least at the 0.05 level.

Figure 1: Overarching Research Question—How does being Latino relate to Self-Reported Diabetes?

Figure 2: Theoretical Framework

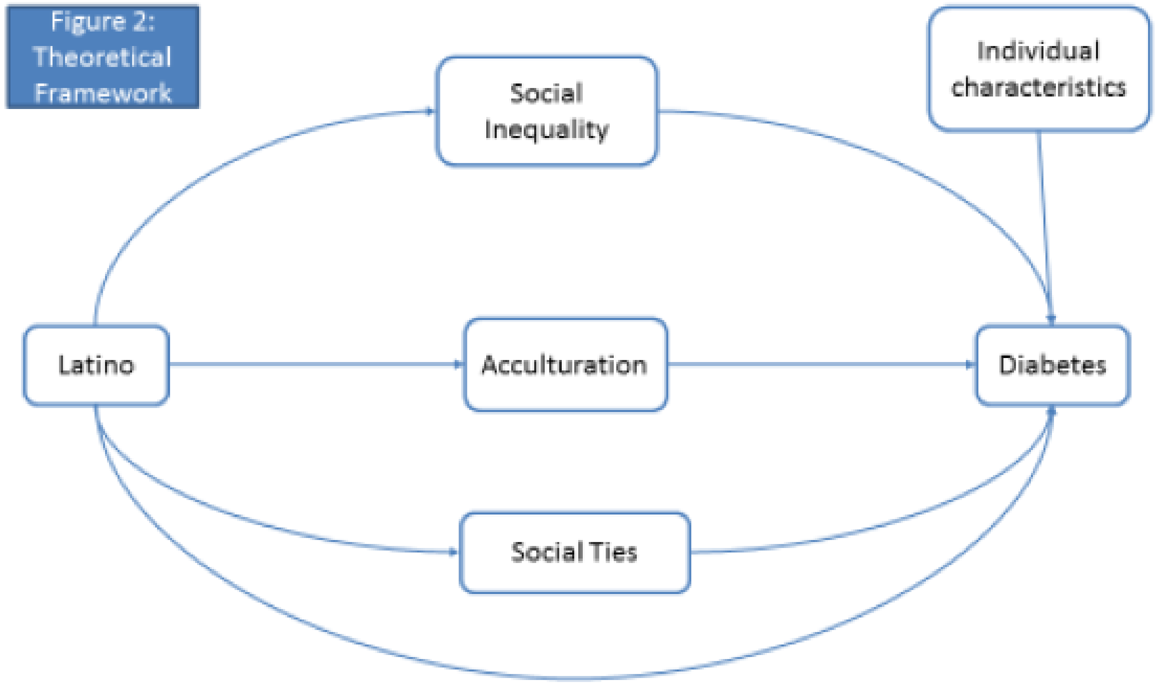

Figure 3: Theoretical Framework with Measures

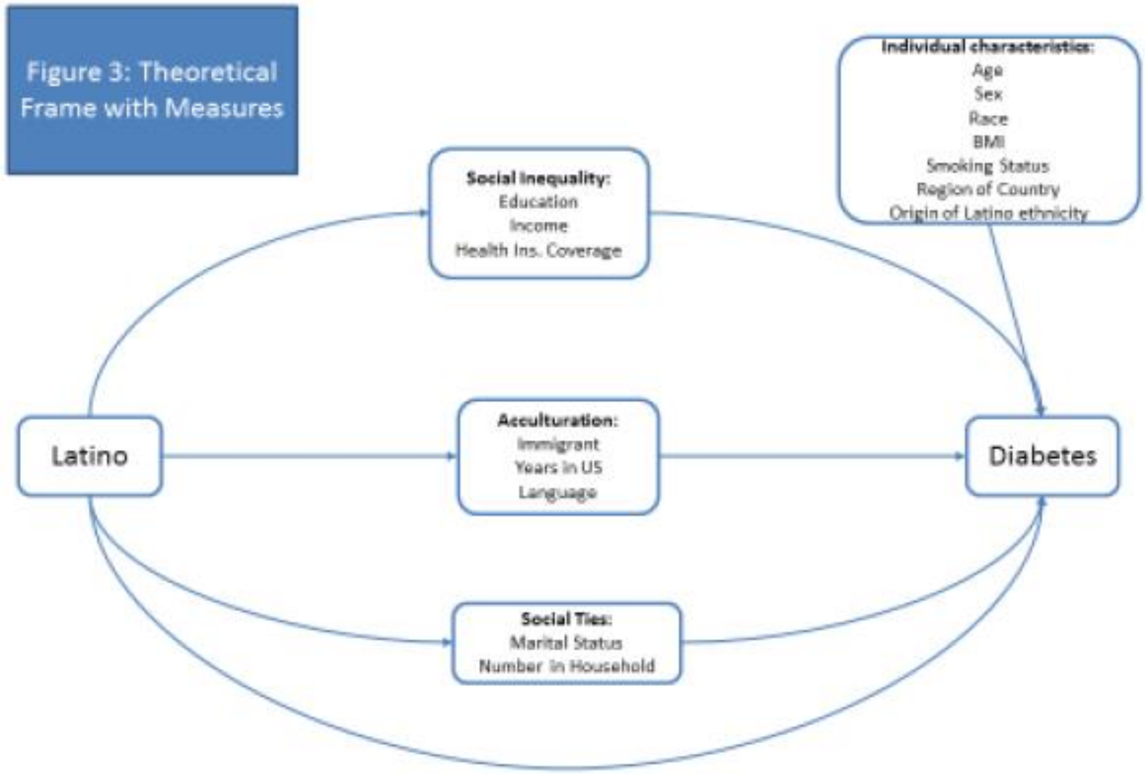

Figure 4: Theoretical Framework with Logistic Regression Results (Significant results in **bold**)

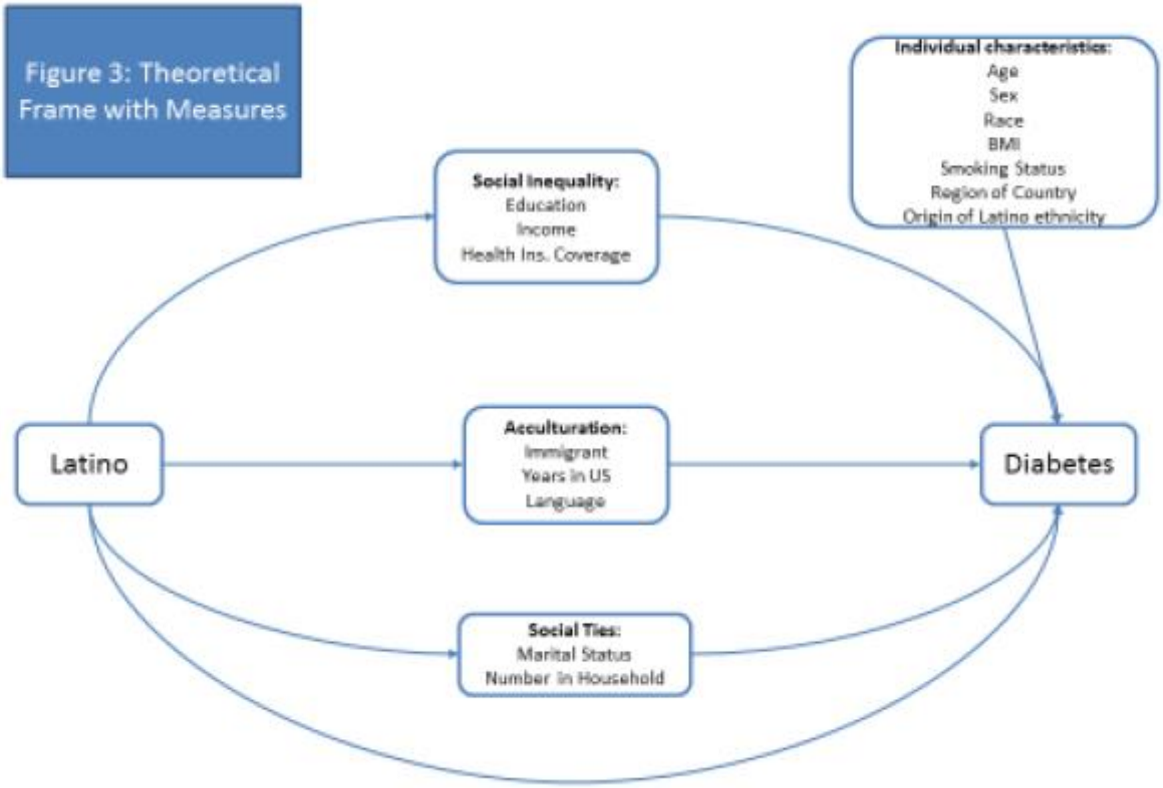

Figure 5: Oaxaca Decomposition of Diabetes Difference, Proportion Explained & Unexplained (Age & Race Only)

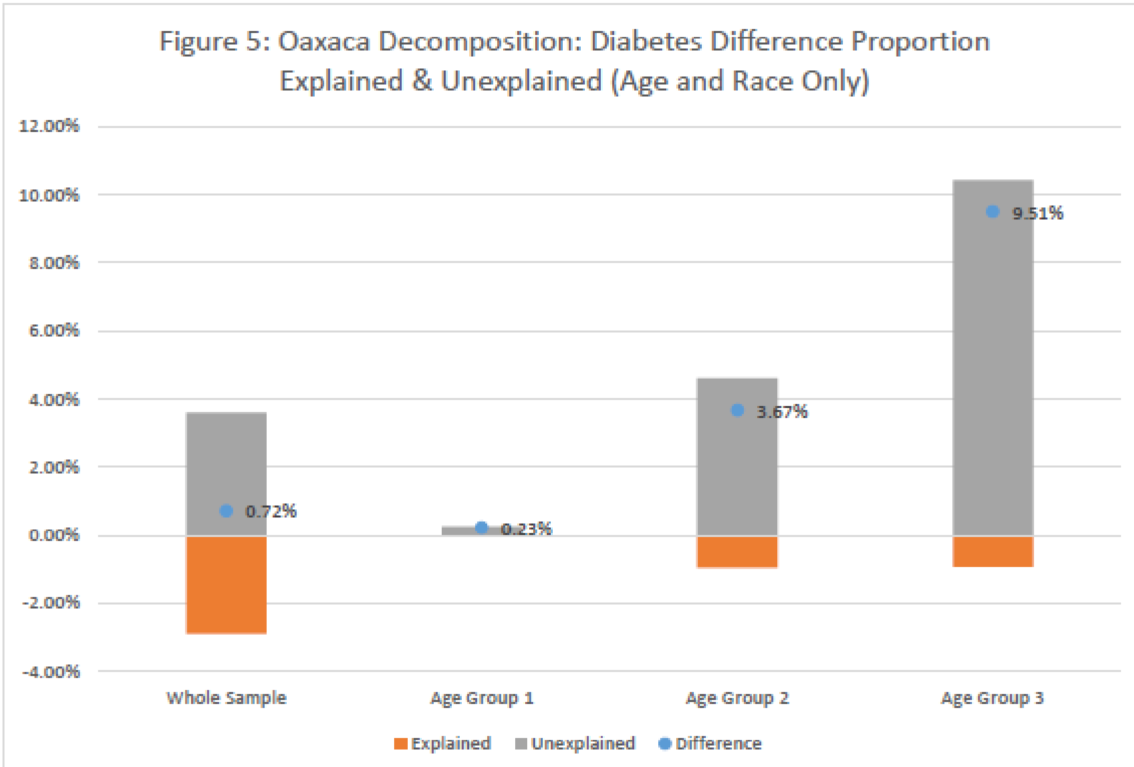

Figure 6: Oaxaca Decomposition of Diabetes Difference--Proportion Explained by Age & Race Only

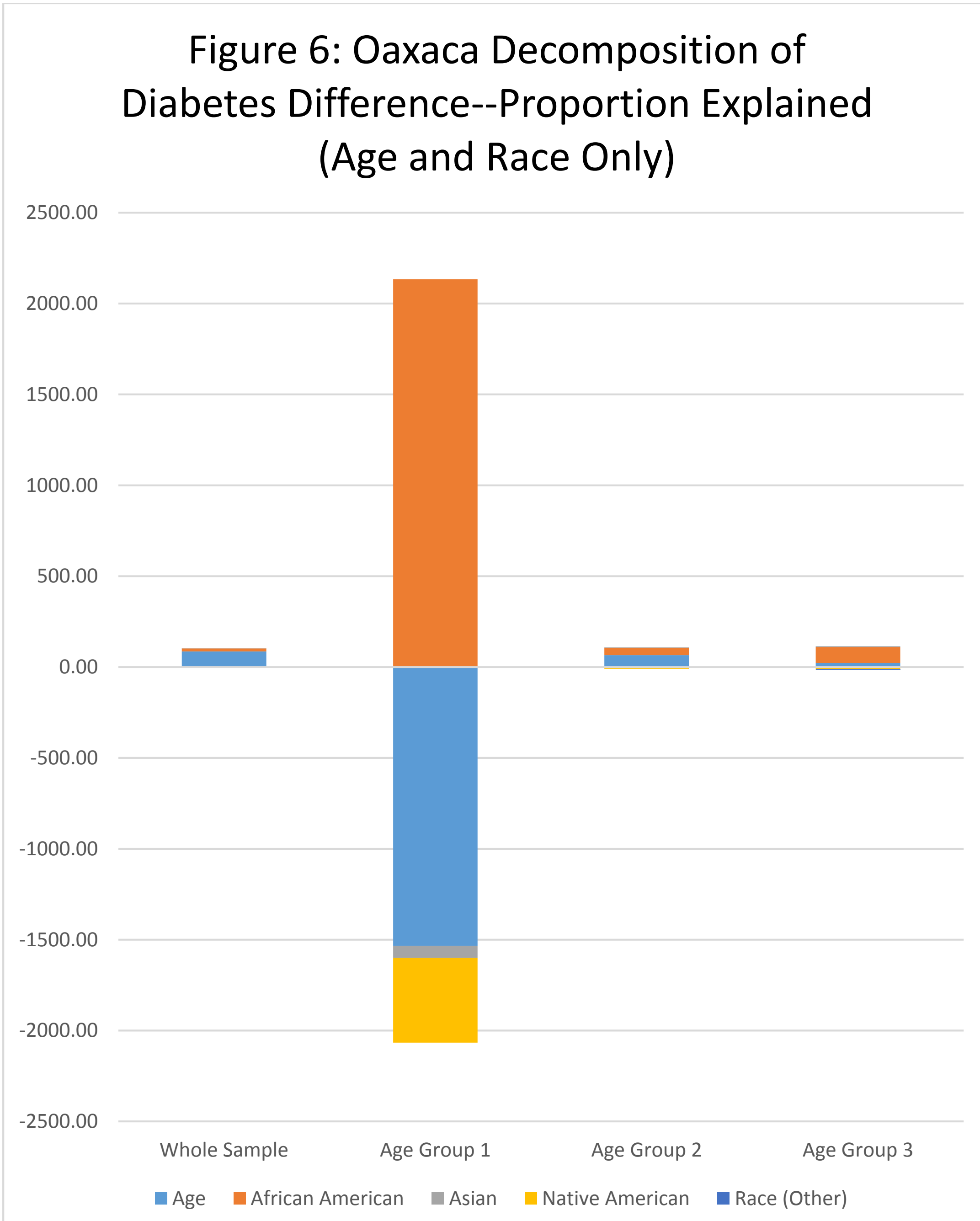

Figure 7: Oaxaca Decomposition of Diabetes Difference—Proportion Explained by Age & Race Only, Absolute Value Adjusted

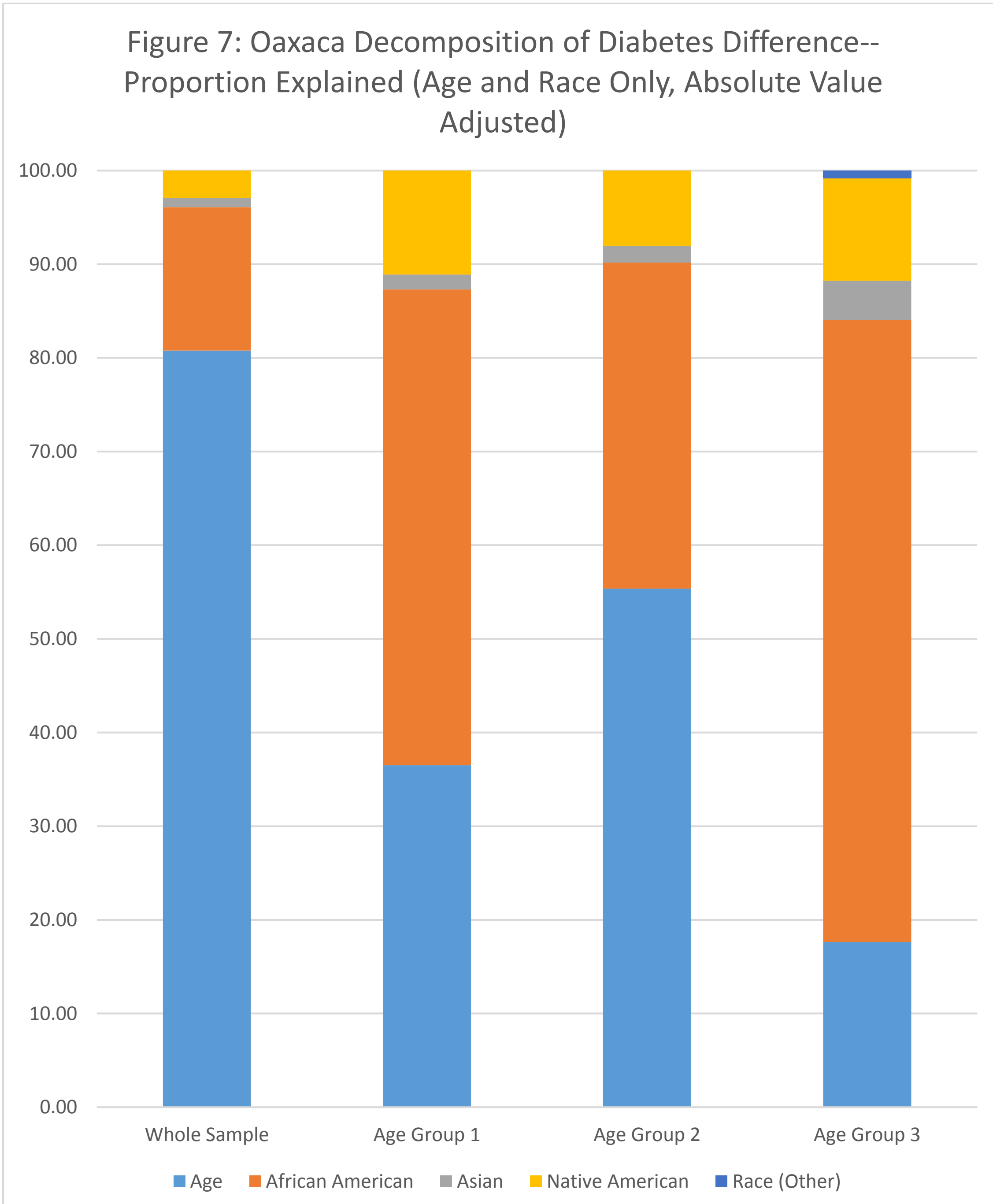

Figure 8: Oaxaca Decomposition of Diabetes Difference—Proportion Explained & Unexplained, Full Model

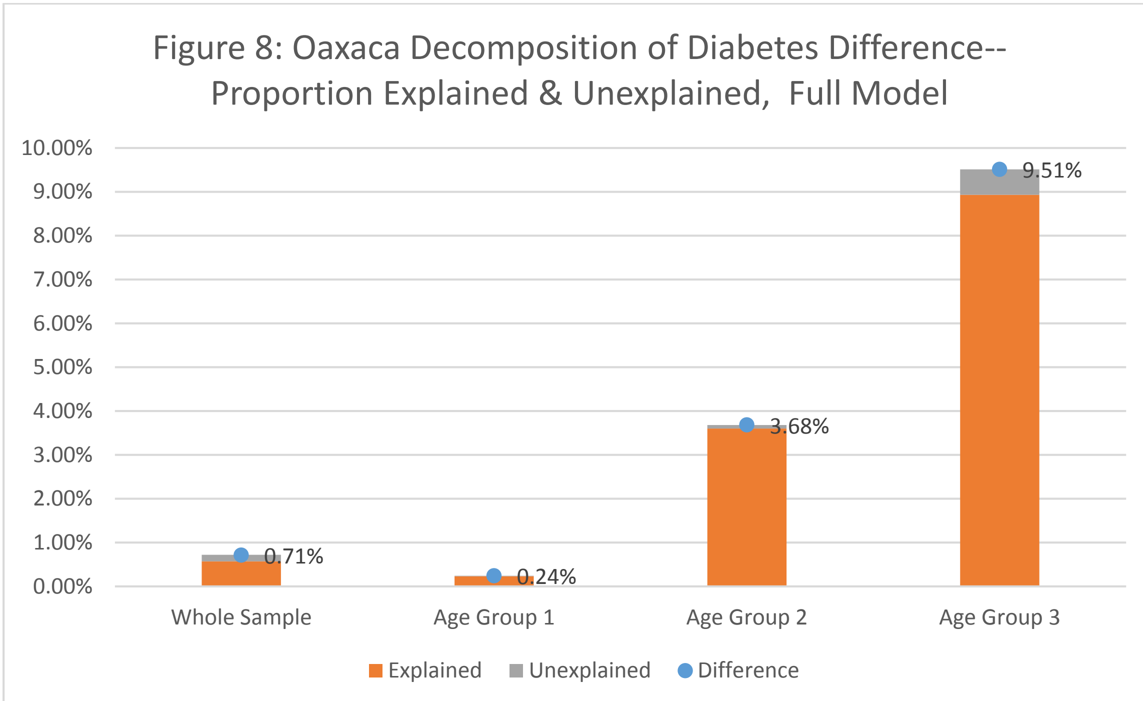

Figure 9: Oaxaca Decomposition of Diabetes Difference—Proportion Explained by Variables, Full Model

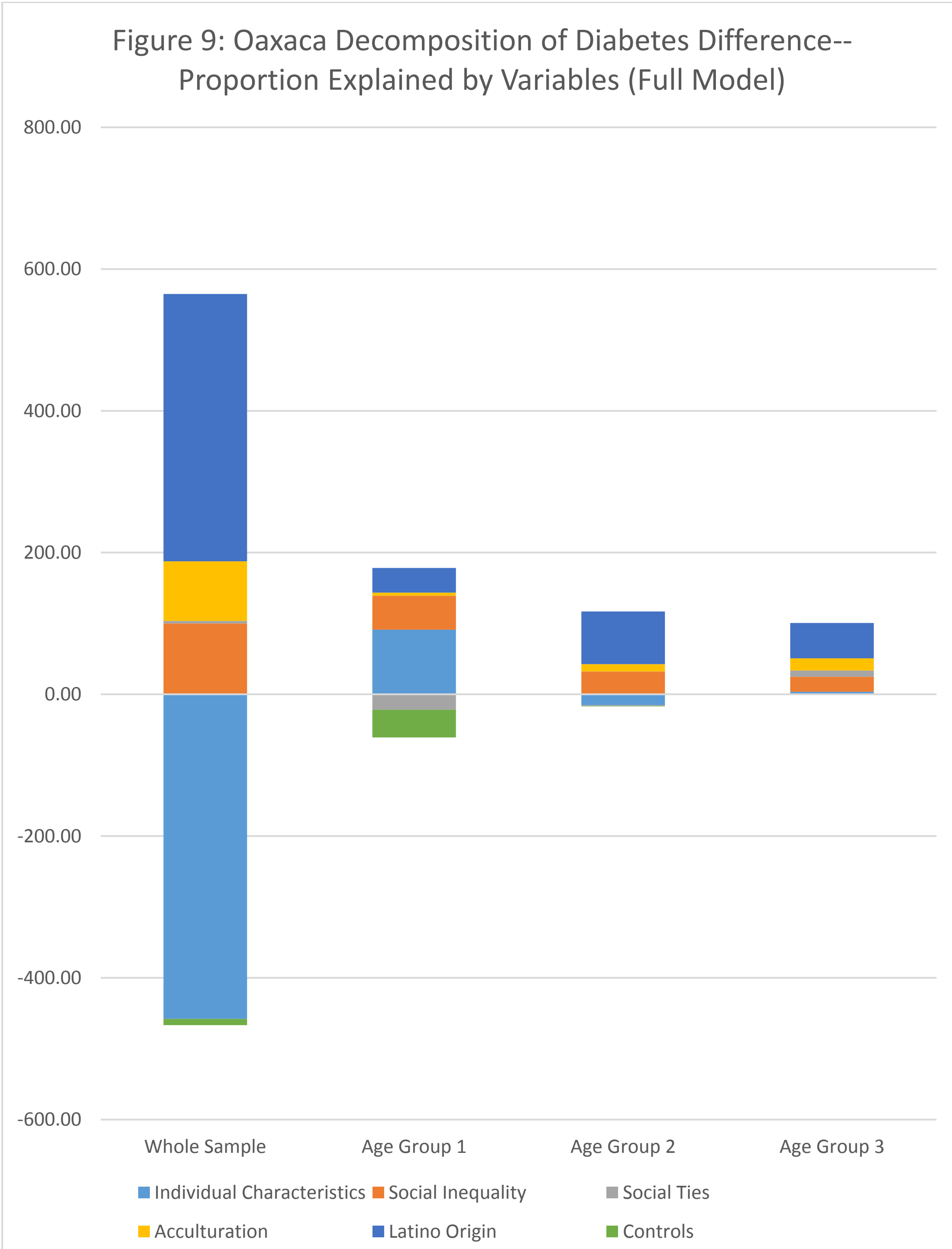

Figure 10: Oaxaca Decomposition of Diabetes Difference—Proportion Explained by Variables, Absolute Value Adjusted, Full Model

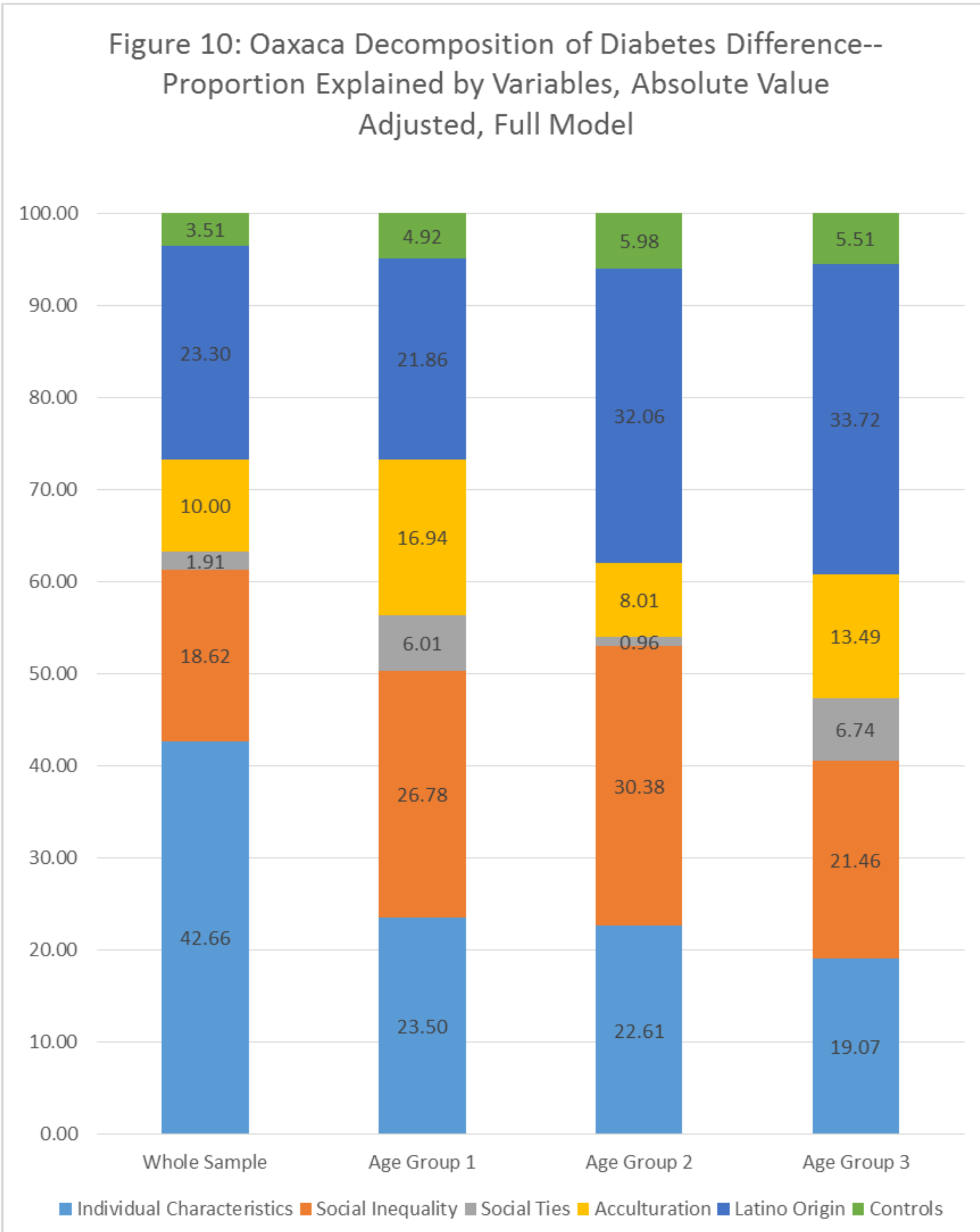

Figure 11: Theoretical Framework with Oaxaca-Blinder Decomposition Results (Full Model) (Significant Variables listed in **bold**)