Why Does the Health of Immigrants Deteriorate?

Osea Giuntella*

Department of Economics, Boston University (BU)

Draft

October 20, 2012

Abstract

Despite their socio-economic status, Hispanic immigrants in the US initially have better health outcomes than natives. However, while their socio-economic status improves over time and across generations, their health deteriorates ("Hispanic health paradox"). There is an open debate about whether the observed convergence is explained by selection on health or by the the adoption of less healthy lifestyles. This paper uses a unique dataset linking the birth records of two generations of Hispanics born in California and Florida (1975-2009), to analyze the mechanisms behind the generational decline in birth outcomes. Using country-level differences in health outcomes to pin down the degree of selection of the first-generation immigrants and existing estimates to pin down the intergenerational transmission of health status, I develop a simple model to interpret the health trajectories of immigrant descendants. Accounting for socio-economic differences between second-generation Hispanics and natives, the model not only explains the apparent paradox, but it over-predicts convergence with respect to the data. The paradox is reversed; contrary to the non-significant difference observed between third-generation immigrants and natives, the calibration exercise predicts a fairly large health advantage for natives. I show that the lower incidence of risk factors among Hispanics can explain 80% of the "reverse paradox". There is evidence of a generational worsening of risky behaviors and health conditions (e.g., smoking, alcohol consumption, and, hypertension), but second-generation Hispanics conserve a sizeable advantage compared to white natives. Among Hispanics, worse behaviors and higher assimilation are associated with poorer third-generation birth outcomes. This holds true even when focusing on a subset of second-generation siblings and controlling for grand-mother fixed effects.

---

*270 Bay State Road, 02215, Boston (MA); phone: 617-353-5616; fax: 617-353-4449; email: osea@bu.edu.
I am grateful to Daniele Paserman, Randall Ellis, Kevin Lang and Claudia Olivetti for their comments and advice. I would like to thank all the participants to the American Society of Health Economists Conference (Minneapolis, 2012), the XII Doctoral Workshop in Economic Theory and Econometrics (EIEF, 2012), the EconCon Conference (Princeton U, 2012), the European Meeting of the Econometric Society (Malaga, 2012), the XV IZA Summer School (Buch-Ammersee, 2012), the European Association of Labor Economists/IZA Conference (Bonn, 2012), as well as all seminar attendees at the BC-BU Green-Line Labor Meeting (2012), the Boston University Empirical Micro-Lunch(2012). Any errors are my own. The project was made possible by generous funding from the Boston University Institute for Economic Development.

Keywords: Hispanic health paradox, Birth Outcomes, Siblings Analysis

JEL Classification Numbers: I10, J15

1 Introduction

A substantial body of research has documented that immigrants are relatively healthier than natives when they arrive to the US, but that this initial advantage deteriorates over time spent in the US and across generations. These findings are particularly striking when focusing on Hispanics. Since they are characterized by lower socio-economic status than natives, one would expect them to be at higher risk of negative health outcomes. Furthermore, despite the positive socio-economic assimilation and the positive socio-economic gradient in health, there is evidence of a downward convergence in health over time and across generations. For these reasons, previous scholars referred to these stylized facts as the "Hispanic health paradox". This apparent paradox has been observed in general health status, life-expectancy, mortality due to cardiovascular diseases, cancer, age of puberty, and infant outcomes. The goal of this paper is to analyze the mechanisms underlying these facts.

There is a general consensus that selection can explain the first-generation advantage (Palloni and Morenoff, 2001; Jasso et al., 2004; Antecol and Bedard, 2006), however researchers are still puzzled about the possible explanations for the subsequent health convergence observed in the second-generation. The observed health patterns could be explained by the fact that immigrants are positively selected on health status (Palloni and Morenoff, 2001), hence first generation immigrants have better health outcomes, and that health status is only weakly correlated across generations (Jasso et al., 2004), so that the second generation essentially loses all the initial advantage, through a process of natural regression towards the mean. Some scholars emphasized the role of behaviors providing evidence of lower incidence risk-factors among immigrants at the time of migration and of a worsening of behaviors with time spent in the US and across generations (Acevedo-Garcia et al., 2005; Antecol and Bedard, 2006; Fenelon, 2012). Overall, the lack of extensive longitudinal data and the small sample sizes of surveys severely limited the ability to clarify the possible channels behind the Hispanic health paradox.

I contribute to these previous studies, by taking advantage of a large longitudinal and intergenerationally linked data set. In particular, this paper analyzes the birth outcomes of the universe of second and third generation Hispanics born in California and Florida, two of the top-immigrant receiving states in the US. Linking the birth records of two generations allows me to overcome some of the limits faced by previous studies and to investigate the factors affecting the generational worsening of birth outcomes among Hispanic immigrant descendants.

To test whether the paradox can be entirely explained by selection and regression towards the mean, I develop a simple model of health transmission. I use country-level difference in health outcomes to account for the degree of selection of the first generation, and existing estimates from the literature to impute the intergenerational transmission of health status. By doing so and assuming that second-generation immigrants are exposed to the same health care and socio-economic environment than the "average non-Hispanic native" and that their behaviors do not worsen across generations, the model confirms the paradox: second generation immigrants should still have an advantage, but they do not. However, while second-generation Hispanics improve their socio-economic status with respect to the first-generation, they still have lower socio-economic status than non-Hispanic natives. Therefore, on average they are not exposed to the same quality of care. Calibrating the differences in the quality of health care to match the differences in socio-economic status, the model not only explains all of the paradox, but, everything else constant, it actually over-predicts convergence and results in a "reverse paradox". Third-generation Hispanics show better birth outcomes than what we would expect, given the relatively low rate of intergenerational transmission observed in the data. The question is then how are second-generation Hispanics able to preserve their health advantage? In the paper, I show that first-generation immigrants have substantially lower incidence of both risky behaviors (smoking and alcohol consumption) and health risk factors (hypertension) that are known to importantly affect birth outcomes and to be importantly related to life style habits (Shireen and Lelia, 2006; Forman et al., 2009; Kaiser and Allen, 2002; Gonzalez, 2011). Behaviors did worsen slightly between first and second generation, but immigrants maintain a sizeable advantage in terms of lower incidence of health risk factors compared to white natives. The high persistence in healthy behaviors and in health conditions importantly related by life-style can explain 80% of the "reversed paradox".

The importance of behaviors is confirmed by the analysis of differences in the health convergence among second-generation Hispanics. I show that third-generation birth outcomes are importantly correlated with quality of care, socio-economic status, risky behaviors and health conditions. To address the potential endogeneity of these covariates, I follow the Currie and Moretti (2007) strategy of linking siblings, and I test whether the correlations are robust to the inclusion of grandmother fixed effects. Analyzing within family variations in the patterns of socio-economic and behavioral assimilation of second generation Hispanics allows me to disentangle the contribution of these factors from all the background characteristics that were common within a family, at the time of their birth (including the migrant's selectivity). Overall, the sibling's analysis confirms that behaviors do matter and do importantly affect differences in the convergence rate among Hispanics. The convergence is more marked among those who are less likely to maintain the health protective behaviors and conditions (eg. smoking, alcohol consumption, hypertension) that characterize the first generation immigrants and, more generally, among those who are less likely to have culturally assimilated. In particular, among second-generation Hispanics, intra-married couples show higher resilience in healthy behaviors, health conditions and birth outcomes. Using ethnic inter-marriage as a broader metric of cultural assimilation, I show that third-generation children of inter-married Hispanic couples are 13% more likely to be low-birth weight than children of intra-married couples. This result is particularly striking given that inter-marriage is usually associated to positive socio-economic outcomes.

The paper is organized as following. Section 2 discusses the explanations put forth by previous literature to understand the Hispanic paradox. In section 3, describes the data and verifies the Hispanic paradox in birth outcomes. Section 4 discusses the possible mechanisms behind these health patterns. In section 5, I examine the heterogeneity in the health convergence within the Hispanic group, exploiting siblings fixed effects. Concluding remarks are in Section 6.

2 Hispanic paradox: selection or worsening of behaviors?

A vast literature investigates the health differences between natives and immigrants. Most papers show that immigrants are healthier upon their arrival and at the same time their advantage erodes over time spent in the US.[1] As mentioned above, there is a consensus on the fact that the initial health advantage can be explained by positive selection into the United States biasing upward the immigrant-native health differences (Palloni and Morenoff, 2001; Jasso et al., 2004; Chiswick et al., 2008; Antecol and Bedard, 2006), but researchers continue to be puzzled by the mechanism underlying the following convergence to native health status.

A first group of scholars (Palloni and Morenoff, 2001; Jasso et al., 2004; Chiswick et al., 2008) argue that the apparent deterioration could be largely attributed to a statistical artefact: the regression towards the mean, following the initial selection. In particular, Palloni and Morenoff (2001) provide a simple model to show how even a moderate degree of selection at migration could explain the second-generation advantage in birth outcomes. Following this line of argument, Jasso et al. (2004) suggest that immigrants might select on transitory health traits and that their inability to fully forecast the evolution of their health might turn in a natural process of reversion towards the average health of the original population. These articles provide empirical support for the selection hypothesis as a plausible explanation of

[1] Gutmann et al. (1998) describe the origin of the epidemiologic paradox. Using data from the 1910 US Census and the 1990 linked birth and death certificate file the authors find that Hispanics suffered higher child mortality than non-Hispanic, but there was already some evidence of an advantage compared to the African American population. Historical data suggest that Hispanics did not show better birth outcomes than white non-Hispanic natives until the early sixties.

the initial health advantage observed among first-generation immigrants and their children compared to natives. However, none of these studies tested directly the implications of selection and regression towards the mean on the adult health of second generation immigrants and on the birth outcomes of their children. [2]

A second strand of the literature emphasized the importance of negative acculturation. According to these studies the unhealthy convergence can be explained by the worsening of dietary styles, the adoption of risky behaviors (smoking, alcohol consumption, substance use) and the erosion of social and cultural protective factors such as familism and religiosity (Guendelman and Abrams, 1995; Acevedo-Garcia et al., 2004; Antecol and Bedard, 2006; Fenelon, 2012). These studies provide evidence of a protective effect of foreign-born status, ethnic density, age at migration, years since immigration, on birth outcomes and infant health risk factors (Acevedo-Garcia et al., 2005; Bates and Teitler, 2008; Guendelman and Abrams, 1995; Hummer et al., 2007; Finch et al., 2007; Osypuk et al., 2010; Shaw et al., 2010). However, these papers did not attempt to disentangle the causal effect of behaviors, accounting for selection and other potential confounding factors.

Indeed, most of these studies were limited in their scope by either the sample size, the cross-sectional nature of the data or the lack of objective and reliable measures on nativity, ethnicity and health. Previous research relied on the use of synthetic cohorts of immigrants (Antecol and Bedard, 2006; Kaushal, 2008) to analyze the effects of time spent in US and age at arrival on obesity and other health outcomes. Due to the lack of information on parental nativity, researchers were often forced to use foreign-born status and self-reported ethnicity to analyze generational changes in health and health related behaviors. While Jasso et al. (2004) point out the need to analyze health trajectories of immigrants across generation, to the best of my knowledge there is no paper analyzing the Hispanic health paradox using individual linked data on two generations of immigrants.

[2] Previous scholars have also argued on whether the low infant mortality observed among Hispanics might be explained by selective re-migration (Palloni and Arias, 2004). While this is a relevant concern, Hummer et al. (2007) show that women of Mexican origin are extremely unlikely to migrate to Mexico with newborn babies.

The large size of the data and the ability to link the records of two generations allows me to address questions that other researchers had eluded. Exploiting the intergenerational nature of the data, I can verify whether the apparent deterioration in birth outcomes can be entirely explained by a positive selection at migration and a subsequent regression towards the mean. Furthermore, the large size of the sample allows me to focus on a sub-sample of second-generation siblings. Using within-family variation, I can partially isolate the original migrant's selectivity in order to analyze the heterogeneity in the path of convergence observed in the birth outcomes of third generation. Finally, one of the important advantages of this paper is that most of the health outcomes considered (pregnancy outcomes and maternal health characteristics) are recorded by medical officials and are therefore not subject to self-reporting bias.

3 Data

The main data used in this paper are drawn from the Birth Statistical Master File provided by the Office of Vital Records of the California Department of health and from the Birth Master Dataset provided by the Bureau of Vital Statistics of the Florida Department of health. These data contain information extracted from the birth certificates for all children born in California and in Florida for the years 1975-1981 and 1989-2009.

Information on mother's country and state of birth, mother's first and maiden name, state of birth occurrence, child's full name, date of birth, gender, parity, race, birth weight, hospital of birth, county of birth are available in both states for all the period considered. However, not all the variables are available in each year and for each of the two states. For instance, mother's age is reported for the entire period in California, but only only since 1989 in Florida, while mother's education is reported for the entire period in Florida, but only since 1989 in California. Data do not contain information on legal marital status, which is inferred. Information on birth-weight is available for the entire period in both states, while unfortunately other important measures of health at birth (eg.: APGAR scores, gestational lenght) are available only in the more recent years . Almond et al. (2005) and Wilcox (2001) cast doubt on the causal effect birth weight might have on mortality and more generally on infant health. Despite the debate on the causal mechanism, there is a general consensus that low birth weight is an important marker of health at birth and that is strongly associated with higher risk of mortality and morbidity (Currie and Moretti, 2007; Conley and Bennett, 2000). Since this study does not analyze the effects of birth weight and given that birth weight is the only measure of birth outcome available for the entire period, I will mostly focus on birth weight and incidence of low birth weight as indicators of health at birth. However, results go in the same direction when using alternative measures (e.g. APGAR scores, infant mortality) of infant health for the years in which other metrics are available. A full description of the variables used in the paper and their availability in each of the two states for the period considered is provided in the Appendix (see Table A1).

Similarly to previous literature (Fryer and Levitt, 2004; Currie and Moretti, 2007; Royer, 2009) using administrative birth records, I am able to link information available at a woman's birth to that of her children if the woman is both born in California (Florida) and also gave birth in California (Florida). Unfortunately father's information is less detailed. In particular, father's first name and father's state of birth are not available for the entire period under analysis in California. To ensure the comparability of the analysis in the two states, I focus on women and in particular on the cohort of women born between 1975 and 1981.[3]

One of the typical drawbacks of administrative vital statistics is the lack of information on individual income and occupation. However, the data contain some information on parental education and on the mother's zip-code of residence, from 1989 onwards in Cal-

[3]Florida's data contain information on father's full name and date of birth which allows me to conduct a parallel analysis using father's information. However, due to the lower quality of information on fathers and their lower likelihood of becoming parents at an early age, the matching rate is considerably lower than what observed among women and the selectivity of the sample increases. Results are similar in direction, but only marginally significant and are available upon request.

ifornia and for the entire period in Florida. Therefore, with the Florida's data I can use grandmother's education, as well as the median income, and the poverty rate in her zip-code of residence. In California, I do not have information on the grandmother's education as well as on grandmother's zip-code of residence, but I can use the socio-economic characteristics of the zip-code of the hospital where the birth occurred as a proxy for the socio-economic status of the grandmother as in Currie and Moretti (2007). Data on zip-code socio-demographic and economic characteristics are drawn from the Census (source: Social Explorer). In particular, I use median family income and poverty rate as of the 1980 Census for the zip-code of mother's birth and grand-mother's residence and as of 1990 Census for the zip-code of child's birth and mother's residence.

To construct the inter-generational sample, I linked the records of all the infants born between 1989 and 2009 whose mother was born in California or Florida between 1975 and 1981 to the birth records of their mothers. I matched the child's birth record to the mother's record by mother's first and maiden name, exact date of birth and state of birth. Whenever I was able to uniquely identify the mother's birth record, I included the observations in the linked sample.[4]

I then identify siblings born between 1975 and 1981 using information on the grandmother (mother's mother). To match grand-mothers (the first generation immigrants) across different birth certificates of their children (second generation immigrants) I use information on grandmother's name, child's last name, mother's race, and mother's state of birth. This implies that children born to the same mother but from different fathers would not be considered in my sample of siblings. I drop individuals for whom the matching variables are

[4]The distribution of the number of children born to mothers born in California and Florida between 1975 and 1981 is similar to the one obtained using the ACS 2010, confirming the quality of the matching. In particular, in the Californian data, 39% of the mothers born between 1975 and 1981 had one child before 2009, 34% had two children, 17% had three children, and the rest had 4 or more. Similarly, in the 2010 ACS, 31% of the mothers born between 1975 and 1981 had one child before 2009 36% two children, 20% had three children, and the rest had 4 or more. In Florida, 40% of the mothers born between 1975 and 1981 had one child before 2009, 33% had two children, 17% had three children, and the rest had 4 or more, similarly to what found in the 2010 ACS: 31% of the mothers born between 1975 and 1981 had one child before 2009 38% two children, 18% three children, and the rest had 4 or more.

missing. Following Currie and Moretti (2007) I drop all the grandmothers who are matched to more than 20 grandchildren according to my matching strategy (9304 observations in California, 5860 observations in Florida).[5] I also analyze data drawn from the Natality Detail Data which collect births data for all births occurred in the United States. The Natality Detail Data, in its public version, does not allow to link the records of two generations, since it does not release information on child and mother's name. Geographic data include only state, county, city, SMSA (1980 onwards), and metropolitan and non-metropolitan counties. The 2005-on data does not include any geographic variables such as state, county, or msa. However, I can use the Natality Detail Data data to conduct cross-sectional analysis for the entire U.S. for the years 1975-2004.

3.1 Matching and selection: descriptive statistics

The quality of matching for children born between in California and Florida between 1989 and 2009, whose parents were born in the same states between 1975 and 1981 is high in Florida (96.6%) and only slightly lower in California (87.5 %). I do not manage to match observations for which names were misspelled or changed across birth certificates, date of birth was mis-reported or could not to be uniquely identified with the available information. Despite the high matching, the linked sample is not representative of women (men) born between 1975 and 1981. The final sample includes 726,837 (44%) of the 1,643,865 female children born between 1975 and 1981 in California and Florida. This reflects the fact that not all the women born in Florida and California between 1975 and 1981 were still living in those states between 1989 and 2009 and not all of them became mothers before 2009. In particular, the Natality Detail Data, which contains information on mother's state of birth

[5] Regarding the matching of mothers to grand-mothers, in California, in 84% of the cases I matched only one daughter, in 12% of the cases I matched two daughters, and in 4% I matched three or more daughters to each grand-mother. In Florida, in 80% of the cases I matched only one daughter, 17% of the cases two daughters, and in around 3% of the cases I matched three or more daughters to each grand-mother. In the overall sample, the average number of children matched to each mother is 1.91, the average number of grandchildren linked to each grandmother is 2.50 grandchildren and 4.20 if conditioning on linking at least two second generation sisters to their offspring.

and state of birth of the child, shows that around 11.5% of women born in California and 13.5% of women born in Florida between 1975 and 1981 had a child in a other US state before 2004 (latest year for which both the information on the state of birth of the mother and the state of birth of the child are available in the Natality Detail Data). At the same time, using the American Community Survey (2010), we know that about 42% women born in California and 39% women born in Florida between 1975 and 1981 had not yet had a child in 2009. Data problems such as misspelling or missing information account for the rest of the attrition.

Table 1 shows the matching rates for the main race and ethnic groups in the sample. The matching is particularly high for African-American (70%) due to both their lower mobility and higher likelihood of having a child at a early age. Children of Hispanic immigrants have a higher rate of matching than natives. The matching rate among children of Mexican origin is 55% and 50% among second generation immigrants of Cuban origin. Children of Puerto-Rican immigrant women are less likely to be linked (37%). It is worth noting that differences in the density of ethnic networks as well in the different type of migration are reflected in different matching rates among Mexicans and Cubans in the two states. The rate of matching also depends on socio-economic background which is clearly associated to infant health, mobility and age at first birth. Children of first-generation mothers who were residing in poor zip-codes (lowest income quartile) are more likely to be linked to the records of their offspring than the children of first-generation mothers who were living in rich zip-codes (highest income quartile).

While these descriptive statistics show evidence of selection on socio-demographic characteristics (see column 3), the differences in initial health endowments between the linked and the non-linked observations are not striking (see column 4-9). If anything they suggest that the linked sample has a slightly lower incidence of low birth weight. The differences in birth weight appear to be negligible and non-systematic. A 100-gram increase in birth weight increases only by 0.5% the probability of a later observation. However, if the mother was born with a weight below the 2,500 grams threshold she would be 13% less likely to be linked. The lower incidence of low birth weight in the linked sample can be explained by higher rates of infant death, higher probability of returning to the family's country of origin ("salmon bias") or by a lower probability of having a child among those children born with poor health outcomes. To partially take into account the selection bias, I used a two-step Heckman selection model using an exclusion restriction based the year of birth of the second generation.[6] The results are substantially unchanged.[7]

3.2 Empirical specification

The focus of the paper is on the mechanisms behind the apparent deterioration in infant health. However, it is important to first verify the Hispanic paradox within the sample of birth records under analysis. To this scope, I use a simple linear probability model that relies on a comprehensive set of individual and contextual controls to study the conditional differences in birth outcomes between immigrants and natives. For expositional ease, for both the immigrants and the natives in the sample, I refer to first generation for all the women delivering between 1975 and 1981, second generation for all the children born between 1975 and 1981 and who delivered between 1989 and 2009, third generation for all the children born between 1989 and 2009. Formally, I consider the following model:

$$H_{izt,2} = \alpha + \beta IMM_{izt,1} + \gamma X_{izt,1} + \tau_{t,2} + \xi_{z,2} + \epsilon_{izt,2} \tag{1}$$

where the subscript 1 and 2 represents first and second generation. H_{izt} is a birth outcome outcome (eg. birth weight, incidence of low birth weight etc.) of the second generation child

[6]The year of birth of second generation women is an important predictor of later observation, while differences in birth outcomes by year of birth are negligible when considering children born less than 6 years apart.

[7]Palloni and Arias (2004) suggested that large part of the lower mortality rates observed in the Mexican population can be explained by selective out-migration ("salmon bias" effect). However, Hummer et al. (2007) argue that selective-out migration is unlikely to explain the advantage observed in the health outcomes of second-generation children, expecially when looking at first-hour, first-day and first-week mortality.

i, whose mother resided (or delivered) in zip-code z at time t. $IMM_{izt,1}$ is a dummy equal to one if the first generation woman delivering between 1975 and 1981 was born outside the US. X_{izt} include a set of individual socio-demographic characteristics of the first generation mothers including education (4 categories: high-school dropout, high-school graduate, some college, college or more), marital status[8], race, age dummies (in Florida mother's age is not available in the period 1975-1981), an index of adequacy of prenatal care based on the month in which prenatal care started, father's age (quadratic), father's education (four groups) and father's race. [9] I include indicators for missing information on parental education and age, marital status, and parity. Finally, I control for both time $\tau_{t,2}$ and zip-code $\xi_{z,2}$ fixed effects. I then turn to the analysis of the linked sample and analyze whether these differences persist over time and are transferred to the third-generation immigrants children. Formally, I estimate the following model:

$$H_{izt,3} = \alpha + \beta IMM_{izt,1} + \gamma X_{izt,2} + \tau_{t,3} + \xi_{z,3} + \epsilon_{izt,3} \tag{2}$$

where the subscript 1, 2 and 3 represents first, second and third generation. $H_{izt,3}$ is a birth outcome of the third generation child, whose mother resided (or delivered) in zip-code z at time t. $IMM_{izt,1}$ is a dummy equal to one if the first generation was born outside US. To ensure the comparability of the analysis, the model includes the same set of controls used in the analysis of second generation birth outcomes.

3.3 Verifying the Hispanic paradox in birth outcomes

Column 1 and 2 in Table 2 report the conditional birth outcomes differences between children of immigrants coming from the three largest Hispanic groups in the US (Mexicans,

[8] Marital status is self-reported in Florida and is inferred in California.

[9] In Florida, the month in which prenatal care started is imputed using the number of visits and the usual relationship between the number of visits and the month in which prenatal care started. However, results are similar when using the number of visits.

Puerto Ricans and Cubans) and children of white US-born mothers.[10] I restricted the sample to children born between 1975 and 1981 to white mothers and Hispanic first-generation immigrant mothers coming from Cuba, Puerto Rico and Mexico. [11]

The final sample includes 2,234,571 births for which information on birth weight and zip-code is not missing.[12] Table A2 in the Appendix shows the sensitivity of the magnitude of the coefficients to the addition of different set of controls. It is important to note that the addition of geographic controls (county, hospital or zip-code fixed effects) is associated with a stronger advantage in terms of lower risk of LBW for children of Mexican origin. This is consistent with the original definition of the epidemiological paradox as the fact that children of Hispanic immigrants fare considerably better than children of non-Hispanic women sharing a similar socio-economic background.

The results presented in column 1 and 2 document no significant differences in birth weight, but a lower incidence of low birth weight among children of Cuban mothers. Children of Mexican mothers are only slightly heavier (about 22 grams), but show a significant lower incidence of low birth weight compared to the children of white native mothers sharing similar socio-economic background. In contrast, Puerto Rican mothers are more likely to give birth to lighter babies.[13] Overall, column 1 and 2 show that while we do observe a healthy immigrant effect when considering the incidence of low birth weight for Mexicans and

[10]In this paper, I focus on immigrant of Hispanic origin, for which the paradox is particularly striking given their socio-economic background characteristics, and who are by far the largest ethnic group in the United States. However, when looking at the same analysis for children of immigrants coming from other countries, I find that the incidence of low birth weight is 12% lower among children of Canadians than among US natives, while it is 20% higher among children of Japanese, and is non-significantly different among children of Chinese and Vietnamese mothers, though the coefficient is negative for the latter.

[11]Mother's ethnicity is not consistently reported before 1989. Restricting the sample to the second-generation mothers I am able to link to their offspring, I can use the ethnicity reported at the time of delivery to further restrict the sample of natives to non-Hispanics. Coefficients differ only slightly in the magnitude, consistently with the patterns of convergence observed among immigrants of Hispanic origin. Results are similar when considering the sample of male and female children separately. These tables are available upon request.

[12]Notice that this number includes male and female births and therefore is about twice as large as the number of observations presented in Table 1, which only includes the birth records of women who could be potentially linked to the birth records of their off-spring. Furthermore, in Table 1 the overall sample also includes black children. Results are similar using only data on women born between 1975 and 1981.

[13]When breaking down the analysis by state, the coefficient for children of Mexican mothers tends to be higher in Florida than in California, probably reflecting higher selection.

Cubans, children of Puerto Ricans fare considerably worse than their native counterparts. This is consistent with the idea that they might be less favorably selected, since Puerto Rico is a U.S. territory. Even among children of Mexicans, for which the advantage in low birth weight is highest, there is only a 22 grams difference in the average birth weight. The differences between the continuous and the discrete outcome variable reflect the independence of the predominant and residual distribution of birth weight and more generally differences in the distribution of term and pre-term births (Wilcox, 2001).[14] Figure 1 describes the distribution of birth weight in California and Florida for the immigrant descendants of Hispanic origin and white natives, over the period analyzed in this study. The distributions show the typical bell shape. However, previous studies noted that the size and the nature of the effects of covariates on the conditional mean might not capture the importance of the effects on the lower tail of the birth weight distribution (Koenker and Hallock, 2001). Indeed, quantile regression shows that the advantage in birth weight (in grams) is more substantial at the left tail of the birth weight distribution. Children of Hispanic origin immigrants are on average 50 grams heavier than children of white natives in the 5% quantile of the distribution, while the differences are much smaller and become even negative in the upper tail of the distribution. In particular, in the 5% quantile of the distribution children of Mexican mothers weigh on average 70 grams more than children of white native mothers and are on average 50 grams heavier in the 10% quantile (see Table A5 in the Appendix). [15] To sum up, column 1 and 2 document that the healthy immigrant effect in infant outcomes is mostly concentrated in the lower tail of the birth weight distribution, and it is heterogeneous across ethnic groups.

Columns 3 and 4 illustrate the differences in birth weight and incidence of low birth weight between third generation children whose grand-mothers were born in Cuba, Puerto

[14]The predominant distribution is substantially equivalent to the distribution of birth weight observed for term births.

[15]The 0.05 quantile roughly corresponds to the traditional threshold of low-birth weight. In the quantile regression, I include gender, marital status, adequacy of prenatal care, parity, type of birth, year fixed effect, state fixed effect, maternal education (Florida), and a quadratic in age. This is substantially equivalent to the specification used in Table 2, without including zip-code fixed effects.

Rico or Mexico and the third-generation white natives.[16] The estimates include the same set of controls used in columns 1 and 2. [17]

The deterioration in birth outcomes is mostly evident when examining the incidence of low birth weight, but even when analyzing differences in birth weight the coefficients are always negative and larger in magnitude compared to those observed among second-generation immigrants. Note that the average incidence of low birth weight is relatively stable among second and third generation white natives (see Table A4 in the Appendix). Third generation children of Mexican origin (column 4, row 2) conserve some of the initial health advantage, but the coefficient shrank significantly (by about 65%) compared to the one observed among second-generation children in column 2. The deterioration with respect to the native birth outcomes is even stronger when analyzing children of Cuban and Puerto Rican origin.

4 A simple model of selection and health transmission

As mentioned earlier in the paper, previous scholars questioned the paradoxical nature of these stylized facts arguing that they could be entirely explained by selection and a subsequent process of regression towards the mean. Immigrants are positively selected on health, but since health is only weakly correlated across generations, we observe convergence in the birth outcomes of the third generation. To verify this hypothesis I build on Palloni and Morenoff (2001) and introduce a very simple model of selection on health at migration and intergenerational health transmission. Due to the limited information available on birth weight distribution in the country of origin, I am not able to provide a direct estimate of the original selection. However, I can calibrate the model using the observed differences

[16]Unfortunately, data do not contain information on the country of origin of the father for the entire period. To be able to compare with the results shown in columns 1 and 2, I included all grand-children of US-born white women. However, one could restrict the sample grand-children of US-born white women whose mothers did not report Hispanic origin. The results (available upon request) would be substantially similar.

[17]Table A3 reports the conditional mean differences obtained using different set of control variables.

in health outcomes between the US and the sending countries to pin down the degree of selection of first-generation and existing estimates from the literature to capture the degree of intergenerational transmission of health. To keep the model simple and intuitive, I focus on the major sending country, Mexico and compare the health distribution of Mexicans and natives in the US.

Similarly to the traditional Roy model of self-selection (Roy, 1951; Borjas, 1988) the decision to migrate can be represented by a dichotomous variable that equals 1 when an individual migrates and takes value 0 otherwise. Here for simplicity, I assume that the migration decision will only depend on the health of the migrant. The underlying idea is that cost of migration is higher for those who are less healthy. Immigrants who have a health above a certain threshold t_1 at the time of migration will be able to migrate, while the rest will stay in the country of origin. Formally:

$$Imm_1 = \begin{cases} 1 & \text{if } h_1 \geq t_1 \\ 0 & \text{if } h_1 < t_1 \end{cases} \tag{3}$$

$$h_1 = u_1 \tag{4}$$

where h_1 is the health of first generation at the time of migration, u_1 is distributed as a random normal (μ_j,1) reflecting the health distribution in the country of origin, μ_j is the average health in country j, and t_1 is the migration threshold. We can think of μ_j as the composite effect of genes, quality of health care, socio-economic environment, and behaviors on health. Individuals with $h_1 \geq t_1$ will be able to migrate. The higher the threshold, the more selected is the sample of migrants. The incidence of low birth weight is determined as follows:

$$BW_2 = \gamma h_1 + v_2 \tag{5}$$

$$LBW_2 = \begin{cases} 1 & \text{if } BW_2 \leq t_2 \\ 0 & \text{if } BW_2 > t_2 \end{cases} \tag{6}$$

where BW_2 is the birth weight of the second generation, h_1 captures maternal health at migration, v_2 is distributed as a random (0,1) normal variable reflecting the effect of other unobservables on the birth weight of second generation, γ captures the effect of maternal health on child's health and t_2 represents the low birth weight threshold. Similarly, third generation birth outcomes can be described as a function of second generation health characteristics and other factors. Formally:

$$h_2 = \rho h_1 + u_2 \tag{7}$$

where h_2 is the health of second generation mothers, u_2 is distributed as a random (μ_{j2},1) normal variable reflecting the effect of other unobservables on second generation mother's health, ρ measures the degree of intergenerational correlation in health between first and second generation. The birth weight of the third generation is then expressed as a function of maternal health. Formally:

$$BW_3 = \gamma h_2 + v_3 \tag{8}$$

$$LBW_3 = \begin{cases} 1 & \text{if } BW_3 \leq t_2 \\ 0 & \text{if } BW_3 > t_2 \end{cases} \tag{9}$$

where BW_3 is the birth weight distribution in the third generation, v_3 is distributed as a random normal (0,1) variable reflecting the effect of other unobservables on the birth weight of third generation, t_2 determines the amount of low birth weight in the third generation. Assuming that the unobserved random shocks to health and birth weiht are not correlated, the intergenerational correlation in birth weight can be then rewritten as following:

$$Cov(BW_3, BW_2) = Cov(\gamma h_2 + v_3, \gamma h_1 + v_2) = Cov(\gamma \rho h_1 + \gamma u_2 + v_3, \gamma h_1 + v_2) = \gamma^2 \rho \tag{10}$$

which implies

$$\rho = \frac{Cov(BW_3, BW_2)}{\gamma^2} \tag{11}$$

Within this framework, I can estimate the extent to which selection and the estimated intergenerational correlation in health can explain the evolution of low birth weight incidence among immigrant descendants. Following the empirical strategy used in the previous sections, I analyze the trajectories of low birth weight incidence among immigrants of Mexican origin and use the white native population as a reference group. I start with a population of 10,000 potential Mexican migrants and 10,000 U.S. natives. Individuals receive a random value for their health at migration which is drawn from a normal distribution with mean identical to the mean of their country of origin (Mexico or the US). I use the native health as a benchmark and set μ_{US} equal to 0. The low birth weight threshold t_2 is set to be -1.57 to match the average incidence of low birth weight observed in the data (0.058) over the entire period studied (1975-2009) in the overall population in US (excluding blacks). The mean of unobservables shocks affecting health mu_{MX} is set to be such that the difference in the low birth weight of the two populations would be equivalent to that implied by the earliest available measure of incidence of low birth weight in Mexico (10.6%, see Buekens et al. (2012)) relative to the average incidence of low birth weight in the US non-black population (5.8%). I consider different values of ρ and γ such that equation 11 would hold for values of Cov (BW_3, BW_2) around the value of 0.2 which is the estimated intergenerational correlation in birth weight, consistent with what found by Currie and Moretti (2007). Previous studies estimated the intergenerational correlation in longevity and mortality to range between 0.2 and 0.3 (see Ahlburg (1998)) and the intergenerational correlation in BMI to be around 0.35 (see Classen (2010)). Based on these estimates, I focus the analysis on values of $\rho \in [0.2, 0.5]$. The above restrictions imply that γ has to be $\in [0.58, 1]$. [18] Here, I use as a baseline the

[18] Alternatively, I can set the intergenerational correlation in health and analyze the relationship between selection and differences in low birth weight for different values of γ and for a range of values of intergenerational correlation in birth weight around the 0.2 estimated in the data and in the literature. The implications of the model do not change substantially.

case in which ρ is equal to 0.35, but the interpretation of the simulation exercise does not change significantly for different values of ρ in the defined range [0.2,0.5].

Using this parametrization and under the assumption of identical effects of health on birth weight in the two populations, I simulate the intergenerational transmission of health between first and second generation immigrants and their children. This allows me to evaluate the importance of selection and regression towards the mean in explaining the observed differences in birth outcomes between natives and immigrant descendants. I repeat these simulations 1,000 times. Figure 2 shows the differences in the incidence of low birth weight between children of Mexican immigrants and children of native (y-axis) by different extent of selectivity at migration (x-axis). The plot in the graph describes the relationship between selectivity and the low birth weight differences with the natives for the baseline specification ($\rho = 0.35, \gamma = 0.75$). The dashed line marks the raw difference in low-birth weight between second-generation Mexicans and white natives as observed in the data (coef: -0.008, s.e.: 0.001). Table 3 reports the raw differences in the birth outcomes of children of Mexican origin and children of white natives born between 1975 and 1981 in in California and Florida (column 1) and for the entire US (column 4).

Figure 2 suggests that the initial advantage can be entirely explained by a relatively moderate selection at the time of migration. In particular, if individuals in the lowest quintile of the health distribution do not migrate due to their health conditions, the positive selection can explain the lower incidence of low birth weight observed among second generation Mexicans. To verify whether even the second part of the paradox -the deterioration of immigrant health- can be entirely explained just by a relatively weak intergenerational correlation in health (regression towards the mean), I consider and test the validity of alternative hypothesis regarding the socio-economic and behavioral assimilation of immigrants across generations.

I first assume that second-generation immigrants fully assimilate socio-economically and are exposed to the same quality of care and environmental characteristics of the average non-Hispanic native. In practice, $mu_{MX_2}= mu_{US_2}$=0. This scenario is described as Case 1 in Figure 3. The dashed lines in the figure report the observed differences in the birth outcomes of third-generation immigrants of Hispanic origin and third-generation natives (see columns 3 and 5 in Table 3). Assuming full assimilation, I do not find evidence of important differences in the incidence of low birth weight of third-generation Mexicans and natives, consistently with what shown in the data. The coefficients reported in columns 3 and 5 have opposite signs, but are both relatively small in magnitude (Table 3). The model fits fairly well the raw-differences in low birth weight observed in data between third-generation Mexicans and white native children (column 2 and 4, Table 3). However, this scenario (Case 1) is not realistic for two reasons. First, it does not consider the potential role of behaviors. Second, it assumes full socio-economic assimilation.

It is known that both substance abuse and dietary practices during pregnancy have important effects on the birth outcomes of children and previous studies provide evidence that immigrants behave significantly better than natives and have lower incidence of risk factors. Administrative records provide only limited information on health behavior during pregnancy and only for the more recent years. Therefore, I am not able to verify directly how the intergenerational changes in important risk factors such as smoking during pregnancy affect the intergenerational transmission of health at birth. However, I can provide cross-sectional evidence of differences between US-born second generation immigrants of Hispanic origin and first-generation immigrants. Information on adult behaviors and health conditions is very limited in California, while the Florida's data report tobacco, alcohol consumption and weight gain during pregnancy 1989 onwards, as well as information on pre-pregnancy BMI (weight and height), chronic hypertension, gestational hypertension and diabetes 2004 onwards. For this reason to analyze the role of behavioral assimilation, I focus on the Florida's sample (Panel A, Table 4), but I integrate the analysis using the information on behaviors and risk factors contained in the Natality Detail Data for the entire US (Panel B, Table 4).

Table 4 illustrates the incidence of risky behaviors and other risk factors that are known to be importantly related to life style among natives (coulumn 1), first (column 2) and second (column 3) generation Hispanics. Column 4-6 report the raw differences in the incidence of these risk factors between first-generation Hispanics and natives, second generation Hispanics and natives, and among Hispanics between second and first generation. First generation immigrants have substantially lower incidence of risk factors compared to non-Hispanic white natives. Second generation immigrants show some convergence towards the less healthy behaviors and higher incidence of risk factors of Americans, but they conserve a fairly sizeable advantage over natives.

Going back to the model and Figure 3, I consider a second scenario (Case 2) assuming that second-generation Mexicans maintain the lower incidence of risk factors observed among first-generation immigrants and they are fully economically assimilated. In practice, I consider $mu_{MX_2} > \epsilon > 0$. Under these assumption, the model would confirm the paradox: third generation children would be expected to show an advantage in low birth weight, but they do not.

However, even this second scenario is not completely realistic. Second-generation Mexicans are not likely to be exposed to the same quality of care, environment and socio-economic characteristics of the 'average non-Hispanic white'. Duncan and Trejo (2011) show that second-generation Mexican wages are on average 18% lower than among white natives, and the differences are even more important when analyzing family income. Therefore, I incorporate in the model the effect of the observed socio-economic differences in birth outcomes.

To isolate the effect of socio-economic factors, I first assume no difference in behaviors. In practice, I estimate what would be the expected difference in birth outcomes due to the observed difference in family income between Hispanics and natives. Previous studies provide few estimates of the causal effect of income on birth weight. Cramer (1995) finds that a 1 percent change in the income-to-poverty ratio increases birth weight by about 1.05 grams. More recently, Almond et al. (2009) find similar marginal effects analyzing the impact of food-stamps on birth outcomes. Using CPS data (1994-2009), I estimate that on average the family income-to-poverty ratio among Mexicans is 42% lower than among US natives (see Table 5, Panel B, column 6).[19] Using Cramer (1995) estimate, holding everything else constant, we should expect the birth weight of Mexicans to be 45 grams lower than the one observed among natives. I can then impute the difference between the health distribution of second-generation Mexicans and that of US natives, assuming that they fully assimilate to US natives on other unobservable characteristics affecting health (including behaviors). Under this scenario (Case 3, in Figure 3) and not considering the effect of behaviors, the model not only can explain the paradox, but actually over-predicts convergence: third-generation birth outcomes are predicted to be worse that they actually are (columns 2 and columns 6, Table 3). The paradox is reversed. What is striking is no longer why third generation children have worse birth outcomes with respect to second generation when compared to white natives, but rather why do they fare better than they should have. Results go in the same direction if considering the entire Hispanic (see Panel A Table 5) group or if using socio-economic information at the zip-code level (see Panel C and D, Table 5).

Table 4 suggests that the persistence of healthy behaviors can explain part of the reverse paradox predicted by the model. To better assess this hypothesis and quantify the importance of behaviors, I compute the difference in low birth weight between Mexicans and natives, conditioning on risky behaviors during pregnancy (tobacco and alcohol consumption and hypertension.Results are similar when including other health conditions such as obesity, diabetes or chronic hypertension. I focus on gestational hypertension because it is shown to significantly affect low-birth weight and it is particularly related to both exercise and dietary practise during pregnancy. When doing so, the coefficient becomes higher and closer to the one predicted by the model (see Table 3, columns 3 and 6). Accounting for the observed behaviors and health conditions, the model can explain about 80% of the reverse paradox

[19] 1994 is the earliest year in which information on father and mother's birth place is available in the CPS surveys.

found after accounting for socio-economic differences. [20]

Notice, that I am only able to account for the contribution of a limited set of behaviors, for which information is available in the data. Dietary practices have been shown to be important determinants of birth outcomes. In particular, there is evidence of the importance of fruit and vegetable intake. I cannot test directly how changes in dietary practices affected changes in birth outcomes across generations. However, having information on these behaviors one could explain a larger share of the deterioration in birth outcomes. Therefore, the unexplained part of the "reverse paradox" is likely to be explained by other behaviors (eg. dietary habits) that I do not observe and that are known to affect importantly birth outcomes.

Taken together, the model suggests that a combination of selection, alongside positive but less than full socio-economic assimilation as well as persistence in healthy behaviors and lower incidence of health risk factors can explain fairly well the observed health trajectories among immigrant descendants. To confirm the importance of socio-economic and behavioral assimilation, in the next section, I analyze the heterogeneity of health convergence among Hispanics.

5 Heterogeneity among second-generation Hispanic: a within family analysis

In this section, I take advantage of the large size of the data to analyze differences in the health trajectories among Hispanics. Examining the differences in the birth outcomes of third generation immigrants of Hispanic origin can further clarify to what extent the deterioration in infant health is inevitable and whether policies can help second generation immigrants maintaining and transmitting their health to their children. I explore the intergenerational

[20] More specifically, depending on whether we consider the low birth weight differences in the US or in the California and Florida's sample, controlling for behaviors and health conditions help us explaining between 66% and 83% of the reverse paradox. Despite these differences, these results show that once we account for both the persistence in healthy behaviors and the less than full socio-economic assimilation, the model fits fairly well the observed pattern in the data.

pathways of infant health following migration restricting the sample to the descendants of first generation immigrants. In particular, I focus on the role of socio-economic characteristics, quality of care and health risk factors. Since the information on behaviors is limited and available only for the Florida's data, I first introduce a broader metric of cultural assimilation, ethnic inter-marriage, which I show to be importantly related with behavioral assimilation. I then turn to analyze the direct role of behaviors using the Florida's data.

To partially isolate selectivity, I control for the birth weight of the second-generation mother and include grand-mother fixed effects exploiting only the within sibling variation in socio-economic, environmental and behavioral factors. Comparing the birth outcomes of third-generation cousins allows to eliminate the bias induced by genetic and environmental factors that are constant within the family and in particular for the common characteristics of mothers (sisters) who grew up in the same family.

Formally, I estimate the following linear probability model:

$$H_{it,3} = \alpha + \beta_2 SES + \beta_3 Acc_{it,2} + \gamma X_{it,2} + \lambda H_{i,2} + \epsilon_{it,3} \tag{12}$$

where $H_{it,3}$ is a metric of infant health of third generation, SES is an indicator of socio-economic status of the second generation, $Acc_{it,2}$ is a metric of cultural assimilation of the second generation or a risky health behavior, $X_{it,2}$ are a set of socio-demographic controls, $H_{it,2}$ is a metric of infant health of the second generation. To account for potential omitted variable bias, I include grand-mother fixed effects and exploit differences among siblings (within a family) in the covariates under analysis.

5.1 Socio-economic status and quality of care

Consistent with what shown in Section 4, the descriptive statistics reported in Table 5 show that first generation immigrants improve their socio-economic status with respect to first-generation immigrants, but do not fully assimilate. Earlier in the paper, I showed that these socio-economic patterns help to explain the observed health patterns, I now analyze how differences in socio-economic assimilation and in the quality of care received can help us to explain differences in the health convergence among Hispanics. This is not the ideal measure of adequacy of care (Kotelchuck, 1994). Indeed, one would like to know not only the initiation of care, but also the number of visits after care started. However, information on the total number of prenatal visits is available only 1989 onwards and the onset of care can be still used as a valid proxy for the quality of prenatal care received (Conway and Deb, 2005). [21] About 85% of the mothers start prenatal care in their first trimester of pregnancy (see Table 5,Panel D, column 4).

Table 6 summarizes the effects of socio-economic and cultural assimilation on the birth outcomes of third-generation of Hispanic origin. Column 1-5 report the simple association of birth weight (incidence of low birth weight) with the indicator of adequacy of care (column 1), a dummy for whether the mother did not graduate from high-school (column 2), an indicator for whether the second-generation mother was residing (column 3) or gave birth (column 4) in a high-poverty (lowest quartile of poverty distribution) zip-code, and a metric of cultural assimilation (inter-marriage), column 5. In Columns 6, I introduce a broader set of socio-demographic characteristics including a set of control variables available at the time of birth of second-generation: second generation birth weight, the poverty share in the zip-code of birth of the second generation mothers, first-generation grand-mother's age dummies, and dummies capturing the interaction of county of residence and year of birth at the time of second-generation's birth; and a more endogenous set of covariates containing information available at the time of birth of third-generation children: marital status, parity, father's education (4 educational groups), a quadratic in father's age, and dummies capturing the interaction of county of residence and year of birth at the time of third-generation's birth. Columns 7 includes grandmother-fixed effects exploiting within-family variation.[22] In column

[21] While Florida's data do not contain information on the month in which prenatal care began for the first period, I imputed the month of prenatal care start using information on the number of total prenatal visits.

[22] A more detailed analysis of the sensitivity of the coefficients to the addition of different batteries of controls is available in the Appendix.

8, I estimate the same model presented in column 7 among non-Hispanic white mothers.

The first row of in Table 6 shows that children of mothers who started care later in the pregnancy or had no prenatal care show significantly lighter birth weights (Panel A) and are at higher risk of low birth weight (Panel B). The coefficient on adequacy of prenatal care remains strong and significant to the addition of socio-demographic controls (column 6) and grand-mother fixed effects (column 7). These results emphasize the malleability of health at birth and support the idea that the regression to the mean might be partially buffered by improvement in adequacy of prenatal care. The adequacy of prenatal care is here defined as starting prenatal care in the first trimester of pregnancy.

While there is some evidence of a lower risk of low birth weight for those who have higher educational attainment, the effects of high-school graduation, college attendance and graduation have opposite sign when controlling for other socio-economic characteristics and introducing grand-mother fixed effects. For space considerations, in Table 6 I report the results including only a dummy for whether the mother graduated from high-school or dropped out. The coefficients on high-school drop-outs reported in column 6 and 7 should however be interpreted with caution given the collinearity of maternal education with adequacy of care, poverty of the neighbourhood of residence and father's education.

More robust evidence of a negative effect of socio-economic status is found when analyzing the poverty of the environment in which the birth occurred. In particular, in the fifth row, I use the poverty level of the zip-code of residence of second-generation mothers at the time of birth of their children to identify mothers who were living in the poorest zip-code (lowest quartile of income distribution, highest quartile of poverty distribution). The marginal effects with respect to the mean correspond to a 3.5% increase among Hispanic descendants. However, coefficients become smaller and non-significant when controlling for grand-mother fixed effects (column 7). Interestingly, the coefficients are larger and more robust when using the poverty rate in the zip-code of the hospital (sixth row) rather than the poverty in the residential zip-code of the mother. Third generation immigrants of Hispanic origin born in a hospital located in a high-poverty zip-code are 12% more likely to be low-birth weight than their counterparts born in hospital located in richer areas. The effect remains relevant and significant when including grand-mother fixed effects and exploiting differences in the adult socio-economic background of second-generation siblings (see column 7). These results might reflect selection on hospital choice as well as differences in the quality of care received. Overall, the evidence presented in Table 6 suggests that the poverty of the environment at birth is associated with poorer birth outcomes. The results are very similar when using the median income of the zip-code rather than the zip-code poverty rate to define socio-economic status.

5.2 The role of acculturation: analyzing ethnic inter-marriage

In the previous section, I showed that behaviors and risk factors play an important role in explaining the health trajectories of immigrant descendants. More generally, previous studies argued that the adoption of less health life style and the erosion of cultural protective factors explain part of the observed differences in the health convergence of immigrants. To overcome the limits imposed by the lack of extensive information on behaviors and risk factors for the entire period and in both states, I start using a broader metric of cultural assimilation: ethnic-inter-marriage.

Inter-marriage has been defined as "the last stage of assimilation" and it is known to importantly affect the process of adaptation in the new country (Furtado and Theodoropoulos, 2009; Furtado and Trejo, 2012). Acevedo-Garcia et al. (2005) provide evidence that ethnic inter-marriage has an important role in influencing the adoption of native life-style behaviors, affecting, for instance, anti-smoking socialization among Latinos. Here, I define inter-marriage as a dummy equal to one if the father is a non-Hispanic white and zero if the father is Hispanic. The sample is therefore restricted to second-generation immigrants women born between 1975 and 1981 and whose mothers migrated from Cuba, Puerto Rico or

Mexico. [23] Since I do not have information on the country of origin of the father for the entire period under study, I use information on father's ethnic origin to define inter-marriage. [24] I include all the observations for which I had information on father' ethnicity and, hereafter, I use inter-marriage to indicate the the different ethnicity of mother and father, regardless of the reported or inferred marital status.

Table 7 confirms that inter-marriage is importantly related to behavioral assimilation. Panel A reports raw correlations for California and Florida, while Panel B shows the same correlation using the Natality Detail Data for the entire US. Second-generation intra-married couples are less likely to smoke and drink during pregnancy and they show higher rates of chronic and gestational hypertension. The rate of smoking among second-generation intra-married couples is the same as the one observed among first-generation immigrants (see Table A6 in the Appendix). In Table 8, I show that these correlation hold up when controlling for socio-demographic characteristics and including grand-mother fixed effects. Panel A shows that the coefficient of inter-marriage on smoking is always positive and significant. It shrinks from 2 percentage points to 0.6 percentage points after including grand-mother fixed effects, but this is still a large and meaningful effect. With respect to the average rate of smoking, being inter-married is associated with a 30% higher likelihood of tobacco use during pregnancy. Results go in the same direction when looking at other risk factors such as alcohol use and gestational hypertension (Panel B and C). However, in this case the magnitude and the significance of the coefficient is less robust to the addition of grand-mother fixed effects. Inter-marriage is of course a very coarse measure of acculturation, but this evidence suggests that it might be importantly correlated to many other important behaviors which are observed to change across generations and are known to importantly affect the health transmission process. [25]

[23] The data contain information on the country of origin only for major sending countries. However, when considering individuals whose mothers were born abroad and who report Hispanic origin results are similar.

[24] Note that Baumeister et al. (2000) found a surprisingly high coincidence (95%) between the ethnicity reported by birth clerks in the California birth certificates and the ethnicity reconstructed in personal interviews.

[25] For instance, using the American Time Use Survey, I find evidence that among immigrants of Hispanic

The coefficients of inter-marriage on birth weight and incidence of low birth weight reported in Table 6 suggest the existence of a paradox within the paradox: children of second-generation Hispanic mothers who inter-married have worse birth outcomes, despite their higher socio-economic status. In particular, children of Hispanic mothers who inter-married with a non-Hispanic white native weigh on average 76 grams less than children of intra-married couples. However, when including grand-mother fixed effects the difference becomes less important with an effect around -0.5% with respect to the average birth weight in the sample. Again, the effects are larger in magnitude when focusing on the lower tail of the birth weight distribution. Inter-marriage is associated to a significantly higher risk of low-birth weight. This effect is robust to the inclusion of socio-demographic controls. In particular, columns 1 and 6 show that inter-marriage is associated with a 1.7 (1.4) percentage points increase in the incidence of low birth-weight. When adding grand-mother fixed effects (column 7) and focusing on sisters the coefficient is halved the coefficient is smaller, but still large and significant. The incidence of low birth weight among children of inter-married couples is 0.007 percentage point higher than among inter-married couples (+13% with respect to the mean of the dependent variable). It is worth noting that this result is not sensitive to the addition of controls for mother's and father's educational dummies which are included in column 7.

In light of the previous literature (Furtado and Theodoropoulos, 2009; Furtado and Trejo, 2012; Wang, 2012) providing evidence of a positive relationship between inter-marriage and socio-economic outcomes, the fact that the children of inter-married couples have worse birth outcomes is particularly striking. Clearly inter-marriage is not an exogenous decision, but the unobservables that are usually associated with the likelihood of marrying a native would be likely to downward bias the estimated coefficient. Wang (2012) reports that Hispanic-White couples on average earn about $20,000 more than Hispanic-Hispanic couples. [26] A

origin, intra-marriage is positively associated with time spent in food-preparation. While this goes out of the scope of this paper, it is important to notice that recent research has shown important correlations between health and consumption of more processed food.

[26] In the data inter-marriage is positively correlated with median family income (+0.15) in the zip-code,

relevant concern is that selection in the marriage market might be substantially different for men and women. However, while there are important gender differences in inter-marriage rates between blacks and whites, Wang (2012) shows that there are no significant gender differences in the inter-marriage rates of Hispanic and whites and that white men who married Hispanic women are not less educated that those who married white women. [27]

The results presented in columns 7 might still reflect the selectivity of migrants on the father's side. By controlling for grand-mother fixed effects and mother's birth weight I am able to eliminate the selection effects coming from the mother's side, but the inter-marriage coefficient might be the result of the genetic advantage carried by the immigrant-descendant father. One could argue that by controlling for mother's birth weight and assuming positive assortative mating in birth weight, father's selectivity would be partially captured. The paucity of information on father's full name does not allow me to provide robust evidence on this. An alternative way to test for the role of father's selectivity is to conduct a placebo test analyzing the effect of inter-marriage among second-generation non-Hispanic white women and focusing on the birth outcomes of their off-spring. If the inter-marriage coefficient is capturing father's selectivity, one should expect a protective role of having a Hispanic father even when analyzing the effects of inter-marriage on birth weight among third-generation native children. When conducting this test, I find that marrying a man of Hispanic origin has no protective effect and if anything increases the risk of low-birth weight (see column 8 in Table6).

Though Wang (2012) documents only a slight difference in the education levels of white women who married with Hispanic men compared to those married to a non-Hispanic white native, one might be concerned that the placebo test might be invalidated by confounding factors affecting the selection in the marriage market. In particular, white women who marry a Hispanic man might be negatively selected with respect to white women who married

with mother's education (0.10) and father's education (.14), while negatively related to the zip-code poverty share (-.17).

[27] In particular, 32.3% of white men married to white women completed college education, compared to 33.1 % of white men who married to Hispanic women.

within the group. However, the placebo test is robust not only to the addition of father's and mother's education which are included in column 6-8, but also to the separate analysis of women who married to equally, more or less educated Hispanic men. Note also that when controlling for grand-mother fixed effects and previous generation birth weight, marrying to a high-school drop-out has no significant effects on the risk of low birth weight third generation white natives.

The placebo test suggests that rather than simply reflecting the selectivity of the Hispanic father, the coefficient reported in column 7 of Table 6 is likely to reflect the correlation of inter-marriage with behavioral assimilation that we observed in Tables 7 and 8. In particular, the comparison of inter-marriage coefficients reported in columns 4 and 5 of Table 7 suggests that observed risk factors capture part of the protective ethnic effect of intra-marriage. While the coefficients are not statistically different in the Florida's sample, the point-estimate is always lower in magnitude when controlling for risky behaviors and hypertension. Furthermore, in the Natality Detail Data,due to the higher number of observations, the differences in the inter-marriage coefficient between column 4 and 5 become statistically significant. The results go in the same direction when controlling for the sample selection induced by missing information on behaviors (see Table A9).

There might be other unobserved factors affecting both selection in the marriage market and birth outcomes. However, the extensive set of controls, the siblings analysis and the overall robustness of the coefficient reduces the concern that these confounding factors might significantly alter the main finding as well as the validity of the falsification test.

5.3 Maternal behaviors

Earlier in the paper, I showed evidence of the important relationship between behaviors and risk-factors (see Table 3 and Table 7). Using the Florida's sample, I can analyze more directly the effect of behaviors on the birth outcomes of third generation immigrants, controlling for a richer set of individual characteristics and exploiting within siblings' variation. Table 9 illustrates the relationship between different risk factors and the incidence of low birth weight. The sample is composed of all the mothers born between 1971 and 1985 in Florida, including both natives and Hispanics, since here I am interested in providing further evidence of the causal effect of behaviors. However, including a set dummies capturing the interaction between risk-factors and country of origin does not affect the results. Columns 1-3 analyze the effect on low birth weight of risk-factors for which information is available since 1989: tobacco and alcohol consumption during pregnancy, normal weight gain which is defined gaining between 24 and 40 pounds.[28] Column 2 control for the same set of socio-demographic controls using in 6. Column 3 includes grand-mother fixed effects.

Smoking during pregnancy has been widely recognized as the most modifiable risk factor for LBW Almond et al. (2005); Currie (2007). Table 9 shows that tobacco use during pregnancy increases the incidence of low-birth weight by about 2 percentage points, which is more than 20% of the average incidence of low birth weight in the Florida's sample. The coefficient is relatively robust to the addition of socio-demographic controls (column 2), grandmother fixed effects (column 3) and other risk-factors for which information is available 2004 onwards (column 4-6). Alcohol use during pregnancy is associated with a 4 percentage points increase in the incidence of low-birth weight, but the coefficient becomes non-significant once controlling for socio-demographic characteristics. Having a normal weight gain (not-adjusted for BMI) is associated to lower likelihood of low-birth weight (columns 1-3). Mother's chronic and gestational hypertension are shown to be important determinants for the incidence of low-birth weight. The coefficients are important in magnitude and robust to the addition of socio-demographic controls (column 5) and grandmother fixed effects (column 6). Hypertension is known to be importantly related to dietary habits and physical exercise. Diabetes increases the risk of low birth weight, but the coefficient is not significant after controlling for grand-mother fixed effects. Since 2004 onwards the Florida's data contain information on pre-pregnancy BMI, I can compute a better measure of adequate gain using the IOM

[28] I do not have information on pre-pregnancy BMI before 2004.

criteria which define as adequate weight gain between 28 and 40 pounds for women with a BMI lower than 18.5, between 25 and 35 pounds for women with a BMI between 18.5 and 24.9, between 15 and 25 pounds for women with a BMI between 25 and 29.9, between 10 and 20 for women with BMI equal to or higher than 30. When adjusting for pre-pregnancy BMI, the coefficient on adequate weight gain becomes non-significant. [29]

Given the cross-sectional differences in risk factors between first and second generation immigrants (see Table 4) and the net effect of each risk factor on low birth weight, one can estimate, with a back of the envelope calculation, that if the behaviors did not worsen, the incidence of low-birth weight would have been 2% lower (with respect to the average low birth weight incidence). This confirms that behaviors did worsen, but very little compared to natives and not enough to explain the Hispanic paradox. [30] As mentioned above, I cannot test directly how changes in dietary practices affected changes in birth outcomes across generations. However, having information on these behaviors one could explain a larger share of the deterioration in birth outcomes.

These results show that policies aimed at preventing the adoption of less-healthy life styles that affect adult and child health, might have non-marginal benefits on the birth outcomes of future Americans. This is particularly important with respect to smoking. While tobacco use has declined dramatically in developed countries, it has increased and become more acceptable in less developed countries and among immigrants living in US, particularly among women. These facts explain the growing attention that the Tobacco Industry has devoted to target US immigrants (Acevedo-Garcia et al., 2004) who became an

[29] While not reported in Table 9, the coefficients of pre-pregnancy obesity is non-significant. If anything the sign of the coefficient shows that obesity might have some protective effect on the risk of low birth weight. However, the analysis of the relationship between maternal pre-pregnancy BMI and risk of low birth weight is problematic, given the non-linearity of the relationship between pre-pregnancy BMI and birth weight. Indeed, when considering as an alternative measure of infant health the likelihood of having a low APGAR score, I find evidence of lower APGAR scores among children of obese mothers.

[30] To conduct the back of the envelope calculation I used the generational differences and for each behavior I estimated the net effect on LBW, obtained in a unique regression including smoking, alcohol consumption, normal weight gain, gestational hypertension, chronic hypertension, chronic diabetes and controlling gender, race, fixed effects for child's year of birth and an the interaction of grand-mother county of residence and year mother's year of birth. Most of the change is explained by changes in smoking and gestational hypertension

appealing market. Policies aimed at discouraging smoking and in particular smoking during pregnancy might have direct effect on the incidence of low-birth weight of second-generation immigrants.

6 Conclusion

This paper confirms that while second generation Hispanics have lower incidence of low birth weight than children of native white mothers, this advantage shrinks substantially in the third generation. These stylized facts are mostly driven by the population of Mexican origin. With the help of a simple model of selection and intergenerational health transmission, I show that selection can explain the better birth outcomes of second-generation children compared to white natives, but if the only factor at play were selection, third generation birth outcomes would actually be worse. Considered the relatively weak intergenerational correlation in health and the less than full socio-economic assimilation, the model not only explains the deterioration, but over-predicts the convergence in the birth outcomes of the third-generation Hispanics, hence the paradox is reversed. However, accounting for the observed differences in risk factors (eg. tobacco and alcohol consumption during pregnancy, gestational hypertension) can account for about 80% of the reverse paradox.

While there is evidence of a generational worsening in risky behaviors, second-generation pregnant women conserve a significantly lower incidence of risk factors with respect to white natives. Between first and second generation behaviors did worsen, but very little compared to natives. The importance of behaviors is confirmed by the analysis of differences in the health convergence among second-generation Hispanics. Children of Hispanic women who are more likely to have culturally assimilated and show higher incidence of risky behaviors and risk factors are more likely to have poor birth outcomes. This holds true even after accounting for potential confounding factors focusing on a sub-sample of second-generation siblings and controlling for grand-mother fixed effects.

Overall, these findings show that the health trajectories observed among Hispanic descendants cannot be entirely explained by a pure mechanical statistical process. While there is evidence of a natural regression towards the mean, environmental and behavioral characteristics mediate the transmission of health across generations and together these factors explain and solve the Hispanic paradox.

Policies aimed at reducing disparities in access and quality of care, and at maintaining the healthy behaviors (substance abuse, dietary practices etc.) can importantly affect these health patterns. Given that second-generation births are overtaking migration as the main source of growth of the American population, the effects of these policies might be not negligible.

References

Acevedo-Garcia, Dolores, Elizabeth Barbeau, Jennifer Anne Bishop, Jocelyin Pan, and Karen Emmons (2004) 'Undoing an epidemiological paradox: The tobacco industry's targeting of us industry.' *American Journal of Public Health* 14(12), 2188–2193

Acevedo-Garcia, Dolores, Jocelyn Pan, Hee-Jin Jun, Theresa L. Osypuk, and Karen M. Emmons (2005) 'The effect of immigrant generation on smoking.' *Social Science and Medicine* 61, 1223–1242

Ahlburg, Dennis (1998) 'Intergenerational transmission of health.' *The American Economic Review, Paper and Proceeding* 88(2), 265–270

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach (2009) 'Inside the war on poverty: The impact of food stamps on birth outcomes.' *Review of Economics and Statistics* XCIII(2), 387403

Almond, Douglas, Kennet Y. Chai, and David S. Lee (2005) 'The cost of low birth weight.' *Quarterly Journal of Economics* 120(3), 1031–1083

Antecol, Heather, and Kelly Bedard (2006) ‘Unhealthy assimilation: Why do immigrants converge to american health status levels?’ *Demography* 43(2), 337–360

Bates, Lisa M., and Julien O. Teitler (2008) ‘Immigration and low birth weight in the us: the role of time and timing.’ *PAA, Mimeo*

Baumeister, Lisa, Kristen Marchi, Michelle Pearl, Ronald Williams, and Paula Braveman (2000) ‘The validity of information on race and ”hispanic ethnicity in california birth certificate data.’ *Health Services Research* 35(4), 869–883

Borjas, George J. (1988) ‘Self-selection and the earnings of immigrants’

Buekens, Pierre, Caitlin Canfield, Nicolas Padilla, Elia Lara Lona, and Rafael Lozano (2012) ‘Low birth weight in mexico: A systematic review.’ *Maternity and Child Health Journal*

Chiswick, Barry R., Yew Liang Lee, and Paul W. Miller (2008) ‘Immigrant selection system and immigrant health.’ *Contemporary Economic Policy* 26(4), 1074–3529

Classen, Timothy J. (2010) ‘Measures of the intergenerational transmission of body mass index between mothers and their children in the united states, 1981-2004.’ *Economics and Human Biology* 8(4), 1074–3529

Conley, Dalton, and Neil G. Bennett (2000) ‘Is biology destiny? birth weight and life chances.’ *American Sociological Review* 65(3), 458–476

Conway, Karen Smith, and Partha Deb (2005) ‘Is prenatal care really ineffective? or, is the devil in the distribution?’ *Journal of Health Economics* 24(3), 489 – 513

Cramer, James (1995) ‘Ratial and ethnic differences in birth weight: The role of income and financial assistance.’ *Demography* 32(2), 231–247

Currie, Janet (2007) ‘Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development.’ *Journal of Economic Literature* XLVII(1), 87–122

Currie, Janet, and Enrico Moretti (2007) 'Biology as destiny? short- and long-run determinants of intergenerational trasmission of birth weight.' *Journal of Labor Economics* 25(2), 231–263

Fenelon, Andrew (2012) 'Revisiting the hispanic paradox in the united states: the role of smoking'

Finch, Brian Karl, Nelson Lim, William Perez, and D. Phuong Do (2007) 'Toward a population health model of segmented assimilation: The case of low birth weight in los angeles.' *Sociological Perspectives* 50(3), 445–468

Forman, JP, MJ Stampfer, and GC Curhan (2009) 'Diet and lifestyle risk factors associated with incident hypertension in women.' *JAMA: The Journal of the American Medical Association* 302(4), 401–411

Fryer, Roland G., and Steven D. Levitt (2004) 'The causes and consequences of distinctively black names.' *The Quarterly Journal of Economics* 119(3), 767–805

Furtado, Delia, and Nikolaos Theodoropoulos (2009) 'Intermarriage and immigrant employment: The role of networks.' *B.E. Journal of Economic Analysis and Policy*

Furtado, Delia, and Stephen J. Trejo (2012) 'Interethnic marriages and their economic effectss.' *CReAM Discussion Paper*

Gonzalez, Carmen (2011) 'Diet for pregnancy and hypertension.' *http://www.livestrong.com/article/370938-diet-for-pregnancy-hypertension/*

Guendelman, Sylvia, and Barbara Abrams (1995) 'Dietary intake among mexican-american women: Generational differences and a comparison with non-hispanic women.' *American Journal of Public Health*

Gutmann, Myron P., W. Parker Frisbie, Peter DeTurk, and K. Stephen Blanchard (1998)

'Dating the origins of the epidemiologic paradox among mexican americans.' *Texas Population Research Center Papers*

Hummer, Robert A., Daniel A. Powers, Starling G. Pullum, Ginger L. Gossman, and W. Parker Frisbie (2007) 'Paradox found (again): Infant mortality among the mexican-origin population in the united states.' *Demography* 44(3), 441–457

Jasso, Guillermina, Douglas S. Massey, Mark R. Rosenzweig, and James P. Smith (2004) 'Immigrant health: Selectivity and acculturation.' *The Institute for Fiscal Studies Workig Paper* 04(23), 1–46

Kaiser, Lucia Lynn, and Lindsay Allen (2002) 'Position of the american dietetic association: Nutrition and lifestyle for a healthy pregnancy outcome.' *Journal of the American Dietetic Association* 102(10), 1479 – 1490

Kaushal, Neeraj (2008) 'Adversities of acculturation? prevalence of obesity among immigrants.' *Health Economics* 18, 291–303

Koenker, Roeger, and Kevin F. Hallock (2001) 'Quantile regression.' *Journal of Economic Perspectives* 15(4), 143–156

Kotelchuck, Milton (1994) 'An evaluation of the kessner adequacy of prenatal care index and a proposed adequacy of prenatal care utilization index.' *American Journal of Public Health* 84(9), 1486–1489

Osypuk, Theresa L., Lisa Bates, and Dolores Acevedo-Garcia (2010) 'Another mexican birthweight paradox? the role of residential enclaves and neighborhood poverty in the birthweight of mexican-origin infants.' *Social Science and Medicine* 70, 550–560

Palloni, Alberto, and Elizabeth Arias (2004) 'Paradox lost: Explaining the hispanic adult mortality advantage.' *Demography* 41(3), 385–415

Palloni, Alberto, and Jeffrey D. Morenoff (2001) 'Paradox lost: Explaining the hispanic adult mortality advantage.' *Annals of The New York Academy of Sciences* 954, 140–174

Roy, A.D. (1951) 'Some thoughts on the distribution of earnings.' *Oxford Economic Papers* 3(2), 135–146

Royer, Heather (2009) 'Separated at girth: Us twin estimates of the effects of birth weight.' *American Economic Journal: Applied Economics* 2(2), 49–85

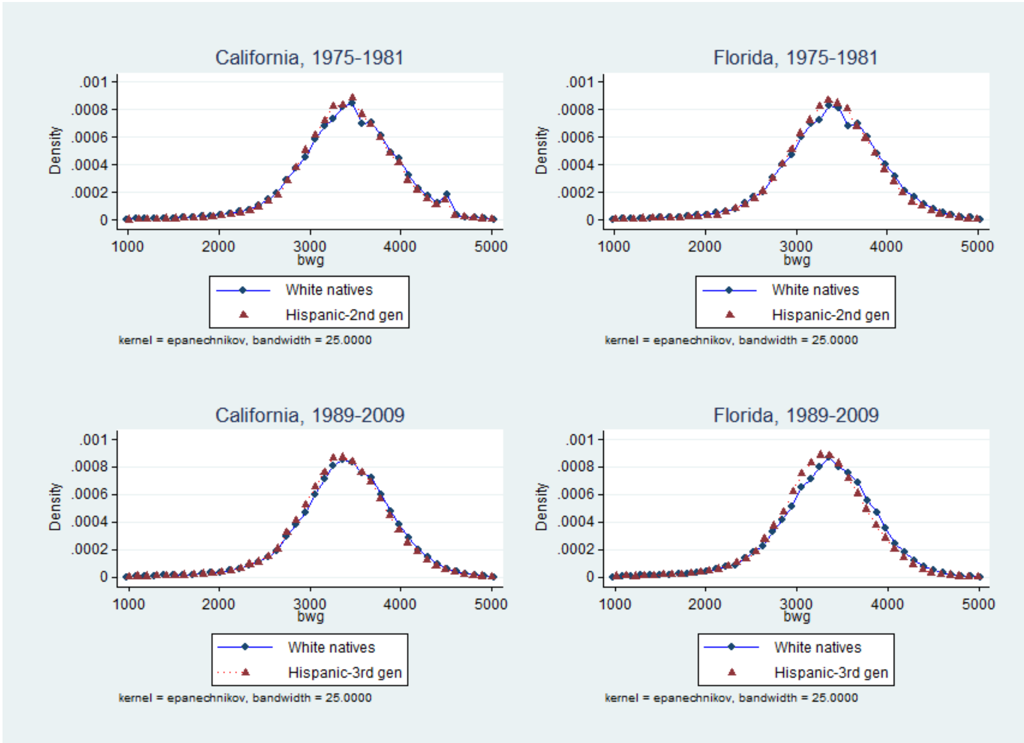
Shaw, Richard J., Kate E. Pickett, and Richard G. Wilkinson (2010) 'Ethnic density effects on birth outcomes and maternal smoking during pregnancy in the us linked birth and infant death data set.' *American Journal of Public Health* 100(4), 707–713

Shireen, Meher, and Duley Lelia (2006) 'Exercise or other physical activity for preventing pre-eclampsia and its complications.' *Cochrane Database of Systematic Reviews* (3), 1233–1241

Wang, Wendy (2012) 'The rise of intermarriage. rates, characteristics vary by race and gender.' *Pew Social and Demographic Trends*

Wilcox, J. Allen (2001) 'On the importance and the unimportance of birthweight.' *International Journal of Epidemiology* 30(6), 1233–1241

Figure 1: Birth Weight (grams) distribution, California and Florida



Notes - Source: California and Florida Vital Statistics, 1975-1981, 1989-2009.

Table 1: Matching quality. Women born in California and Florida, 1975-1981

	Observations			**Birth Weight (grams)**			**Low Birth Weight (below 2500 grams)**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample:	Overall	Linked	Matching rate	Overall	Linked	Non-linked	Overall	Linked	Non-linked
Overall	1,643,865	726,837	0.44	3,288	3,289	3,288	0.069	0.065	0.072
US born	1,273,023	558,921	0.44	3,283	3,277	3,288	0.073	0.070	0.075
Us born black	175,493	123,472	0.70	3,072	3,081	3,050	0.130	0.124	0.144
Us born white	1,097,530	435,449	0.40	3,317	3,333	3,307	0.064	0.055	0.070
Foreign born	370,842	167,916	0.45	3,304	3,328	3,285	0.055	0.046	0.061
Hispanic	231,741	124,267	0.54	3,327	3,345	3,305	0.051	0.044	0.060
Cuban	17,290	8,695	0.50	3,301	3,309	3,293	0.056	0.047	0.065
Mexican	198,264	109,661	0.55	3,341	3,356	3,323	0.049	0.042	0.057
Puerto Rican	16,187	5,911	0.37	3,175	3,186	3,168	0.074	0.064	0.080
1st income quartile	273,285	131,932	0.48	3,265	3,266	3,264	0.074	0.069	0.078
2nd income quartile	293,879	140,486	0.48	3,261	3,260	3,262	0.077	0.073	0.081
3rd income quartile	442,946	194,460	0.44	3,289	3,291	3,287	0.068	0.064	0.071
4th income quartile	367,580	144,510	0.39	3,321	3,324	3,320	0.060	0.056	0.062

Notes - Data are drawn from the California and Florida Vital Statistics, 1975-1981. The linked sample is composed of all the women born between 1975 and 1981 for whom I was able to link the information available at their birth to the birth records of their children born in California and Florida between 1989 and 2009.

Table 2: Hispanic Health Paradox in Birth Outcomes

	2^{nd} generation 1975-1981		3^{rd} generation 1989-2009	
	(1)	(2)	(3)	(4)
Dependent variable:	birth weight (grams)	low birth weight (< 2500 grams)	birth weight (grams)	low birth weight (<2500 grams)
Country of origin:				
Cuba	-0.285	-0.009***	-44.807***	0.001
	(4.266)	(0.002)	(7.242)	(0.003)
Mexico	22.647***	-0.016***	-8.078***	-0.004***
	(1.326)	(0.001)	(1.927)	(0.001)
Puerto Rico	-91.191***	0.005	-158.849***	0.022***
	(7.406)	(0.003)	(6.974)	(0.003)
Mean of Dep. Var.	3394.289	0.055	3347.184	0.058
std.err.	572.415	0.228	571.599	0.233
Observations	2,234,571	2,234,571	1,043,636	1,043,636

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

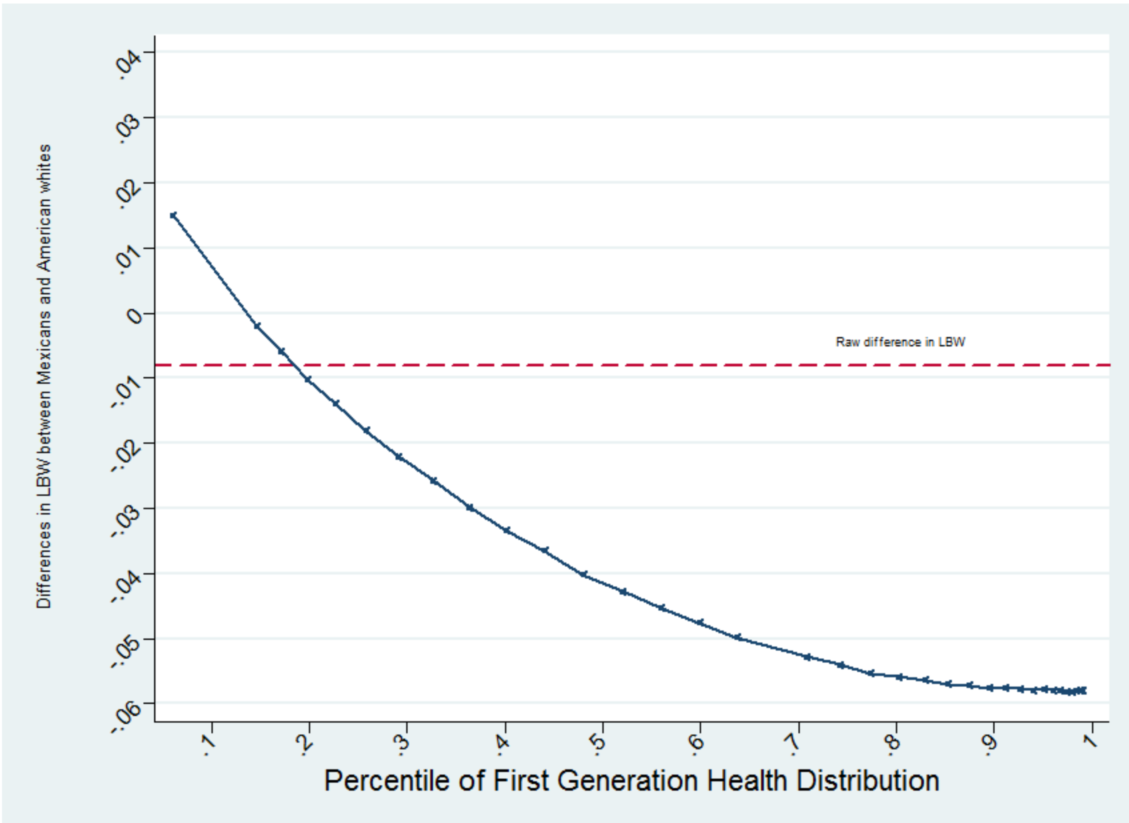
Notes - Data are drawn from the California and Florida Birth Records (1975-1981, 1989-2009). All the estimate include controls for child's gender, parity, type of birth, year of birth fixed effects, mother's age dummies, father's age (quadratic), mother's marital status,an indicator of adequacy of prenatal care, mother's education (4 groups dummies), father's education (4 group dummies), zip-code fixed effects, and indicators for missing variables: mother's age, father's age, mother's education, father's education, marital status, parity.

Table 3: Differences in the incidence of low birth weight between 1^{st}, 2^{nd} generation Mexicans and US white natives

Data Source:	CA-FL Vital Statistics			Natality Detail Data		
Years	1975-1981	1989-2009	1989-2009	1975-1981	1989-2004	1989-2004
	(1)	(2)	(3)	(4)	(5)	(6)
States:	CA-FL	CA-FL	FL^a	US	US^b	US
1st gen Mexican	-0.008***			-0.009***		
	(0.000)			(0.000)		
2nd gen Mexican		-0.001**	0.011***		0.001***	0.008***
		(0.001)	(0.004)		(0.000)	(0.000)
tobacco consumption			0.040***			0.040***
(during pregnancy)			(0.001)			(0.000)
alcohol consumption			0.018***			0.010***
(during pregnancy)			(0.006)			(0.001)
Gestational hypertension			0.114***			0.091***
			(0.003)			(0.001)
State F.E.	YES	YES	NO	YES	YES	YES
Year. F.E.	YES	YES	YES	YES	YES	YES
Observations	2,505,218	1,074,449	427,776	15,494,118	5,904,672	5,059,298

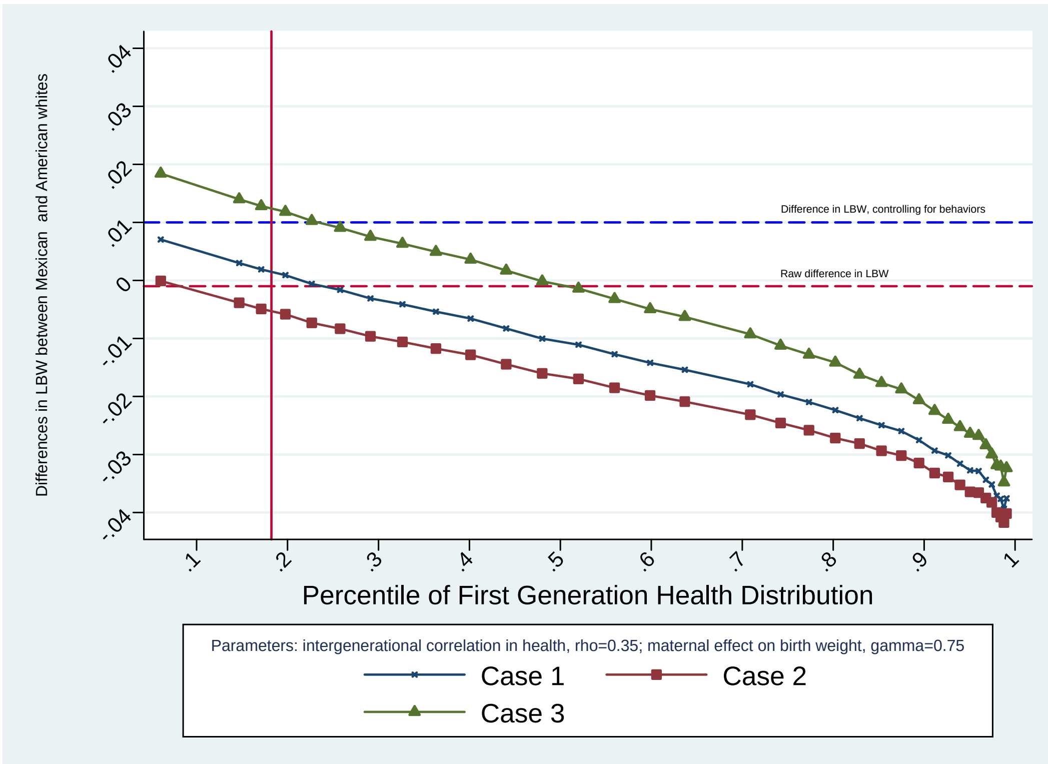
Notes - Data for columns 1-3 are drawn from California and Florida Vital Statistics (1975-1981; 1989-2009). Data for columns 4-6 are drawn from the Natality Detail Data (1975-1981; 1989-2004). The dependent variable is the incidence of low birth weight (birth weight < 2500 grams). [a]Data drawn from the California Vital Statistics do not contain information on these risk factors for the period under analysis. [b] The Natality Detail Data does not allow to distinguish second or higher generation since it does not contain information on parental nativity of the mothers.

Figure 2: Selection on health at migration and differences in the incidence of low-birth weight betweeen 2^{nd} generation Mexicans and white American ($\gamma = 0.75, \rho = 0.35$)



Notes - The plotted curve reports the predicted low-birth weight differences between 2^{nd} generation Mexicans and white Americans for each level of selection on health at migration assuming that the intergenerational correlation in health ρ is equal to 0.35 and the effect of maternal health on birth weight, γ, is equal to 0.75. The dashed line describes the observed raw difference in the incidence of low-birth weight between 2^{nd} generation Mexicans and white Americans born between 1975 and 1981, in California and Florida (see Table 3, col. 1).

Figure 3: Regression towards the mean and differences in the incidence of low-birth weight betweeen 3^{rd} generation Mexicans and white American ($\gamma = 0.75, \rho = 0.35$)



Notes - The plotted curves report the predicted low-birth weight differences between 3^{rd} generation Mexicans and white Americans for each level of selection on health at migration assuming that the intergenerational correlation in health ρ is equal to 0.35 and the effect of maternal health on birth weight, γ, is equal to 0.75. Each curve represents different assumptions on socio-economic and behavioral assimilation of the $2^{n}d$ generation. Case 1 illustrates a scenario in which second generation Mexicans fully assimilate socio-economically and behaviorally (μ_{MX_2}=0). Case 2 illustrates a scenario in which second generation Mexicans fully assimilate socio-economically but maintain their superior behaviors and health conditions (μ_{MX_2}=0.1). Case 3 assumes that Mexicans fully assimilate in behaviors but incorporates the estimated effect on birth weight of the observed socio-economic differences between second generation Mexicans and white Americans (less than full socio-economic assimilation, $\mu_{MX_2} = -0.15$). The lower dashed line (y=-0.001) describes the observed raw difference in the incidence of low-birth weight between 3^{rd} generation Mexicans and white Americans born between 1989 and 2009 in California and Florida (see Table 3, col. 2). The upper dashed line (y=0.010) describes the observed raw difference in the incidence of low-birth weight between 3^{rd} generation Mexicans and white Americans born between 1989 and 2009, after controlling for tobacco and alcohol consumption during pregnancy and gestational hypertension, (see Table 3, col. 3). The vertical solid line represents the level of selection that would explain the low birth weight difference observed in the data between 2^{nd} generation Mexicans, see Figure 2.

Table 4: Cross-sectional differences in Risk Factors between 1^{st}, 2^{nd} generation Hispanics and US white natives

		(1)	(2)	(3)	(4)	(5)	(6)
				Panel A			
Data	**Risky behaviors**	Natives N	1^{st} gen. Hispanics (H_1)	2^{nd} gen. Hispanics (H_2)	H_1-N	H_2-N	H_2-H_1
Source: Florida Vital Statistics		Mean	Mean	Mean			
Florida 1989-2009	Tobacco consumption	0.170	0.011	0.020	-0.159***	-0.150***	0.009***
	Observations	404,976	220,234	39,677	(0.001)	(0.002)	(0.001)
Florida 1989-2009	Alcohol consumption	0.004	0.001	0.001	-0.003***	-0.003***	0.000**
	Observations	407,053	219,357	39,622	(0.000)	(0.000)	(0.000)
	Health Conditions						
Florida 2004-2009	Hypertension	0.013	0.007	0.008	-0.006***	-0.005***	0.001
	Observations	172,435	108,230	19,249	(0.000)	(0.001)	(0.001)
Florida 2004-2009	Gestational hypertension	0.053	0.033	0.037	-0.020***	-0.016***	0.004***
	Observations	172,435	108,230	19,249	(0.001)	(0.002)	(0.001)
Florida 2004-2009	Diabetes	0.007	0.008	0.007	0.002***	-0.000	-0.002**
	Observations	172,435	108,230	19,249	(0.000)	(0.001)	(0.001)
Florida 2004-2009	Normal Weight Gain	0.301	0.332	0.305	0.032***	0.005	-0.027***
	Observations	158,276	95,488	17,017	(0.002)	(0.004)	(0.004)
Florida 2004-2009	Obesity	0.213	0.185	0.223	-0.028***	0.010***	0.038***
	Observations	160,699	95,488	17,180	(0.002)	(0.003)	(0.003)
				Panel B			
Data	**Risky behaviors**	Natives N	1^{st} gen. Hispanics (H_1)	2^{nd} gen Hispanics $(H_2)^a$	H_1-N	H_2-N	H_2-H_1
Source: Natality Detail Data		Mean	Mean	Mean			
US 1989-2004	Tobacco consumption	0.231	0.012	0.065	-0.219***	-0.166***	0.053***
	Observations	5,853,235	925,032	757,800	(0.000)	(0.000)	(0.000)
US 1989-2004	Alcohol consumption	0.009	0.003	0.008	-0.006***	-0.001***	0.005***
	Observations	5,981,571	929,651	762,412	(0.000)	(0.000)	(0.000)
	Health Conditions						
US 1989-2004	Gestational hypertension	0.046	0.022	0.030	-0.024***	-0.016***	0.008***
	Observations	6,379,454	1,432,244	1,160,089	(0.000)	(0.000)	(0.000)

Notes - Data for Panel A are drawn from the California and Florida Vital Statistics (1989-2009). For Panel B, data are drawn from the US Natality Detail Data (1989-2004).
[a] The Natality Detail Data does not allow to distinguish second or higher generation since it does not contain information on parental nativity of the mothers.

Table 5: Socio-economic differences between 1^{st} (2^{nd}) generation Hispanics and white native Americans

	(1) Natives (N)	(2) 1^{st} gen. Hispanics (H_1)	(3) $H_1 - N$	(4) Natives (N_2)^a	(5) 2^{nd} gen. Hispanics (H_2)	(6) $H_2 - N_2$
Panel A: Intergenerational assimilation in income-to-poverty ratio, Current Population Survey, Hispanics vs American white natives						
Years:	1994-2011	1994-2011	1994-2011	1994-2011	1994-2011	1994-2011
log (family income / poverty)	1.252	0.532	-0.726***	1.252	0.793	-0.459***
			(0.003)			(0.004)
Observations	1,278,594	166,296	1,444,890	1,278,594	56,828	1,335,422
Panel B: Intergenerational assimilation in income-to-poverty ratio, Current Population Survey, Mexicans vs American white natives						
log (family income / poverty)	1.252	0.508	-0.744***	1.252	0.831	-0.421***
			(0.003)			(0.006)
Observations	1,278,594	82,189	1,360,783	1,278,594	30,191	1,308,785
Panel C: Intergenerational assimilation in poverty (hospital-zip-code), California and Florida Vital Statistics, Hispanics vs American white natives						
Years:	1975-1981	1975-1981	1975-1981	1989-2009	1989-2009	1989-2009
high-poverty zip-code	0.222	0.392	0.170***	0.143	0.217	0.074***
			(0.001)			(0.001)
Observations	1,779,538	341,668	2,121,206	586,759	224,006	810,765
Panel D: Intergenerational assimilation in adequacy of prenatal care, California and Florida Vital Statistics, Hispanics vs American white natives						
Years:	1975-1981	1975-1981	1975-1981	1989-2009	1989-2009	1989-2009
prenatal care started 1^{st} trimester of pregnancy	0.489	0.311	-0.178***	0.856	0.850	-0.006***
			(0.001)			(0.001)
Observations	2,039,202	421,225	2,460,427	832,043	248,019	1,080,061

Notes - Data for Panel A and B are drawn from the Current Population Survey (1994-2011). Panel B restricts the analysis to natives and first (second) generation immigrants of Mexican origin. All statistics reported in Panel A and B are computed using analytical weights. In Panel C and D, the data are drawn from California and Florida Vital Statitics (1975-1981; 1989-2009). [a] The subscript 2 refers to the fact that in Panel B and C, I consider native women born between 1975 and 1981 who gave birth between 1989 and 2009. In Panel A, given that information on parental birth place is available in the CPS only since 1994, I include all natives in the working-age population (15-64) whose both parents were born in the US.

Table 6: Differences in birth outcomes of 3^{rd} generation Hispanics, within-family analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Hispanics	Hispanics	Hispanics	Hispanics	Hispanics	Hispanics	Hispanics	Non-Hispanic Whites
Panel A: Birth Weight (in grams)								
prenatal care	194.970***					182.003***	153.765***	163.668***
started in first trimester of pregnancy	(4.229)					(4.417)	(4.927)	(4.350)
high school drop-out		-34.236***				-1.134	4.698	2.794
		(2.989)				(3.747)	(5.121)	(5.392)
high-poverty (zipcode of residence)			-8.802***			-4.261	-8.954*	-4.621
			(3.051)			(3.518)	(5.009)	(5.596)
high poverty (hospital zip-code)				-17.840***		-10.671**	-11.073*	-22.109***
				(3.501)		(4.199)	(5.763)	(7.281)
non-Hispanic white father					-88.901***	-61.047***	-20.743**	50.689***
					(5.321)	(5.706)	(8.268)	(14.215)
Panel B: Incidence of Low Birth Weight								
prenatal care	-0.051***					-0.049***	-0.041***	-0.046***
started in first trimester of pregnancy	(0.002)					(0.002)	(0.002)	(0.002)
high school drop-out		0.003**				-0.004***	-0.004*	0.001
		(0.001)				(0.001)	(0.002)	(0.002)
high-poverty (zipcode of residence)			0.003***			0.003*	0.001	0.002
			(0.001)			(0.001)	(0.002)	(0.003)
high poverty (hospital zip-code)				0.007***		0.005***	0.006**	0.010***
				(0.001)		(0.002)	(0.003)	(0.003)
non-Hispanic white father					0.018***	0.014***	0.008**	-0.011
					(0.002)	(0.003)	(0.004)	(0.007)
Socio-demographic controls	NO	NO	NO	NO	NO	YES	YES	YES
Mother's birth weight	NO	NO	NO	NO	NO	YES	YES	YES
Grand-mother F.E.	NO	NO	NO	NO	NO	NO	YES	YES
Observations	228,461	232,501	225,373	222,872	233,440	201,754	201,754	336,439

Notes - Data are drawn from the California and Florida Vital Statistics (1975-1981, 1989-2009). The sample is restricted to third-generation Hispanics born between 1989 and 2009 in columns 1-7 and to third-generation white natives born between 1989 and 2009 in column 8. Socio-demographic controls include 3^{rd} generation child's gender, mother's birth weight (LBW), dummies for the interaction of county and year of birth of second generation children (mothers), first generation's (grand-mothers) age (at delivery) dummies, second generation's age (at delivery) dummies, second generation parity, poverty rate in zip-code of birth of second generation, marital status, father's and mother's education (4 groups), parity, dummies for the interaction of county and year of birth of third generation children, indicators for missing information on first generation's (grand-mothers) age (at delivery) (FL), father's education and age. Columns 7 and 8 include grand-mother fixed effects. Standard errors are clustered at the grand-mother level.

Table 7: Ethnic inter-marriage and low birth weight risk factors among among 2^{nd} generation Hispanic mothers

Panel A: Florida Vital Statistics					
Dependent Variables	tobacco consumption	alcohol consumption	gestational hypertension	low birth weight	low birth weight
Years	1989-2009	1989-2009	2004-2009	1989-2009	2004-2009
non-Hispanic white father	0.0250***	0.00129***	0.00646**	0.00595**	0.00303
	(0.00147)	(0.000385)	(0.00291)	(0.00266)	(0.00400)
tobacco consumption during pregnancy					0.0383**
					(0.0153)
alcohol consumption during pregnancy					0.0676
					(0.0522)
Gestational hypertension					0.193***
					(0.00985)
Observations	39,662	39,608	19,249	39,703	19,084

Panel B: Natality Detail Data					
Dependent Variables	tobacco consumption	alcohol consumption	gestational hypertension	low birth weight	low birth weight
	(1)	(2)	(3)	(4)	(5)
Years	1989-2004	1989-2004	1989-2004	1989-2004	1989-2004
non-Hispanic white father	0.0666***	0.00711***	0.00397***	0.00859***	0.00456***
	(0.000640)	(0.000231)	(0.000372)	(0.000552)	(0.000689)
tobacco consumption during pregnancy					0.0337***
					(0.00138)
alcohol consumption during pregnancy					0.0136***
					(0.00381)
Gestational hypertension					0.119***
					(0.00184)
Observations	626,017	630,675	956,882	967,194	617,255

Notes - Data are drawn from the Florida Birth Master File (1971-1985, 1989-2009). The sample is restricted to second-generation Hispanic mothers born in Florida between 1971 and 1985. In Panel A, the sample is restricted to years 2004-2009 when using information on gestational hypertension which is available since 2004.

Table 8: Ethnic inter-marriage and low birth weight risk factors among among 2^{nd} generation Hispanic mothers, within-family analysis

	(1)	(2)	(3)	(4)
Dependent Variable:	Tobacco consumption during pregnancy			
non-Hispanic white father	0.025***	0.020***	0.006**	0.006**
	(0.002)	(0.002)	(0.003)	(0.003)
Mean of dep. Var.	0.07			
St.Er.	0.25			
Observations	39,662	38,952	38,952	37,243
Dependent Variable:	Alcohol consumption during pregnancy			
non-Hispanic white father	0.001***	0.001***	0.001*	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Mean of dep. Var.	0.07			
St.Er.	0.25			
Observations	39,608	38,900	38,900	37,368
Dependent Variable:	Gestational hypertension			
non-Hispanic white father	0.006**	0.005	0.010	0.010
	(0.003)	(0.003)	(0.007)	(0.007)
Mean of dep. Var.	0.07			
St.Er.	0.25			
Observations	19,249	19,044	19,044	18,346
Child's gender, race		YES	YES	YES
Mother low birth weight		YES	YES	YES
Mother's and child's year of birth F.E.		YES	YES	YES
Mother's county of residence F.E.		YES	YES	YES
Grandmother's age at delivery		YES	YES	YES
Grandmother F.E.			YES	YES
County-year at birth of mother and child F.E.				YES
Mother's education, age dummies, parity, marital status				YES
adequate care				YES
Poverty of the zipcode of residence at mother's birth				YES

Notes - Data are drawn from the Florida Birth Master File (1971-1985, 1989-2009). The sample is restricted to second-generation Hispanic mothers born in Florida between 1971 and 1985.

Table 9: The effect of health risk factors on low-birth weight, Florida 1989-2009, within-family analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Years		1989-2009			2004-2009	
tobacco consumption	0.01835***	0.02797***	0.01852***	0.03558***	0.02895***	0.01849***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.007)
alcohol consumption	0.04057***	0.00494	0.00608	0.06372***	0.02059	0.00521
	(0.006)	(0.008)	(0.010)	(0.013)	(0.014)	(0.028)
gestational hypertension				0.14340***	0.13737***	0.08788***
				(0.004)	(0.004)	(0.007)
chronic hypertension				0.10802***	0.08572***	0.02667**
				(0.006)	(0.007)	(0.013)
diabetes				0.03163***	0.01798**	0.00456
				(0.008)	(0.008)	(0.015)
normal weight gain	-0.03255***	-0.02140***	-0.01875***			
	(0.001)	(0.001)	(0.001)			
normal weight gain (adjusted for BMI)				0.00063	0.00200	-0.00154
				(0.001)	(0.001)	(0.002)
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes
Grandmother F.E.	No	No	Yes	No	No	Yes
Observations	793,005	542,517	542,517	284,528	214,711	214,711
Avg. Incidence of LBW	0.07					
Std.dev.	0.25					

Notes - Data are drawn from Florida Birth Master Dataset (1989-2009). The sample is restricted the children of mothers born between 1971 and 1985 in Florida. Socio-demographic controls include 3^{rd} generation child's gender, race, country of origin dummies, mother's birth weight (LBW), dummies for the interaction of county and year of birth of second generation children (mothers), second generation's age (at delivery) dummies, second generation parity, poverty rate in zip-code of birth of second generation, marital status, father's and mother's education (4 groups), parity, dummies for the interaction of county and year of birth of third generation children, father's education and age. Standard errors are clustered at the grand-mother level.

Appendix A

Table A1: Availability of Variables in CA and FL, across years

	CA Sample	FL Sample
Birth weight	1975-2009	1971-2009
APGAR score	1978, 2008-2009	1989-2009
infant death	N/A	1971-2009
gestational length	1978-2009	1971-2009
Mother's state of birth	1975-2009	1975-2009
Child's race	1975-2009	1975-2009
Mother's race	1975-2009	1975-2009
Father's race	1975-2009	1975-2009
Sex	1975-2009	1975-2009
Mother's age	1975-2009	1989-2009
month prenatal care started	1975-2009	2004-2009
number of prenatal visits	2989-2009	1975-2009
parity	1975-2009	1975-2009
plural	1975-2009	1975-2009
Father's age	1975-2009	1989-2009
Zip-code of residence	1989-2009	1975-2009
Zip-code of the hospital	1975-2009	N/A
Mother's date of birth	1989-2009	1989-2009
Family name of mother	1975-2009	1975-2009
First name of mother	1989-2009	1975-2009
Full name of father	N/A	1975-2009
Full name of child	1975-2009	1975-2009
Exact residential address of mother	1997-2009	1975-2009
Exact residential address of father	N/A	1975-2009
Smoking during pregnancy	N/A	1989-2009
Alcohol use during pregnancy	N/A	1989-2009
Weight gain	N/A	1989-2009
Pre-pregnancy BMI	N/A	2004-2009
Hypertension	N/A	2004-2009
Gestational Hypertension	N/A	2004-2009
Diabetes	N/A	2004-2009

Table A2: Hispanic Health Paradox, Immigrant-Native differences in birth outcomes, 1975-1981

	(1)	(2)	(3)	(4)	(5)	(6)
Birth Weight (in grams)						
Cuba	35.407*	-13.266	-12.378***	-7.799*	-7.623	-5.529
	(20.246)	(15.271)	(4.431)	(4.520)	(4.728)	(4.728)
Mexico	-2.129*	3.972*	3.883***	19.195***	21.025***	19.866***
	(1.123)	(2.209)	(1.374)	(1.283)	(1.449)	(1.448)
Puerto Rico	-74.826***	-113.207***	-113.355***	-95.185***	-111.611***	-104.625***
	(18.714)	(15.251)	(8.748)	(8.044)	(8.048)	(8.052)
Mean of dep.var.	3382.816					
st.dev.	559.6369					
Incidece of Low Birth Weight						
Cuba	-0.004	0.003	-0.005***	-0.005***	-0.006***	-0.007***
	(0.012)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Mexico	-0.008***	-0.011***	-0.010***	-0.016***	-0.015***	-0.015***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Puerto Rico	0.014	0.018***	0.014***	0.006	0.009***	0.007**
	(0.011)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
Mean of dep.var.	0.054					
st.dev.	0.225					
Observations	2,560,258	2,474,983	2,473,231	2,474,983	2,312,229	2,250,517
Socio-demographic characteristics		YES	YES	YES	YES	YES
County F.E.			YES	YES	YES	YES
Hospital F.E.				YES		
Zip-code F.E.					YES	YES
Father's age and race						YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes - Data are drawn from the California and Florida Birth Certificates (1975-1981; 1989-2009). Socio-demographic controls include an indicator for reported unmarried status, adequacy of care, parity, type of birth (twin), gender, race, year fixed effects, age dummies. Zip-code refers to the zip-code of the hospital for California and to the zip-code of residence in Florida. The sample is restricted to descendants of 1^{st} generation immigrants from Cuba, Puerto Rico and Mexico and 1^{st} generation US white natives. The reference group for the reported coefficients are children of US-born 1^{st} generation. Column 6 includes mother's education available, father's age and father's race. Estimates include an indicator for missing information on mother's age, marital status, parity and type of birth (column 2-6), father's age,and mother's education (column 6).

Table A3: Hispanic Health Paradox, Immigrant-Native differences in birth outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Birth Weight (in grams)						
Cuba	-50.778***	-69.649***	-54.407***	-41.935***	-48.651***	-48.651***
	(5.501)	(5.468)	(6.599)	(7.113)	(7.438)	(7.438)
Mexico	-27.983***	-22.547***	-13.278***	-7.731***	-7.211***	-7.211***
	(1.735)	(1.753)	(1.852)	(1.913)	(1.915)	(1.915)
Puerto Rico	-172.970***	-162.344***	-163.704***	-154.051***	-156.097***	-156.097***
	(6.884)	(6.890)	(6.961)	(6.945)	(6.980)	(6.980)
Mean of dep. var.	3315.811					
st.dev.	575.201					
Incidece of Low Birth Weight						
Cuba	-0.001	0.003	0.003	0.000	0.002	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Mexico	-0.001**	-0.002***	-0.003***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Puerto Rico	0.028***	0.025***	0.026***	0.021***	0.022***	0.022***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mean of dep. var.	0.054					
st.dev.	.225					
Observations	1,102,271	1,064,209	1,063,271	1,064,209	1,054,966	1,054,966
Socio-demographic characteristics		YES	YES	YES	YES	YES
County F.E.			YES	YES	YES	YES
Hospital F.E.				YES		
Zip-code F.E.					YES	YES
Father's age and race						YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes - Data are drawn from the California and Florida Birth Certificates (1975-1981; 1989-2009). Socio-demographic controls include an indicator for reported unmarried status, adequacy of care, parity, type of birth (twin), gender, race, year fixed effects, age dummies. Zip-code refers to the zip-code of the hospital for California and to the zip-code of residence in Florida. The sample is restricted to descendants of 1^{st} generation immigrants from Cuba, Puerto Rico and Mexico and 1^{st} generation US white natives. The reference group for the reported coefficients are children of US-born 1^{st} generation. Column 6 includes mother's education available, father's age and father's race. Estimates include an indicator for missing information on mother's age, marital status, parity and type of birth (column 2-6), father's age,and mother's education (column 6).

Table A4: Summary Statistics

	1975-1981 Birth Weight	1989-2009 Birth Weight	1975-1981 LBW	1989-2009 LBW	1975-1981 Macrosomic	1989-2009 Macrosomic
Overall Sample	3394.289	3347.184	0.055	0.058	0.119	0.097
	(572.416)	(571.599)	(0.228)	(0.233)	(0.323)	(0.296)
White natives	3394.570	3353.500	0.057	0.058	0.121	0.100
	(576.791)	(574.696)	(0.231)	(0.235)	(0.326)	(0.300)
Cuba	3369.977	3272.216	0.052	0.067	0.097	0.066
	(553.112)	(552.060)	(0.222)	(0.250)	(0.297)	(0.247)
Mexico	3397.320	3336.142	0.047	0.054	0.110	0.089
	(549.328)	(559.711)	(0.211)	(0.225)	(0.313)	(0.285)
Puerto Rico	3269.531	3180.003	0.071	0.087	0.077	0.052
	(575.005)	(570.753)	0.257	(0.281)	(0.266)	(0.222)
California	3399.949	3355.867	0.053	0.055	0.120	0.100
	(568.312)	(570.616)	(0.224)	(0.228)	(0.325)	(0.300)
Florida	3374.915	3313.358	0.062	0.070	0.115	0.085
	585.829	(574.165)	0.240	(0.254)	0.319	(0.279)

Table A5: Quantile Regression, Differences in Birth Weight, 1975-1981

	(1) OLS	(2) QR5	(3) QR10	(4) QR25	(5) QR50	(6) QR75	(7) QR90	(8) QR95
Cuba	-23.943***	33.143***	18.484***	-3.366	-1.549	-93.501***	-60.231***	-64.930***
	(3.266)	(9.587)	(5.775)	(3.282)	(3.175)	(3.912)	(4.825)	(6.916)
Mexico	4.704***	75.208***	49.067***	16.921***	3.447***	-51.203***	-30.332***	-32.131***
	(0.994)	(2.929)	(1.763)	(1.000)	(0.966)	(1.164)	(1.463)	(2.096)
Puerto Rico	-106.635***	-58.545***	-83.244***	-101.097***	-102.226***	-74.032***	-118.722***	-132.823***
	(6.739)	(19.807)	(11.932)	(6.775)	(6.553)	(7.427)	(9.944)	(14.255)
Observations	2,447,951	2,447,951	2,447,951	2,447,951	2,447,951	2,447,951	2,447,951	2,447,951

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes - Data are drawn from California and Florida Birth Statistical Master Files (1975-1981). All the regression include controls for gender, marital status, adequacy of prenatal care, parity, type of birth, 4 educational dummies, a quadratic in age state and year of birth fixed effects

Table A6: Cross-Sectional Differences in Smoking Consumption among Hispanic Immigrants, by Generation and Intermarriage

Smoking Behavior	**1st gen (1989-2009)**	**2nd gen (1989-2009)**	**Non-Hispanic White natives**
% of Women Smoking During Pregnancy	0.011	0.020	0.114
Intramarriage			
Married to a Hispanic husband	0.007	0.011	
Married to a Non-Hispanic husband	0.025	0.035	
Education			
High-school dropouts	0.012	0.048	
High-school graduates	0.012	0.020	
Some College	0.011	0.013	
College or more	0.006	0.006	
By country of origin			
Cuba	0.011	0.017	
Mexico	0.003	0.013	
Puerto Rico	0.029	0.043	

Notes - Data are drawn from Florida Vital Statistics (1989-2009). The sample is restricted to mothers born between 1971 and 1985.

Table A7: Cross-Sectional Differences in Alcohol Consumption among Hispanic Immigrants, by Generation and Intermarriage

Hypertension during pregnancy	**1st gen (1989-2009)**	**2nd gen (1989-2009)**	**Natives**
% of Women drinking during pregnancy	0.0010	0.0015	0.0042
Intramarriage			
Married to a Hispanic husband	0.0007	0.0010	
Married to a Non-Hispanic husband	0.0022	0.0023	
Education			
High-school dropouts	0.0013	0.0019	
High-school graduates	0.0007	0.0018	
Some College	0.0011	0.0008	
College or more	0.0008	0.0013	
By country of origin			
Cuba	0.0007	0.0014	
Mexico	0.0008	0.0007	
Puerto Rico	0.0021	0.0026	

Notes - Data are drawn from Florida Vital Statistics (1989-2009). The sample is restricted to mothers born between 1971 and 1985.

Table A8: Cross-Sectional Differences in Hypertension among Hispanic Immigrants, by Generation and Intermarriage

Hypertension during pregnancy	**1st gen (1989-2009)**	**2nd gen (1989-2009)**	**Natives**
% of Women Hypertense (pregnancy)	0.033	0.037	0.0550
Intramarriage			
Married to a Hispanic husband	0.033	0.035	
Married to a Non-Hispanic husband	0.033	0.041	
Education			
High-school dropouts	0.028	0.032	
High-school graduates	0.036	0.035	
Some College	0.040	0.043	
College or more	0.037	0.038	
By country of origin			
Cuba	0.036	0.035	
Mexico	0.028	0.047	
Puerto Rico	0.044	0.042	

Notes - Data are drawn from Florida Vital Statistics (1989-2009). The sample is restricted to mothers born between 1971 and 1985.

Table A9: Ethnic inter-marriage and low birth weight risk factors among among 2^{nd} generation Hispanic mothers

Dependent Variables	tobacco consumption	alcohol consumption	gestational hypertension	low birth weight	low birth weight	low birth weight	low birth weight	low birth weight
Panel A: Florida Vital Statistics								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years	1989-2009	1989-2009	2004-2009	1989-2009	1989-2009	1989-2009	2004-2009	2004-2009
non-Hispanic white father	-0.0234***	-0.00129***	-0.00660**	-0.00630**	-0.00545*	-0.00548*	-0.00464	-0.00263
	(0.00203)	(0.000432)	(0.00319)	(0.00304)	(0.00303)	(0.00304)	(0.00459)	(0.00455)
tobacco consumption					0.0323***	0.0318***		0.0384*
					(0.0114)	(0.0113)		(0.0198)
alcohol consumption						0.0277		0.0676
						(0.0437)		(0.0736)
gestational hypertension								0.193***
								(0.0184)
Observations	39,662	39,608	19,249	39,703	39,537	39,467	19,249	19,084
Panel B: Natality Detail Data, US								
Years	1989-2004	1989-2004	1989-2004	1989-2004	1989-2004	1989-2004	1989-2004	1989-2004
non-Hispanic white father	0.0593***	0.00661***	0.00397***	0.00859***	0.00456***	0.00398***	0.00707***	0.00456***
	(0.000435)	(0.000157)	(0.000372)	(0.000552)	(0.000565)	(0.000567)	(0.000685)	(0.000689)
	(0.000640)	(0.000231)	(0.000372)	(0.000552)	(0.000567)	(0.000685)	(0.000685)	(0.000689)
tobacco consumption					0.0341***	0.0342***		0.0337***
					(0.00130)	(0.00133)		(0.00138)
alcohol consumption						0.0145***		0.0136***
						(0.00366)		(0.00381)
gestational hypertension						0.134***		0.119***
Observations	967,765	967,765	956,882	967,194	967,194	956,360	617,255	617,255

Notes - Data are drawn from the Florida Birth Master File (1971-1985, 1989-2009). The sample is restricted to second-generation Hispanic mothers born in Florida between 1971 and 1985. In Panel A, the sample is restricted to years 2004-2009 when using information on gestational hypertension which is available since 2004. In columns 5 and 6, I used missing indicators for hypertension (Panel A) and for tobacco and alcohol consumption (Panel B).