Intermarriage and the Unhealthy Assimilation of Immigrant Descendants

Osea Giuntella*

University of Oxford

Draft

February 24, 2015

Abstract

This paper studies the effects of assimilation on the health of Hispanics in the US using ethnic intermarriage as a metric of acculturation. Despite their higher socioeconomic status, third-generation children of second-generation intermarried Hispanic women are more likely to have poor health at birth, even after I account for second-generation health at birth, employ only within-family variations in the extent of assimilation, and consider the endogeneity of intermarriage. These results do not appear to be driven by father's selectivity nor by individual unobservable characteristics associated with intermarriage and largely reflect the higher incidence of risky behaviors among intermarried Hispanic women.

Keywords: Intermarriage, Birth Outcomes, Risky behaviors

JEL Classification Numbers: I10, J15

1 Introduction

Despite being poorer and having lower rates of education and employment, recent Hispanic immigrants are healthier than the average American. They tend to live longer, have a lower incidence of heart disease and cancer, and give birth to healthier children than do US-born natives. However, this population's health advantage erodes with time spent in the US, despite socio-economic assimilation. Previous scholars have referred to this phenomenon as the Hispanic health paradox (HHP) because of the expected positive socio-economic gradient in health.

A large amount of literature has documented these health trajectories and pointed to immigrant selectivity and sociocultural protection as the main explanations for the immigrants initial health advantage and the subsequent health deterioration (Giuntella, 2013; Riosmena et al., 2013; Elder et al., 2013; Ullmann et al., 2011; Antecol and Bedard, 2006; Jasso et al., 2004). Yet, we know relatively little about the heterogeneity of this assimilation process in the immigrant population, as previous studies focused on immigrant-native differences and were based on small sample size cross-sectional data. This paper focuses on second-generation Hispanics and analyzes the relationship between a metric of cultural assimilation, ethnic intermarriage, and the birth outcomes of their offspring, the third generation.

As of 2012, Hispanics comprise 17% of the US population and approximately 25% of all births. Several studies provide evidence of a positive relationship between health at birth and adult health and socio-economic outcomes, and a negative relationship between infant health and health care costs (Conley and Bennett, 2000; Case et al., 2005; Almond et al., 2005; Currie and Moretti, 2007). Children of immigrant mothers, despite having a lower socio-economic status, have better birth outcomes than native children, but second generation immigrant women bear less healthy children.[1] For these reasons, there is growing concern

---

[1]In a companion paper (Giuntella, 2013), I show that modest positive selection on health at the time of migration can account for the initial advantage in birth outcomes of second-generation Mexicans. Moreover, a simple process of regression toward the mean reverses the apparent paradox and predicts a greater deterioration than that observed in the data. Yet, the convergence would occur more quickly if immigrants had not socio-economically assimilated or if they had fully assimilated in the incidence of risky behaviors during

that the deterioration of birth outcomes in a large and growing segment of the American population may have important consequences on fiscal budgets (Kaushal and Kaestner, 2010) and, more generally, on intergenerational mobility (Currie, 2011). Furthermore, in light of the enduring debate on whether to enact more or less assimilation oriented policies, it is relevant to investigate whether the acculturation process - which is usually associated with positive economic outcomes (Nekby, 2010; Bisin et al., 2011) - may be deleterious to the health of immigrants and their offspring (Finch et al., 2007). Finally, in accordance with the "epidemiological approach" (Fernandez, 2011), studying the effects of assimilation among immigrant descendants helps to further the understanding of the health risks associated with urbanization, the adoption of sedentary habits and the westernization of habits in developing countries that are undergoing major health transitions (Marmot and Syme, 1976; Delavari et al., 2013).

There has been a growing interest among economists in studying the determinants of intermarriage (Bisin and Verdier, 2000; Bisin et al., 2004) and the effect of family structure on the economic assimilation of immigrants (Furtado and Trejo, 2012; Meng and Gregory, 2005; Eckstein and Weiss, 2002; Baker and Benjamin, 1997). Interethnic intermarriage is viewed by many as the final step in the assimilation process as it is importantly related with the resilience of traditional behaviors and the adoption of the social norms of the host country (Gordon, 1964; Qian and Lichter, 2007; Sassler, 2005). Therefore, intermarriage is often used as a proxy for cultural assimilation. Moreover, marrying a native has direct effects on language abilities, knowledge of the host country's social norms and labor market conditions, and the ethnic composition of one's social circle. Most studies focus on first-generation immigrants and suggest that marrying a native is positively associated with the success of immigrants in the labor market. These studies have provided evidence that intermarriage may improve job prospects and increase the rate of economic assimilation by facilitating the adoption of host country customs, improving language proficiency, and expanding social

_____

pregnancy.

connections (Furtado and Theodoropoulos, 2010). Social norms and ethnic identity influence the acceptability of risky behaviors in a network (Unger et al., 2001; Bethel and Schenker, 2005; Fenelon, 2013). Intermarriage may therefore have direct effects on risky behaviors and health by changing the reference group and the relevant social norms affecting individual behavior. However, we know relatively little about the relationship between intermarriage and health assimilation. It has been shown that US-born individuals with two foreign-born parents are less likely to smoke than children of inter-married couples (Acevedo-Garcia et al., 2005). Yet, to the best of my knowledge, this is the first paper to analyze the effects of intermarriage on health. In addition, this paper focuses on second-generation Hispanics, while previous studies analyzing the effects of intermarriage on labor market outcomes mostly focused on the first-generation.[2]

Exploiting unique administrative birth records data, I analyze the birth outcomes of third-generation Hispanics born in California and Florida, two of the top immigrant destination states in the US. Specifically, I obtained confidential information on names and dates of birth that allows me to link the birth records of mothers born in California and Florida between 1970 and 1985 to those of their off-spring born between 1989 and 2009. Following an approach similar to Currie and Moretti (2007), I identify second-generation siblings and link their records to those of the third-generation.

To study the effects of acculturation on the health of immigrant descendants, I use the ethnicity of the father to identify women who are in exogamous relationships (Chiswick and Houseworth, 2011).[3] Marriage choice is endogenous and unobserved factors influence who we marry (Nekby, 2010). Individuals who choose to intermarry are not a random sample of second-generation Hispanics, and a natural concern is that the intermarriage decision may be correlated with unobserved determinants of health. I use a variety of techniques to address this endogeneity issue. As I am able to identify a subset of second-generation sisters,

---

[2]See Furtado and Trejo (2012) for a comprehensive review of previous literature on interethnic marriages and their economic effects.

[3]Throughout the paper I use the term intermarriage loosely because I only use the father's ethnicity regardless of the actual marital status.

I can use grandmother fixed effects to control for the time-invariant characteristics of the family across the births of the two siblings and analyze within family differences in socio-economic characteristics and risky behaviors. Nevertheless, as in any study using siblings, there may still be systematic differences across siblings that may confound the relationship of interest. Therefore, I also adopt an instrumental variable approach and use marriage market conditions to instrument for intermarriage.

I find that third-generation children of Hispanic women married with non-Hispanic men are 9% more likely to be of low birth weight than children of endogamously married Hispanics. However, there is no evidence of significant effects of intermarriage on non-Hispanic U.S. born white women, nor on second-generation Hispanic women marrying into other (non-Hispanic) immigrant groups, suggesting that the results are not driven by unobservable characteristics associated with the decision to marry outside ones ethnicity, nor by unobservable traits of Hispanic fathers. The higher incidence of low birth weight among children of second-generation Hispanic women in exogamous relationships is surprising because intermarriage is, consistent with previous literature (Wang, 2012; Furtado and Trejo, 2012; Meng and Gregory, 2005), positively associated with socio-economic outcomes and because of the positive socio-economic gradient in health (Case et al., 2002). Second-generation hispanic women who marry non-Hispanics men live in better neighborhoods, have higher education and are married to husbands with higher socio-economic status. As third-generation birth outcomes correlate significantly with quality of care, socio-economic status, and risky behaviors, we would expect children of intermarried Hispanics to have, if anything, healthier birth outcomes. 2SLS estimates confirm the main findings that children of intermarried couples are more likely to be of low birth weight and, consistent with the existence of a positive socio-economic gradient in health, are larger than the OLS estimates. On the contrary, however, I find no evidence of significant effects of intermarriage on birth outcomes when focusing on non-Hispanic native mothers, suggesting that the intermarriage "health penalty" among children of Hispanic mothers is not driven by confounding factors that could be peculiar to

Hispanic fathers. In addition, there is no evidence of significant effects of intermarriage on second-generation Hispanics who intermarried into Hispanics coming from a different country (e.g., women of Mexican origin married to men of Puerto Rican origin). This evidence suggests that the negative effects of intermarriage on third-generation Hispanics are not explained by unobservable individual characteristics associated with exogamous marriages.

The "health penalty" observed among children of exogamously married second-generation Hispanic mothers can largely be attributed to their higher incidence of risk factors (such as higher rates of smoking, alcohol consumption, and hypertension) comparerd to their endogamously married counterparts. Intermarrried second-generation Hispanic women are much less likely to maintain the health-protective behaviors and conditions that characterize the first-generation immigrant Hispanic mothers. The results point in the same direction when alternative metrics of cultural assimilation are used (e.g., the ethnic content of the third-generation child's first name) and are robust to different model specifications. Overall, these findings suggest that, despite its positive effects on economic outcomes, cultural assimilation may accelerate the adoption of riskier behaviors and, therefore, have negative effects on the health of immigrant descendants.

The paper is organized in the following manner. Section 2 discusses the data and analyzes the determinants of intermarriage among second-generation Hispanics. Section 3 analyzes the effects of intermarriage on the birth outcomes of third generation immigrants. Robustness checks are presented in Section 4. I analyze the possible mechanisms underlying the negative effect of intermarriage on third-generation birth outcomes in Section 5. Section 6 presents the concluding remarks.

2   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 of all children born in the years 1970–1985 (1970–1981 in California and 1971-1985 in Florida) and 1989–2009.[4] In the sample, I define first-generation all women who were born outside the U.S. and delivered in the U.S. between 1970 and 1985; second-generation their children born between 1970 and 1985; third-generation their grandchildren born between 1989 and 2009. Children are classified based on mother's country of origin, as information on father's country of origin is not available before 1989 and often missing thereafter. Thus the sample of second-generation children includes children born to first-generation Hispanic women and U.S. born fathers. As information on the country of origin is reported in both states only for the major sending countries, I restrict the second-generation sample to children of Cuban, Mexican, and Puerto Rican mothers.[5]

Unfortunately, I do not have information on the country of origin of the father that spans the entire period under study and in both states. Thus, as mentioned earlier, the sample includes second-generation children born to Hispanic mothers and U.S. born fathers. For the same reason, the main variable of interest "intermarriage" is based on the father's ethnicity, regardless of nativity and marital status, and is defined as the union of a second-generation Hispanic woman with a non-Hispanic man.[6] Recent studies (Duncan and Trejo, 2007, 2009, 2012, 2011) assessed the importance of selective ethnic attrition among Mexican Americans and pointed out that selective intermerriage and ethnic attrition can produce significant downward bias in tracking the socio-economic success of U.S. immigrant groups. Spanish

---

[4]I obtained data from the California Department of Public Health for the years 1970-1981 and 1989-2009.

[5]However, when considering individuals whose mothers were born abroad and reported Hispanic origin, the results are substantially unchanged. In addition, only 0.06% of second-generation Hispanic women in the sample reported black race. Excluding them from the analysis does not change the results.

[6]As the information on father's race is missing for 40% of the sample, I do not use it in the main analysis. For those observations for which information on father's race is available, only 2.37% of second-generation Hispanic women had a black partner. Similarly using national data from the Natality Detail data (see Section 4.3), I find that bout 96% of the non-Hispanic fathers of children born to Hispanic mothers report a white race, 3% black race and 1% accounts for other reported (American Indian, Alaskan, Asian, Pacific Islander). Among those for whom information on father's country of birth is available, 98% of non-Hispanics were born in the U.S. Controlling for father's race or excluding black fathers from the analysis does not affect the main results discussed in the paper.

surnames, though imperfect, provide a more objective measure of ethnic origin and allow me to identify individuals of Hispanic origin who did not self-report as Hispanic. Therefore, to identify the father's ethnicity, I use the father's surname and the latest list of US Census surnames,[7] which tabulates surnames that are classified by self-reported race/ethnicity and are based on 270 million individuals with valid surnames on the 2000 US Census (Elliott et al., 2008; Word et al., 2008).[8]

I restrict the sample of third-generation children to those with non-missing information on their father's surname, and whose records were successfully linked to those of their mothers born in California and Florida between 1971 and 1985.[9] After being subjected to these restrictions, the sample consists of 346,962 third-generation children born between January 1989 and December 2009 in California and Florida.

2.1 Matching and Sample Selection

As in the previous literature (Fryer and Levitt, 2004; Currie and Moretti, 2007; Royer, 2009) that used administrative birth records, I am able to link information available at a woman's birth to that of her children if the woman was born in California (Florida) and also gave birth in California (Florida).[10] To construct the intergenerational sample, I linked the records of all the infants born between 1989 and 2009 whose mothers were born in California or Florida between 1970 and 1985 to the birth records of these mothers. I matched the child's birth record to the mother's record using the mother's first and maiden name, exact date of

---

[7]Data are available at http://www.census.gov/topics/population/genealogy.html. Accessed on June 20, 2014.

[8]Using the 2000 Census Bureaus surname list, I consider as Hispanic any surname with a posterior probability of being associated with Hispanic ethnicity higher than 0.7, but the results are robust to the use of different thresholds. Furthermore, using self-reported ethnicity to define intermarriage yields similar findings.

[9]Information on father's surname is missing for approximately 13% of third-generation births.

[10]Florida data contain information on the father's full name and date of birth, allowing me to conduct a parallel analysis using the father's information. However, because of the lesser quality of information on fathers and because men are less likely to become parents at an early age, the matching rate for men is considerably lower than that for women and thus the selectivity of the sample increases. The results from using the fathers information are similar in direction to the results from using the mothers information but are only marginally significant and are available upon request.

birth, and state of birth. Using information on mother's first and maiden names names and date and state of birth, I linked 87.5% (96.6%) of the children born in California (Florida) of all the children born between 1989 and 2009 and whose mothers were born in California (Florida) between 1970 and 1981 (1971 and 1985) to the information available on the birth records of their mothers. I am unable to match observations for names that were misspelled or changed across birth certificates or for dates of birth that were misreported or could not be uniquely identified with the information available. Whenever I was able to uniquely identify the mother's birth record, I included it in the linked sample. The linked sample is not representative of women born between 1970 and 1985. The final sample includes 193,261 (56%) of the 348,164 female children born between 1970 and 1985 in California and Florida. This fact reflects the reality that not all the women born in California and Florida between 1970 and 1985 were still living in those states between 1989 and 2009 and that not all of these women became mothers before 2009. The quality of the matching is discussed in detail in Giuntella (2013).[11]

Table A.1 in the Appendix shows evidence of selection on sociodemographic characteristics. Children of first-generation mothers who were residing in poor zip codes (in the 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 wealthier zip codes (in the highest income quartile). The differences in birth weight (in grams) appear to be negligible and nonsystematic. A 100-gram increase in birth weight increases the probability of a later observation only by 0.6%. However, if the mother was born with low birth weight, she is 13% less likely to be linked. The lower incidence of low birth weight (LBW) in the linked sample can be

---

[11]When including both children of U.S. born and immigrant mothers, the final sample includes 1,355,896 (46%) of the 2,952,909 female children born between 1970 and 1985 in California and Florida. Using the Natality Detail Data (see Section 4.3), which contain information on the mother's state of birth and childs state of birth, we know that approximately 13% of the women born in California and in Florida between 1970 and 1985 had a child in another U.S. state before 2004 (the last 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 this database). By using the American Community Survey (2010), we know that approximately 37% of women born in California and Florida between 1970 and 1985 had not yet had a child by 2009. Data problems such as misspelling or missing information account for the rest of the attrition.

explained by higher rates of infant mortality, higher probabilities of returning to the family's country of origin ("salmon bias"), or by lower probabilities of having a child among those children born with poor health outcomes.

To address the concern of selection bias arising from a matching process that selects a sample of women who were both born and have given birth in either California and Florida, I verify the external validity of the results using data from the Natality Detail Data (see Section 4.3), which contain detailed data on all births in the United States. Using these data allows me to conduct cross-sectional analyses for the entire United States for the 1970–2004 period and address the concern that the results obtained using the California and Florida data may suffer from selection bias due to attrition in the matching process.[12] Furthermore, using a Heckman selection model with the child's year of birth as the excluded variable yields essentially identical results (see Table A.5 in the Appendix).[13]

2.2 Descriptive Statistics

Table 1 reports the descriptive statistics for Hispanic women in exogamous and endogamous relationships. The intermarriage rate in the sample is 18.5%. One of the typical drawbacks of administrative vital statistics is the lack of information on individual income and occupation. However, the data contain information on parental education, on the zip-code of residence, and on the zip-code of the hospital in which the birth occurred. Data on zip code sociodemographic and economic characteristics are drawn from the U.S. Census (source: Social Explorer). In particular, I examine the median family income and the poverty rate as of the 1980 Census for the zip code of the mother's birth and grandmother's residence; and the median family income and the poverty rate as of the 1990 Census for the zip code of the child's birth and mother's residence. Consistent with previous studies, I find that intermarried Hispanic mothers are positively selected with respect to socio-economic

---

[12]The Natality Data from the National Vital Statistics System of the National Center for Health Statistics provide demographic and health data for births occuring during the calendar year.

[13]The year of birth of second-generation women is a significant predictor of later observations, while differences in birth outcomes by year of birth are negligible.

status. They are better educated, less likely to live or give birth in poor zip-codes, less likely to live in ethnic enclaves, and more likely to receive adequate prenatal care.[14] Hispanics in exogamous relationships are also slightly older, and less likely to be unmarried.[15] Similarly, non-Hispanic fathers have higher education and are slightly older than Hispanic fathers in the sample.

Despite the higher socio-economic status, children of second-generation intermarried Hispanics have a substantially higher incidence of low birth weight (birth weight <2500 grams) and infant mortality[16] than children of endogamously married Hispanics. They are also more likely to be macrosomic (birth weight >4500 grams), to present abnormal conditions at birth, to be born before term (<37 weeks) and to have a low APGAR score.[17] Complications during pregnancy are more likely to occur among intermarried Hispanic mothers.

Information on birth weight is consistently available throughout the period studied and in both states. As there is a general consensus that low birth weight is an important marker of health at birth and strongly associated with increased mortality and morbidity risk (Currie and Moretti, 2007; Conley and Bennett, 2000), I focus on the incidence of low birth weight as the main indicator of fitness at birth. However, in the robustness checks, I consider alternative measures of infant health (see Table A.2 and Table A.3).

Administrative records provide only limited information on health behavior during pregnancy and only for more recent years. Information on adult behaviors and health conditions is not available for the California data, whereas the Florida data report tobacco use, alcohol consumption, and weight gain during pregnancy from 1989 onwards, and pre-pregnancy weight and height, chronic hypertension, gestational hypertension, and diabetes from 2004

---

[14]Adequate prenatal care is defined as beginning prenatal care in the first trimester of pregnancy. First-trimester care access is a conventional way to assess the access to care for pregnant women (http://www.hrsa.gov/quality/toolbox/introduction/coreclinical/index.html).

[15]Note that the information on marital status is available only for a subsample of the observations and relies on apparent marital status.

[16]For this measure I consider children who were reported dead within a year. However, most of the children who died in their first year of birth died in their first week.

[17]The Apgar scale is determined by evaluating the newborn baby on five criteria (appearance, pulse, grimace, activity, respiration) on a 10 point scale.

onwards. Intermarriage is associated with a higher incidence of risk-factors that are known to importantly affect birth outcomes.

3 Intermarriage and Birth Outcomes

3.1 Empirical Specification

Overall, Table 1 suggests that intermarriage of a second-generation Hispanic woman with a non-Hispanic man is associated with a significant positive socio-economic premium (Meng and Gregory, 2005) but also with a negative "health penalty". Table 2 investigates whether the associations presented in Table 1 hold up once I account for differences in human capital endowments, socio-demographic characteristics and health at birth between second-generation intermarried Hispanic women and second-generation Hispanic women in endogamous relationships. Furthermore, by focusing on a subsample on second-generation sisters, I exploit only variations in intermarriage within families and verify the robustness of the associations presented in Table 1 to the inclusion of grandmother fixed effects. To this scope, I first estimate the following linear probability model with grandmother fixed effects:[18]

$$LBW_{3,t} = \beta_1 INTM_{2,t} + \beta_2 LBW_{2,t} + \gamma X_{2,t} + \lambda_{GM} + \epsilon_{3,t} \tag{1}$$

where $LBW_{3,t}$ is an indicator for whether the third-generation child was born with a birth weight below 2,500 grams. $INTM_{2,t}$ is an indicator for whether the father's ethnicity was non-Hispanic. $X_{2,t}$ includes a set of standard socio-demographic controls: dummies for the mother's and father's education (less than high school, high school, some college, college or more), indicators for prenatal care adequacy, parity,[19] marital status, maternal age dummies, a quadratic in father's age, dummies for the interaction of county and year of birth for the second and third generations, indicators for whether mother resided in a zip-code in the

---

[18]As a robustness check, I also estimate a probit model, see Table A.5 in the Appendix.
[19]The number of times a woman has given birth.

highest quartile of poverty rate or gave birth in a hospital located in a zip-code in the highest quartile of poverty rate, the poverty rate in the zip-code of birth of the second-generation, and the share of second-generation Hispanic mothers in the zip-code. $LBW_{2,t}$ is an indicator for mother born with low birth weight. λ_{GM} are grandmother fixed effects.

3.2 Main Results

Table 2 confirms that intermarriage is correlated with a significantly higher risk of low birth weight. The incidence of low birth weight for children of intermarried couples is approximately 0.8 percentage points higher (column 1) than that of children of intramarried couples. Including socio-demographic controls, local area characteristics, and grandmother fixed effects reduces the coefficient by approximately 40%, yet the effect remains economically important and statistically significant (+9% with respect to the mean of the dependent variable).[20] It is noteworthy that this result is not sensitive to the addition of the father's educational dummies. The results point in the same direction when alternative birth outcomes are considered, such as indicators for premature birth, birth weight below 1,500 grams, abnormal conditions at birth, and infant mortality (see Table A.2).

In light of the previous literature (Meng and Gregory, 2005; Furtado and Theodoropoulos, 2010; Furtado and Trejo, 2012; Wang, 2012) that provides evidence of a positive relationship between intermarriage and socio-economic outcomes, the fact that the children of intermarried couples have worse birth outcomes is particularly striking. Clearly, intermarriage is not an exogenous decision, but the unobservables that are usually associated with the likelihood of marrying a non-Hispanic would be likely to downward bias the estimated coefficient. Wang (2012) reports that Hispanic–White couples on average earn approximately $20,000 more than Hispanic–Hispanic couples. Indeed, in the data, intermarriage is positively cor-

---

[20]Breaking down the estimates by country of origin, the coefficient on intermarriage is positive and significant for Mexicans (coef., 0.004; s.e., 0.002), while it is positive but not statistically significant for Cubans (coef., 0.006; s.e., 0.007) and Puerto Ricans (coef., 0.003; s.e., 0.012). Using information on father's country of origin to define intermarriage and excluding U.S. born Hispanic men results are similar (coef., 0.009; s.e., 0.009), but the results are not precisely estimated as we only have information on father's country of origin for a subsample of the data.

related with median family income (0.15) in the zip code and with the mother's education (0.10) and father's education (0.14), whereas it is negatively correlated with the zip codes poverty rate (−0.17). Table 3 shows how the positive correlation between intermarriage and socio-economic status holds even after including standard demographic controls (Panel A) and grandmother fixed effects (Panel B). Intermarriage is associated with higher education (defined as having attended some college, having college degree or more), higher average income in the zip-code of residence, lower poverty in the zip code of residence and in the hospital zip code, and a higher likelihood of moving to a neighborhood better than the zip code of birth with respect to the quartiles of poverty rate. Including grandmother fixed effects (Panel B), the intermarriage coefficient on all the socio-economic outcomes becomes smaller in absolute value, but it still economically and statistically significant. Moreover, using the data from Florida that contain limited information on risky behaviors, Table A.4 shows that tobacco consumption during pregnancy is negatively correlated with high education (having attended some college or more) among Hispanics and non-Hispanics and results go in the same direction if considering zip code level economic indicators (e.g, median income or poverty rate). Thus, if anything one may expect that if fixed effects are biased, they are biased towards finding that intermarriage reduces the incidence of low birth weight.[21]

3.3   Instrumental Variable

The extensive set of controls is likely to capture most of the sources of bias on the intermarriage coefficient; however, even after including grandmother fixed effects, one may be concerned about systematic unobserved differences across siblings that may be correlated with both health outcomes and the likelihood of marrying a non-Hispanic man. For these reasons, I instrument for intermarriage using the probability of marrying within the group (Angrist, 2002; Meng and Gregory, 2005; Furtado and Theodoropoulos, 2010). The instrument is the share of immigrant men in a woman's age group residing in her metropolitan

---

[21]Similar results are obtained when analyzing alcohol consumption, while the differences are not significant for gestational hypertension. Results are available upon request.

statistical area (MSA) who come from her own country of origin; this instrument can be formally defined as follows:

$$P_{it}^f(MI) = \frac{n_t^m}{N_t^m} \tag{2}$$

where n_t^m is the number of men from the same country of origin as the woman in a given MSA-age group and N_t^m is the total number of men in the MSA-age group. As the number of men sharing the same country of origin in a given MSA-age group declines, women become more likely to marry outside of their ethnic group.[22] [23]

Age is defined in 9 year intervals so that for a 25-year-old woman, the marriage market will include all men between the ages of 23 and 31 and all the women between the ages of 21 and 29. The key identification assumption is that the instrument affects third-generation births only through its effect on intermarriage. Immigrants do not locate randomly and the existence of ethnic enclaves may be importantly correlated with the maintenance of healthy behaviors. However, the extensive set of controls and, in particular, the inclusion of MSA fixed effects, time-varying zip-code economic characteristics and the yearly share of Hispanic mothers in the zip-code of residence should mitigate the concern that the exclusion restriction may be violated.

Column 1 of Table 4 presents the first stage statistics. Standard errors are clustered at the MSA level. A 10 percentage point (about one standard deviation, see Table 1) increase in the percent of first-generation immigrant men from the same country of origin as the woman in the same MSA-age group decreases the probability of the woman marrying a non-Hispanic by approximately 10 percentage points. The F statistic is well above the conventional thresholds for weak instruments.

---

[22] Using only first-generation immigrants or including all individuals reporting a given ethnic origin, which includes second and later generations, yields similar results.

[23] Alternatively, as in Meng and Gregory (2005) and Furtado and Theodoropoulos (2010), I include as a second instrument the sex ratio within each age-msa-country of origin group; this instrument is defined as follows:

$$SexRatio: SRATIO_{it}^f = \frac{n_t^m}{n_t^f} \tag{3}$$

The results from using this instrument do not substantially differ from the 2SLS estimates reported in Table 4 and are available upon request.

Second-stage results (column 4) confirm the positive coefficient of intermarriage on low birth weight incidence. The point estimate is larger than the ones obtained using OLS (column 2) and grandmother fixed effects (column 3), suggesting that Hispanics who marry within the group are, if anything, negatively selected on unobservables. These results are consistent with the existence of a positive socio-economic gradient in health and the fact that intermarriage has positive effects on socio-economic assimilation (Furtado and Trejo, 2012; Meng and Gregory, 2005).

The result is robust to the inclusion of MSA (column 5) or county fixed effects (column 6). Thus despite there may still be differential selection of women across cities over time, the identification strategy is not subjected to omitted variable bias induced by persistent MSA characteristics that may be correlated with both women's location choices and their health. Exploiting only cohort variation in the share of immigrant men in a woman's age group residing in her metropolitan statistical area (MSA), the coefficient increases in absolute value. Again, these results are consistent with the conjecture that Hispanic women are positively selected into exogamous relationships.

3.4 Fathers' selectivity?

3.4.1 A placebo test

The results presented in Table 2 may still reflect the selectivity of migrants on the father's side. By controlling for grandmother-fixed effects and the mother's birth weight, I am able to partially account for maternal selectivity, but the intermarriage coefficient might nonetheless be the result of the genetic advantage carried by the Hispanic father. Ideally, one way to test whether the results obtained simply reflect negative acculturation (e.g., adoption of unhealthy behaviors) or the effect of Hispanic father's genes would be to compare the birthweight of siblings with the same mothers but different fathers: Hispanic versus non-Hispanic. Conducting this test confirms the positive relationship between intermarriage and the incidence of low birth weight (coef., 0.003, st.dev., 0.006). However, the sample of

children with same mothers but different fathers is simply to small to yield precise estimates. Therefore, to test for the role of the father's selectivity, I conduct a placebo test examining the effect of intermarriage on U.S. born non-Hispanic white women (born in the same years as second-generation Hispanic women analyzed so far) and focusing on the birth outcomes of their offspring. If the intermarriage coefficient captures the father's selectivity, one should expect to find a protective role of having a Hispanic father even when analyzing the effects of intermarriage on the birth weight of their children. When conducting this test, I find that marrying a non-Hispanic (Hispanic) man does not have significant effects on the risk of low birth weight and if anything, is associated with a lower (higher) incidence of low birth weight (see column 3 and 4 in Table 5).

The placebo test has some limitations. There may be other unobserved factors affecting both selection in the marriage market and birth outcomes. However, the extensive set of controls and the overall robustness of the coefficient to the inclusion of grandmother fixed effects reduce the concern that these confounding factors could significantly alter the primary finding and the validity of the falsification test.[24]

3.4.2 Selection in the marriage market

A relevant concern is that selection in the marriage market might be substantially different for men and women. However, while there are significant gender differences in intermarriage rates between blacks and whites, Wang (2012) shows that there are no significant gender differences in the intermarriage rates of Hispanics and whites; white men who married Hispanic women are not less educated than those who married white women. In particular, 32.3% of white men married to white women completed college education, compared to 33.1% of white men who married Hispanic women. One may still be worried about the

---

[24]The placebo test is robust not only to the addition of the father's and mother's education but also to the separate analysis of women who married equally, more, or less educated Hispanic men (results available upon request). Note also that when controlling for grandmother fixed effects and the previous generation's birth weight, one finds that marrying a high school dropout has no significant effect on the risk of low birth weight among children born to non-Hispanic U.S. born mothers.

selection of non-Hispanic men on unobservables (e.g., behaviors) which may not necessarily be correlated with socio-economic status. Given the limited information on paternal behaviors we cannot directly test whether the intermarriage "health penalty" is driven by the negative selection of non-Hispanic men marrried to Hispanic women along other dimensions. Yet, it is worth remarking that while there are marked differences in the incidence of risky behaviors (e.g., smoking) between Hispanic and non-Hispanic white women, the differences are less marked and non-significant among men (Agaku et al., 2014). Furthermore, using the Tobacco Use Supplement of the Current Population Survey (TUS-CPS), after controlling for standard socio-demographic characteristics (e.g., age, education, state of residence) I find no significant differences in the prevalence of tobacco use between non-Hispanic men married to Hispanic women and non-Hispanic men in endogamous relationships.[25][26]

4 Robustness checks

4.1 Alternative Models

Given the binary nature of the main outcome variable, the incidence of low birth weight, I replicate the main results using a correlated random effect probit that includes the grandmothers average education level, adequacy of prenatal care, maternal age, and zip-code level characteristics. Table A.5 reports the average partial effect (APE, see Contoyannis et al. (2004); Wooldridge (2005)). The results obtained with the probit model (column 3) are very similar to those estimated using the linear fixed-effects model (column 1) or a linear model including grandmother level averages for the main socio-demographic characteristics. Column 4 reports the coefficient obtained using Heckman's two-step sample selection correction with the child's year of birth as the excluded variable. Again, the average partial effect is

---

[25]Results are available upon request.

[26]The TUS-CPS is an NCI-sponsored survey of tobacco use and policy information that has been administered as part of the Current Population Survey (CPS) since 1992. The CPS is a monthly survey conducted by the U.S. Census Bureau for the U.S. Department of Labor, Bureau of Labor Statistics. Results are available upon request.

not substantially different from, and if anything, is larger than those estimated using the linear fixed effects model.

4.2 Alternative Metrics of Cultural Assimilation

In Table 6, I consider as alternative metrics of assimilation the ethnic content of the third-generation child's first name, and the share of Hispanic mothers in the zip-code.

Following previous studies on first names, I use two metrics to define the ethnic content of the name: the first indicates whether the third-generation first name appears in a dictionary that includes the most common Hispanic first names (Lansky, 2009) and the second considers the likelihood that a given name will be used by U.S. born non-Hispanics (Fryer and Levitt, 2004).[27]

After including socio-demographic controls and grandmother fixed effects, I find that third-generation children with Hispanic names are 5% less likely to have low birth weight than those whose names do not appear in the Spanish names dictionary. Similarly, the "more American" the name, the higher the incidence of low birth weight.

I then analyze whether the incidence of low birth weight among third-generation children is associated with immigrant co-residence in the neighborhood. Consistent with the findings of Osypuk et al. (2010); Finch et al. (2007), living in an ethnic enclave is a protective factor for birth outcomes. This effect holds true even after including grandmother fixed effects and becomes larger once one controls for zip-code poverty and median income. A one standard deviation increase in the share of Hispanic mothers in the zip code is associated with a 9% lower incidence of low birth weight.

---

[27]Formally, the index is defined as follows:

$$\text{American Name Index} = \frac{\text{Pr(Name| US born mother)}}{\text{Pr(Name|US born mother)} + \text{Pr(Name| Foreign born mother)}}$$

where $Pr(Name|$US born mother) measures the frequency of a given name among children born in the period 1989-2009 to a US-born mother, and $Pr(Name|$ Foreign born mother) measures the frequency of the name among children born to foreign mothers who come from the same country of origin.

4.3 External Validity: Evidence from Natality Detail Data (1989-2004)

To verify the external validity of the results obtained using the data from California and Florida, I conduct a similar analysis using the Natality Detail Data, which contain detailed data on all births in the United States. The Natality data are drawn from the National Vital Statistics System of the National Center for Health Statistics and provide demographic and health data for births occuring during the calendar year. The microdata are based on information abstracted from birth certificates filed in vital statistics offices of each State and District of Columbia. Demographic data include variables such as date of birth, age and educational attainment of parents, marital status, live-birth order, race, sex, and geographic area. Health data include items such as birth weight, gestation, prenatal care, attendant at birth, and Apgar score. The public version of the Natality Detail Data does not allow for cross-generational record linking because the dataset does not release information regarding the names of the child and mother. Geographic data include state, county, city, standard metropolitan statistical area (SMSA, 1980 onwards), and metropolitan and non-metropolitan counties. From 2005 onwards, the data do not include any geographic variables, such as state, county, or SMSA. These data do not contain confidential information for identifying siblings, and therefore, I rely on self-reported ethnicity to define intermarriage. Because of these limitations, I cannot use grandmother fixed effects and I cannot distinguish second-generation Hispanics from later generation immigrants for this analysis. To facilitate comparability with the main results presented in Section 3, I restrict the sample to mothers born between 1970 and 1985.

Panel A of Table 7 shows that the OLS estimates and the 2SLS estimates are very close to the estimates presented in columns 1 and 4 of Table 4 for California and Florida. Again, intermarriage is associated with a higher incidence of low birth weight. This correlation holds up when an extensive set of socio-demographic controls are included. 2SLS estimates are larger than the OLS estimates. Including county fixed effects and exploiting only variation

20

across cohorts in the availability of men from one's country of origin increases the coefficient. Again, these results suggest positive selection on intermarriage, which is consistent with the existence of a positive socio-economic gradient in health and the observed positive association between intermarriage and socio-economic status.

Furthermore, Table 8 shows that there is no significant effect of marrying a non-Hispanic on children's low birth weight, when we analyze the sample of US born non-Hispanic white native women (columns 1 and 2), and no evidence of negative intermarriage effects among Hispanic women (e.g., women of Mexican descendance) married to Hispanic men reporting a different origin (e.g., men of Puerto Rican origin, see columns 3 and 4). These results suggest that the intermarriage "health penalty" is not explained by unobserved characteristics of Hispanic men, nor by unobservable individual characteristics associated with exogamous marriages.

5   Possible Mechanisms

Intermarriage may directly affect the adoption of risky behaviors by changing the reference group and facilitating the adoption of the social norms prevailing among U.S. born non-Hispanic women. To further investigate the possible mechanisms underlying the negative effect of intermarriage on third-generation birth outcomes, I take advantage of the Natality Detail Data, which allows me to exploit information on maternal behaviors and risk factors that is available in the birth records of all U.S. states.

Table 9 illustrates the incidence of these risk factors among non-Hispanic white natives, non-Hispanic blacks, and immigrants of Hispanic origin. First-generation immigrants have a substantially lower incidence of risk factors compared to non-Hispanic white natives. Hispanic descendants show some convergence towards the less healthy behaviors and higher incidence of risk factors of non-Hispanic white natives, but they retain a fairly sizeable health advantage. Yet, for both first and second-generation, intermarriage is associated with

21

a higher incidence of risk factors and less distance to the native behavioral norm.[28]

Table 10 focuses on smoking during pregnancy, which has been widely recognized as the most modifiable risk factor for low birth weight (Almond et al., 2005; Currie and Schmieder, 2009). The results presented in Table 10 confirm that intermarriage is importantly associated with smoking during pregnancy and that this correlation holds up when socio-demographic controls are included (columns 1-3). Using the data from Florida that contain limited information on risky behaviors, I investigate whether the correlation between intermarriage and tobacco consumption during pregnancy is robust to the inclusion of grandmother fixed effects (columns 4-7). Columns 6 and 7 show that when including grandmother fixed effects, the coefficient decreases from 2 percentage points to 1.3 percentage points, but the effect remains economically important and statistically significant. Being intermarried is associated with a 60% higher likelihood of tobacco use during pregnancy. The results for other risk factors, such as alcohol use and gestational hypertension, tend toward the same direction. However, in these cases, the magnitude and the significance of the coefficient are less robust to the inclusion of grandmother fixed effects and socio-demographic controls.[29]

Table 11 shows that when we control for this set of risky behaviors, the effect of intermarriage on low-birth weight reduces substantially. Panel B focuses on smoking, while Panel C includes smoking, alcohol consumption and gestational hypertension.[30] When controlling for observable risk factors, the OLS estimate becomes non-significant, while the 2SLS coefficient shrinks by approximately 32%. This evidence suggests that the higher incidence of observed risk factors (such as higher rates of smoking, alcohol consumption, and hypertension) among intermarried Hispanic women explains a substantial part the negative relationship between second-generation intermarriage and third-generation birth outcomes.

---

[28]Note while that the prevalence of smoking among Hispanics is lower than the prevalence among non-Hispanic whites, smoking prevalence among Hispanic men is substantially higher than among Hispanic women and not statistically different from the prevalence observed among non-Hispanic whites Agaku et al. (2014). Hispanic women are also much more likely than non-Hispanic white women to quit smoking during pregnancy (Osterman et al., 2013).

[29]These results are available upon request.

[30]Including information on maternal weight gain and maternal diabetes leaves the estimates practically unchanged. The results are available upon request.

It is worth noting that I am able to account for the contribution of only a limited set of behaviors and that there are other important behaviors that I do not observe which may be correlated with intermarriage and birth outcomes.[31] The unexplained part of the "intermarriage penalty" is likely to be related to other types of behavior, such as dietary habits, for which I do not have data but that are known to significantly affect birth outcomes and to deteriorate with time spent in the US and across generations.

6 Conclusion

This paper analyzed the effects of second-generation assimilation on the birth outcomes of third-generation Hispanics in the US. Among second-generation Hispanics, intramarried couples exhibit higher resilience in terms of healthy behaviors, health conditions, and birth outcomes. Using ethnic intermarriage as a metric of cultural assimilation, I show that third-generation children of intermarried Hispanic couples are 9% more likely to be of low birth weight than children of intramarried couples. This holds true even after accounting for potential confounding factors, focusing on a subsample of second-generation siblings and controlling for grandmother fixed effects. 2SLS estimates confirm these main findings and, consistent with the existence of a positive socio-economic gradient in health, are larger than the OLS estimates. Furthermore, the results do not appear to be driven by unobservable traits associated with the father's ethnicity and intermarriage. The father's Hispanic ethnicity has no protective effect when analyzing children of non-Hispanic women, nor does it have a significant effect on Hispanic women intermarried with Hispanic men of different origin. Instead, the evidence suggests that these differences are explained by the higher prevalence of risky behaviors (e.g., smoking) among second-generation mothers in exogamous relationships.

---

[31]For instance, maternal dietary practices have been shown to be significant determinants of health at birth, and there is evidence of unhealthy assimilation with respect to dietary habits (Guendelman and Abrams, 1995). Using the American Time Use Survey, I find evidence that among immigrants of Hispanic origin, intramarriage is positively associated with time spent on food preparation and consumption.

Overall, these results highlight the importance of policies that encourage the maintenance of healthy behaviors among immigrant descendants. Educational and prevention programs targeting Hispanic mothers may have important effects, particularly on those who are highly acculturated, and thus have a significant potential to reduce these risks. These policies may partially counteract the tobacco industrys increasing efforts to market cigarettes to Hispanic immigrants in the U.S. (Osypuk and Acevedo-Garcia, 2010; Acevedo-Garcia et al., 2004).

Finally, these results are also relevant to sending countries, such as Mexico, that are undergoing important epidemiological transitions. As health transitions occur more rapidly in a migrating population, the health trajectories observed among immigrants can help predict the risks associated with the typical public health problems related to urbanization, the diffusion of sedentary habits, and substance abuse in developing countries.

References

Acevedo-Garcia, Dolores, Elizabeth Barbeau, Jennifer Anne Bishop, Jocelyn Pan, and Karen M Emmons (2004) 'Undoing an epidemiological paradox: the tobacco industry's targeting of US immigrants.' *American Journal of Public Health* 94(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

Agaku, Israel T, Brian A King, Shanta R Dube, Centers for Disease Control, Prevention (CDC) et al. (2014) 'Current cigarette smoking among adultsunited states, 2005–2012.' *MMWR Morb Mortal Wkly Rep* 63(2), 29–34

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

Angrist, Josh (2002) 'How do sex ratios affect marriage and labor markets? evidence from america's second generation.' *Quarterly Journal of Economics* pp. 997–1038

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

Baker, Michael, and Dwayne Benjamin (1997) 'The role of the family in immigrants' labor-market activity: an evaluation of alternative explanations.' *The American Economic Review* pp. 705–727

Bethel, Jeffrey W, and Marc B Schenker (2005) 'Acculturation and smoking patterns among Hispanics: a review.' *American journal of preventive medicine* 29(2), 143–148

Bisin, Alberto, and Thierry Verdier (2000) 'Beyond the melting pot: cultural transmission, marriage, and the evolution of ethnic and religious traits.' *The Quarterly Journal of Economics* 115(3), 955–988

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou (2011) 'Ethnic identity and labour market outcomes of immigrants in europe.' *Economic Policy* 26(65), 57–92

Bisin, Alberto, Giorgio Topa, and Thierry Verdier (2004) 'Religious intermarriage and socialization in the united states.' *Journal of Political Economy* 112(3), 615–664

Case, Anne, Angela Fertig, and Christina Paxson (2005) 'The lasting impact of childhood health and circumstance.' *Journal of Health Economics* 24, 365–389

Case, Anne, Darren Lubotsky, and Christina Paxson (2002) 'Economic status and health in childhood: The origins of the gradient.' *The American Economic Review* 92(5), 1308–1334

Chiswick, Barry R, and Christina Houseworth (2011) 'Ethnic intermarriage among immigrants: Human capital and assortative mating.' *Review of Economics of the Household* 9(2), 149–180

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

Contoyannis, Paul, Andrew M Jones, and Nigel Rice (2004) 'The dynamics of health in the british household panel survey.' *Journal of Applied Econometrics* 19(4), 473–503

Currie, Janet (2011) 'Inequality at birth: Some causes and consequenses.' *NBER working paper*

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

Currie, Janet, and Johannes Schmieder (2009) 'Fetal exposures to toxic releases and infant health.' *American Economic Review, American Economic Association* 99(2), 117–83

Delavari, Maryam, Anders Larrabee Sønderlund, Boyd Swinburn, David Mellor, and Andre Renzaho (2013) 'Acculturation and obesity among migrant populations in high income countries–a systematic review.' *BMC Public Health* 13(458), 2–11

Duncan, Brian, and Stephen J Trejo (2007) 'Ethnic identification, intermarriage, and unmeasured progress by Mexican Americans.' In *Mexican immigration to the United States,* ed. George J. Borjas University of Chicago Press pp. 229–268

_ (2009) 'Ancestry versus ethnicity: The complexity and selectivity of Mexican identification in the united states.' *Research in Labor Economics* 29, 31–66

_ (2011) 'Tracking intergenerational progress for immigrant groups: The problem of ethnic attrition.' *The American Economic Review* 101(3), 603–608

_ (2012) 'The complexity of immigrant generations: Implications for assessing the socioeconomic integration of Hispanics and Asians.' *IZA Discussion Paper No. 6276*

Eckstein, Zvi, and Yoram Weiss (2002) 'The integration of immigrants from the former Soviet Union in the Israeli labor market.' In *A special volume in memory of Michael Bruno,* ed. A. Ben-Basat (MA: MIT Press) pp. 349–77

Elder, Todd E., John H. Goddeeris, and Steven J. Haider (2013) 'Racial and ethnic infant mortality gaps and the role of ses.' `https://www.msu.edu/~telder/SES_Current.pdf`. Accessed Feb. 2015

Elliott, Marc N, Allen Fremont, Peter A Morrison, Philip Pantoja, and Nicole Lurie (2008) 'A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity.' *Health services research* 43(5p1), 1722–1736

Fenelon, Andrew (2013) 'Revisiting the Hispanic mortality advantage in the United States: The role of smoking.' *Social Science & Medicine* 82, 1–9

Fernandez, Raquel (2011) 'Does culture matter?' In *Handbook of Social Economics,* ed. Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, vol. 1A (The Nederlands: North-Holland) pp. 481–510

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

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

Furtado, Delia, and Nikolaos Theodoropoulos (2010) 'Why does intermarriage increase immigrant employment? the role of networks.' *The BE Journal of Economic Analysis and Policy*

Furtado, Delia, and Stephen Trejo (2012) 'Interethnic marriages and their economic effects.' CReAM Discussion Paper Series 1205, Centre for Research and Analysis of Migration

(CReAM), Department of Economics, University College London, Feb. `http://ideas.repec.org/p/crm/wpaper/1205.html`. Accessed Nov. 2012

Giuntella, Osea (2013) 'Why does the health of immigrants deteriorate?' *IZA Discussion Paper 7588*

Gordon, Milton A. (1964) *Assimilation in America* (New York: Oxford University Press)

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* 85(1), 20–25

Jasso, Guillermina, Douglas S. Massey, Mark R. Rosenzweig, and James P. Smith (2004) 'Immigrant health: Selectivity and acculturation.' *The Institute for Fiscal Studies Working Paper* 4(23), 1–46

Kaushal, Neeraj, and Robert Kaestner (2010) 'Health and health insurance trajectories of Mexicans in the US.' Working Paper 16139, National Bureau of Economic Research, June

Lansky, Bruce (2009) *100,000+ Baby Names: The Most Complete, Fascinating, and Helpful Name Book You Can Find* (Simon and Schuster)

Marmot, Michael G., and S.Leonard Syme (1976) 'Acculturation and coronary heart disease in Japanese-Americans.' *American Journal of Epidemiology* 104(3), 225–247

Meng, Xin, and Robert G. Gregory (2005) 'Intermarriage and the economic assimilation of immigrants.' *Journal of Labor Economics* 23(1), 135–174

Nekby, Lena (2010) 'Inter-and intra-marriage premiums revisited: Its probably who you are, not who you marry!' *IZA Discussion Paper No. 5317*

Osterman, MJ, JA Martin, SC Curtin, TJ Matthews, EC Wilson, and S Kirmeyer (2013) 'Newly released data from the revised us birth certificate, 2011.' *National vital statistics*

reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System 62(4), 1–22

Osypuk, Theresa L, and Dolores Acevedo-Garcia (2010) 'Support for smoke-free policies: A nationwide analysis of immigrants, US-born, and other demographic groups, 1995–2002.' *American journal of public health* 100(1), 171

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

Qian, Zhenchao, and Daniel T Lichter (2007) 'Social boundaries and marital assimilation: Interpreting trends in racial and ethnic intermarriage.' *American Sociological Review* 72(1), 68–94

Riosmena, Fernando, Rebeca Wong, and Alberto Palloni (2013) 'Paradox lost: Explaining the Hispanic adult mortality advantage.' *Demography* 50, 1039–1064

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

Sassler, Sharon (2005) 'Gender and ethnic differences in marital assimilation in the early twentieth century.' *International Migration Review* 39(3), 608–634

Ullmann, S.Heidi, Noreen Goldman, and Douglas Massey (2011) 'Healthier before they migrate, less healthy when they return?' *Social Science and Medicine* 73, 421–428

Unger, Jennifer B, Louise Ann Rohrbach, Tess Boley Cruz, Lourdes Baezconde-Garbanati, Paula H Palmer, C Anderson Johnson, and Kim Ammann Howard (2001) 'Ethnic variation in peer influences on adolescent smoking.' *Nicotine & Tobacco Research* 3(2), 167–176

Wang, Wendy (2012) 'The rise of intermarriage. rates, characteristics vary by race and

gender.' *Pew Social and Demographic Trends.* `http://www.pewsocialtrends.org/2012/`
`02/16/the-rise-of-intermarriage/`. Accessed Nov. 2012

Wooldridge, Jeffrey M (2005) 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity.' *Journal of Applied Econometrics* 20(1), 39–54

Word, David L, Charles D Coleman, Robert Nunziata, and Robert Kominski (2008) 'Demographic aspects of surnames from census 2000.' *Unpublished manuscript, Retrieved from http://citeseerx. ist. psu. edu/viewdoc/download*

Table 1: Summary Statistics

	Endogamous Mean	Endogamous Std.Dev	Exogamous Mean	Exogamous Std.Dev
Child's characteristics (3rd-generation)				
Birth outcomes				
Low birth weight (<2500 grams)	5.28	(22.36)	6.04	(23.83)
Very low birth weight (<1500 grams)	0.54	(7.36)	0.68	(8.23)
Birth weight (in grams)	3354.44	(527.79)	3343.81	(543.86)
Infant mortality	0.05	(2.27)	0.12	(3.44)
Complications during pregnancy	50.09	(50.00)	52.20	(49.95)
Macrosomic (>4000 grams)	9.11	(28.78)	9.16	(28.84)
Abnormal conditions	3.17	(17.51)	3.82	(19.18)
Premature	9.88	(29.83)	10.00	(30.00)
Low APGAR* (below 4)	0.45	(6.73)	0.53	(7.28)
Other child's characteristics				
Parity	0.98	(1.10)	0.83	(1.02)
Female	48.88	(49.99)	48.45	(49.98)
Spanish First-Name	34.15	(47.42)	17.34	(37.86)
American Name Index	27.96	(26.84)	36.87	(30.54)
Mother's characteristics (2nd-generation)				
Low birth weight	4.39	(20.49)	4.82	(21.42)
Birth weight (in grams)	3343.10	(496.76)	3305.81	(496.75)
Less than high-school	28.92	(45.34)	16.59	(37.20)
High school	38.39	(48.63)	32.24	(46.74)
Some college	22.53	(41.78)	28.65	(45.21)
College or more	10.16	(30.21)	22.52	(41.77)
Adequate prenatal care	85.22	(35.49)	87.22	(33.39)
Married	47.41	(49.93)	51.87	(49.97)
Age	23.98	(4.82)	25.46	(5.09)
Father's characteristics				
Less than high-school	36.22	(48.06)	14.40	(35.11)
High school	41.74	(49.31)	39.76	(48.94)
Some college	14.36	(35.07)	22.63	(41.85)
College or more	7.68	(26.62)	23.21	(42.22)
Age	26.41	(5.53)	28.16	(6.19)
Zip-code characteristics at birth (3rd-generation)				
Born in poorest quartile (residence)	32.53	(46.85)	18.66	(38.96)
Born in poorest quartile (hospital)	21.65	(41.18)	13.18	(33.82)
Median Zipcode income 1990 (residence)	11.04	(0.32)	11.15	(0.31)
Share of Hispanic mothers in Zipcode (residence)	28.97	(19.07)	21.13	(17.52)
Poverty share in mother's Zipcode of birth (hospital)	19.75	(39.66)	17.36	(37.73)
Marriage market conditions				
Share of Men in MSA-age-country of origin cell	23.31	(8.65)	17.13	(11.75)
Maternal behaviors and risk-factors (FLORIDA)				
Smoking during pregnancy	1.28	(11.22)	3.50	(18.38)
Alcohol during pregnancy	0.07	(2.70)	0.23	(4.75)
Gestational hypertension	3.58	(18.58)	4.13	(19.90)
Observations	282,794	(81.5%)	64,168	(18.5%)

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). The sample consists of 346,962 observations.*Information on APGAR score is available only for the Florida sample.

Table 2: Intermarriage and Low Birth Weight (LBW)

Methodology	(1) OLS	(2) OLS	(3) GM F.E.	(4) GM F.E.
Non-Hispanic father	0.008***	0.006***	0.005**	0.005**
	(0.001)	(0.001)	(0.002)	(0.002)
Socio-demographic controls	NO	YES	YES	YES
Grand-mother (GM) F.E.	NO	NO	YES	YES
Zip-code characteristics	NO	NO	NO	YES
Mean of the Dep.Var.	0.054	0.054	0.054	0.054
Std. Dev.	0.226	0.226	0.226	0.226
Observations	346,962	346,962	346,962	346,962

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). Socio-demographic controls include 3rd generation child's gender, 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, a quadratic in father's age, second generation parity, an indicator for adequate prenatal care, marital status, father's and mother's education (4 groups), dummies for the interaction of county and year of birth of third-generation children, indicators for missing information on maternal and paternal education, adequacy of prenatal care, mother's birth weight, father's age. Column 3 and 4 include grandmother-fixed effects. Column 4 includes an indicator for whether the mother was residing in a high-poverty zip-code, the hospital was located in a high-poverty zip-code, the share of Hispanic mothers in the zip code of residence, and the poverty rate in the zip-code of birth of second-generation mothers. Standard errors are clustered at the grand-mother level.

Table 3: Intermarriage and Socio-Economic Status

Dependent Variable	(1) High-Education	(2) Log (Zipcode Income)	(3) High-Poverty Zipcode Hospital	(4) High-Poverty Zipcode Residence	(5) Better Neighborhood
Panel A: OLS					
Non-Hispanic father	0.084***	0.051***	-0.027***	-0.071***	0.037***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Socio-demographic controls	YES	YES	YES	YES	YES
Panel B: Grand-Mother Fixed Effects					
Non-Hispanic Father	0.044***	0.028***	-0.017***	-0.039***	0.024***
	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
Socio-demographic controls	YES	YES	YES	YES	YES
GM F.E.	YES	YES	YES	YES	YES
Mean of the Dep.Var.	0.356	11.062	0.202	0.300	0.458
Std. Dev.	0.479	0.323	0.401	0.458	0.498
Observations	346,962	295,754	295,754	330,847	346,962

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). All estimates include 3rd generation child's gender, 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, an indicator for prenatal care, marital status, dummies for the interaction of county and year of birth of third-generation children, indicators for missing information on adequacy of prenatal care, and mother's birth weight. Panel B includes grandmother-fixed effects. High-education is defined at having some college, college degree or more. In columns 3 and 4, the dependent variable is a dummy for whether the zip code was in the highest quartile of poverty rate; in column 5 the dependent variable is a dummy for whether the mother was living in a neighborhood better than the zip code of birth with respect to the quartiles of poverty rate. Standard errors are clustered at the grand-mother level.

Table 4: 2SLS Estimates

Methodology Dependent Variable	(1) First Stage Non-Hispanic father	(2) OLS LBW	(3) GM F.E. LBW	(4) 2SLS LBW	(5) 2SLS LBW	(6) 2SLS LBW
Non-Hispanic father		0.008***	0.005**	0.025**	0.052***	0.050***
		(0.001)	(0.002)	(0.011)	(0.010)	(0.010)
Share of Men in MSA-age- country of origin cell	-1.025*** (0.117)					
First Stage F	126.42					
Socio-demographic controls	YES	NO	YES	YES	YES	YES
Grand-mother (GM) F.E.	NO	NO	YES	NO	NO	NO
Zip-code characteristics	YES	NO	YES	YES	YES	YES
MSA F.E.	NO	NO	NO	NO	YES	NO
County F.E.	NO	NO	YES	NO	NO	YES
Mean of the Dep. Var.	0.185	0.054	0.054	0.054	0.054	0.054
Std. Dev.	0.388	0.226	0.226	0.226	0.226	0.226
Observations	346,962	346,962	346,962	346,962	346,962	346,962

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). Column 1 presents the first-stage estimate. Columns 2 and 3 replicate the estimates presented in Colums 1 and 4 of Table 2. Columns 4-6 reports 2SLS estimates. Socio-demographic controls include 3rd generation child's gender, mother's birth weight (LBW), dummies for the year of birth of second-generation children (mothers), second generation's age (at delivery) dummies, a quadratic in father's age, second generation parity, an indicator for prenatal care, marital status, father's and mother's education (4 groups), dummies for the year of birth of third-generation children, indicators for missing information on maternal education, adequacy of prenatal care, mother's birth weight, state fixed effects and zip code level characteristics. Zip-code level characteristics include an indicator for whether the mother was residing in a high-poverty zip-code, the hospital was located in a high-poverty zip-code, the share of Hispanic mothers in the zip code of residence, and the poverty rate in the zip-code of birth of second-generation mothers. Standard errors are clustered at the MSA level.

Table 5: Placebo Test: Intermarriage and Low Birth Weight, US-born Non-Hispanic White Women

Methodology	(1) Hispanics OLS	(2) Hispanics GM F.E.	(3) Non-Hispanic Whites OLS	(4) Non-Hispanic Whites GM F.E.
Non-Hispanic father	0.008***	0.005**	-0.001	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)
Socio-demographic controls	NO	YES	NO	YES
GM F.E.	NO	YES	NO	YES
Mean of Dep.Var.	0.054	0.054	0.056	0.056
Std.Dev.	0.226	0.226	0.230	0.230
Observations	346,962	346,962	1,327,470	1,327,470

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). Column 1 and 3 replicate the analysis conducted in column 1 of Table 2. Columns 2 and 4 include grandmother-fixed effects and the same set of controls used in column 4 of Table 2. Standard errors are clustered at the grand-mother level.

Table 6: Alternative Metrics of Cultural Assimilation

Methodology	(1) OLS	(2) OLS	(3) GM F.E.
		Panel A	
Spanish first-name (3rd gen)	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
Mean of the Dep.Var.	0.054	0.054	0.054
Std. Dev.	0.226	0.226	0.226
Observations	346,881	346,881	346,881
		Panel B	
American Name Index (3rd gen)	0.008***	0.005***	0.005**
	(0.002)	(0.002)	(0.002)
Mean of the Dep.Var.	0.054	0.054	0.054
Std. Dev.	0.226	0.226	0.226
Observations	346,881	346,881	346,881
		Panel C	
Share of Hispanic Mothers in Zipcode	-0.011***	-0.012***	-0.024***
	(0.002)	(0.002)	(0.004)
Socio-demographic controls	NO	YES	YES
Grand-mother (GM) F.E.	NO	NO	YES
Zip-code characteristics	NO	YES	YES
Mean of the Dep.Var.	0.054	0.054	0.054
Std. Dev.	0.226	0.226	0.226
Observations	319,549	319,549	319,549

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). Socio-demographic controls include 3rd generation child's gender, 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, a quadratic in father's age, second generation parity, an indicator for adequate prenatal care, marital status, father's and mother's education (4 groups), dummies for the interaction of county and year of birth of third-generation children, indicators for missing information on maternal and paternal education, adequacy of prenatal care, mother's birth weight, father's age, an indicator for whether the mother was residing in a high-poverty zipcode, the hospital was located in a high-poverty zip-code, the share of Hispanic mothers in the zip code of residence, and the poverty rate in the zip-code of birth of second-generation mothers. Column 3 includes grand-mother fixed effects. Standard errors are clustered at the grand-mother level.

Table 7: Intermarriage and Low Birth Weight, Natality Detail Data (1989-2004)

Methlology	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Non-Hispanic father	0.008***	0.002***	0.037**	0.060***
	(0.001)	(0.001)	(0.017)	0.019
Socio-demographic controls	NO	YES	YES	YES
State fixed effects	NO	YES	YES	NO
County fixed effects	NO	NO	NO	YES
First Stage F			32.23	48.87
Mean of Dep.Var	0.070	0.070	0.070	0.070
Std.Dev.	0.069	0.069	0.0690.069	
Observations	2,326,886	2,326,886	2,326,886	2,326,886

Notes - Data are drawn from the Natality Detail Data (1989-2004). Socio-demographic controls include indicators for child's gender, indicators for adequate prenatal care, marital status, mother's education, parity, dummies for mother's and child's year of birth. Columns 2 and 3 include state fixed effects. Column 4 includes county fixed effects. Standard errors are clustered at the MSA level.

Table 8: Placebo Test: Natality Detail Data (1989-2004)

	(1) NH White	(2) NH White	(3) Hispanics	(4) Hispanics
Non-Hispanic father	0.001*	0.001		
	(0.001)	(0.000)		
Hispanic father reporting different country of origin			0.003***	0.001
			(0.001)	(0.001)
Socio-demographic controls	NO	YES	NO	YES
State F.E.	NO	YES	NO	YES
Mean of the Dep. Var.	0.066	0.066	0.067	0.067
Std.Dev.	0.249	0.249	0.254	0.254
Observations	15,813,352	15,813,352	1,531,578	1,531,578

Notes - Data are drawn from the Natality Detail Data (1989-2004). Columns 1 and 2 restrict the analysis to non-Hispanic white women. Columns 3 and 4 exclude Hispanics women whose partner reported to be a non-Hispanic white. Data are drawn from the Natality Detail Data (1989-2004). Socio-demographic controls include indicators for child's gender, parity, indicators for adequate prenatal care, marital status, mother's education, dummies for mother's and child's year of birth, state fixed effects. Standard errors are clustered at the MSA level.

Table 9: Risk Factors and Assimilation (%), Natality Detail Data (1989-2004)

	Smoking during pregnancy	Alcohol use during pregnancy	Gestational Hypertension
Non-Hispanic U.S. born whites			
	21.05	1.01	4.32
	(40.77)	(10.02)	(20.33)
Non-Hispanic U.S. born blacks			
	8.86	1.00	3.36
	(28.42)	(10.12)	(18.80)
First Generation Hispanics			
	1.36	0.28	2.10
	(11.57)	(5.30)	(14.35)
First Generation - Exogamously Married			
	2.93	0.59	2.38
	(16.87)	(7.64)	(15.23)
First Generation - Endogamously Married			
	0.94	0.20	2.05
	(9.65)	(4.47)	(14.17)
*Second Generation Hispanics**			
	6.34	0.86	2.94
	(24.37)	(9.23)	(16.90)
Second Generation - Exogamously Married			
	10.23	1.35	3.16
	(30.31)	(11.52)	(17.50)
Second Generation - Endogamously Married			
	3.83	0.54	2.83
	(19.19)	(7.34)	(16.59)

Notes - Data are drawn from the Natality Detail Data (1989-2004). The sample is restricted to mothers born between 1970 and 1985 and includes 15,813,352 non-Hispanic US born mothers, 3,346,280 first-generation Hispanic mothers, and 2,326,886 second-generation Hispanic mothers. All numbers are expressed as percentages of total population. *In the Natality Detail Data, we cannot distinguish second-generation Hispanics from later generation immigrants.

Table 10: Intermarriage and Smoking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Natality Detail Data (1989-2004)			Florida (1989-2009)			
Non-Hispanic father	0.064***	0.047***	0.045***	0.022***	0.014***	0.013**	0.013**
	(0.005)	(0.004)	(0.006)	(0.002)	(0.002)	(0.006)	(0.006)
Socio-demographic controls	NO	YES	YES	NO	YES	NO	YES
Zip-code characteristics	NO	NO	NO	NO	YES	NO	YES
STATE F.E.	NO	YES	NO	NO	NO	NO	NO
MSA F.E.	NO	NO	YES	NO	NO	NO	NO
County F.E.	NO	NO	NO	NO	YES	NO	YES
Grand-mother F.E.	NO	NO	NO	NO	NO	YES	YES
Mean of Dep. Var.	0.063	0.063	0.063	0.0205	0.0205	0.0205	0.0205
Std. Dev.	0.244	0.244	0.244	0.142	0.142	0.142	0.142
Observations	1,499,783	1,499,783	1,499,783	33,834	33,834	33,834	33,834

Notes - The sample is restricted to Hispanic mothers born between 1970 and 1985. For columns 1-3, data are drawn from the Natality Detail Data (1989-2004). Socio-demographic controls include indicators for child's gender, parity, indicators for adequate prenatal care, marital status, mother's education, dummies for mother's and child's year of birth, state fixed effects. For columns 4-7 data are drawn from the Florida Vital Statistics (1989-2009). Socio-demographic controls include 3rd generation child's gender, 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, a quadratic in father's age, second generation parity, an indicator for adequate prenatal care, marital status, father's and mother's education (4 groups), dummies for the interaction of county and year of birth of third-generation children, indicators for missing information on maternal and paternal education, adequacy of prenatal care, mother's birth weight, father's age, an indicator for whether the mother was residing in a high-poverty zipcode, the hospital was located in a high-poverty zip-code, the share of Hispanic mothers in the zip code of residence, and the poverty rate in the zip-code of birth of second-generation mothers. Standard errors are clustered at the MSA level.

Table 11: Intermarriage, Risk Factors and Low Birth Weight, Natality Detail Data (1989-2004)

	(1) OLS	(2) OLS	(3) 2SLS
Methology			
Panel A: Intermarriage and low birth weight			
Non-Hispanic father	0.008***	0.002***	0.037**
	(0.001)	(0.001)	(0.017)
First Stage F			32.23
Panel B: Intermarriage and low birth weight, controlling for smoking			
Non-Hispanic Father	0.004***	0.000	0.030*
	(0.001)	(0.001)	(0.016)
First Stage F			32.20
*Panel C: Intermarriage and low birth weight, controlling for risk-factors**			
Non-Hispanic Father	0.003***	-0.000	0.025
	(0.001)	(0.001)	(0.015)
First Stage F			32.21
Socio-demographic controls	NO	YES	YES
State F.E.	NO	YES	YES
Mean of Dep.Var	0.070	0.070	0.070
Std.Dev.	0.069	0.069	0.069
Observations	2,326,886	2,326,886	2,326,886

Notes - Data are drawn from the Natality Detail Data (1989-2004). Socio-demographic controls include indicators for child's gender, indicators for adequate prenatal care, marital status, mother's education, parity, dummies for mother's and child's year of birth, state fixed effects. Columns 3 reports 2SLS estimates. Standard errors are clustered at the MSA level. *Risk factors include smoking, alcohol use and gestational hypertension.

Appendix A

Table A.1: Matching quality. Women born in California and Florida, 1970–1985

	Observations			Birth Weight (grams)			Low Birth Weight (below 2500 grams)		
Sample:	(1) Overall	(2) Linked	(3) Matching rate	(4) Overall	(5) Linked	(6) Nonlinked	(7) Overall	(8) Linked	(9) Nonlinked
Overall	2,952,909	1,355,896	0.46	3,274	3,275	3,272	0.072	0.067	0.076
US born	2,401,162	1,097,301	0.46	3,269	3,266	3,271	0.075	0.071	0.079
Hispanics	348,164	193,261	0.56	3,315	3,332	3,294	0.053	0.046	0.063
Zip code level income:									
1st income quartile	471,251	236,068	0.50	3,252	3,255	3,248	0.076	0.071	0.082
2nd income quartile	542,832	267,325	0.49	3,251	3,253	3,249	0.079	0.074	0.084
3rd income quartile	796,457	360,497	0.45	3,273	3,276	3,271	0.072	0.067	0.076
4th income quartile	700,271	296,500	0.42	3,299	3,300	3,298	0.064	0.059	0.068

Notes - Data are drawn from the California and Florida Vital Statistics, 1970–1985. The linked sample is composed of all the women born between 1970 and 1985 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 A.2: Other Birth Outcomes

Dependent Variable:	(1) Low Birth Weight	(2) Very Low Birth Weight	(3) Premature Birth	(4) Child Death	(5) Abnormal Conditions	(6) Complicated Pregnancy	(7) Macrosomic	(8) Low Apgar
OLS	0.006*** (0.001)	0.001*** (0.000)	0.006*** (0.002)	0.000*** (0.000)	0.003*** (0.001)	0.017*** (0.003)	0.001 (0.001)	0.002** (0.001)
GM F.E.	0.005** (0.002)	0.001 (0.001)	0.004 (0.003)	0.001*** (0.000)	0.003* (0.002)	0.014** (0.006)	0.009*** (0.003)	0.001 (0.001)
Mean of Dep.Var.	0.054	0.006	0.099	0.0006	0.032	0.504	0.912	0.0047
Std. Dev.	0.226	0.068	0.298	0.025	0.178	0.499	0.287	0.0689
Observations	346,962	346,962	328,639	346,962	330,956	246,611	346,962	63,823

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). The first row report OLS estimates, the second row includes grand-mother fixed effects. All estimates include the same set of controls used in column 4 of Table A.2. Standard errors are clustered at the grand-mother level.

Table A.3: Other Birth Outcomes, Natality Detail Data

Dependent Variable:	(1) Low Birth Weight	(2) Very Low Birth Weight	(3) Premature Birth	(4) Complicated Pregnancy	(5) Macrosomic	(6) Assisted Ventilation (More than 30min)	(7) Very Low Apgar
Non-Hispanic Father	0.002***	0.000	0.004***	0.015***	0.010***	0.001***	0.001***
	(0.001)	(0.000)	(0.001)	(0.003)	(0.001)	(0.000)	(0.000)
Sociodemographic Controls	YES	YES	YES	YES	YES	YES	YES
State F.E.	YES	YES	YES	YES	YES	YES	YES
Mean of Dep.Var.	0.070	0.012	0.128	0.256	0.073	0.006	0.005
Std. Dev.	0.255	0.108	0.334	0.436	0.260	0.079	0.068
Observations	2,326,886	2,326,886	2,126,805	2,326,242	2,326,886	2,219,688	833,451

Notes - Data are drawn from the Natality Detail Data (1989-2004). Socio-demographic controls include indicators for child's gender, indicators for adequate prenatal care, marital status, mother's education, parity, dummies for mother's and child's year of birth, state fixed effects. Standard errors are clustered at the MSA level.

Table A.4: Socio-economic status and smoking during pregnancy

Dependent variable	(1) Smoking during pregnancy	(2) Smoking during pregnancy	(3) Smoking during pregnancy	(4) Smoking during pregnancy
	Panel A: Non-Hispanic White Natives			
High education	-0.178***	-0.113***	-0.0261***	-0.0284***
	(0.00149)	(0.00165)	(0.00339)	(0.00347)
Mean of Dep.Var	0.181	0.181	0.181	0.181
Std. Dev.	0.384	0.384	0.384	0.384
Observations	350,772	350,772	350,772	350,772
	Panel B: Second-generation Hispanics			
High education	-0.0240***	-0.0127***	-0.00747	-0.00650
	(0.00199)	(0.00213)	(0.00472)	(0.00503)
Mean of Dep.Var	0.0201	0.0201	0.0201	0.0201
Std. Dev.	0.141	0.141	0.141	0.141
Observations	33,834	33,834	33,834	33,834
Socio-demographic controls	NO	YES	NO	YES
Grand-mother F.E.	NO	NO	YES	YES

Notes - Data are drawn from the Florida Vital Statistics (1989-2009). Socio-demographic controls include 3rd generation child's gender, 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, an indicator for prenatal care, marital status, dummies for the interaction of county and year of birth of third-generation children, indicators for missing information on adequacy of prenatal care, and mother's birth weight. Columns 3 and 4 include grand-mother fixed effects. Standard errors are clustered at the grandmother level.

Table A.5: Other Robustness Checks

Methodology	GM F.E.	G.M. Q.F.E.	PROBIT	Heckman-Correction Model
Non-Hispanic father	0.005**	0.006***	0.007***	0.006***
	-0.002	0.001	(0.002)	(0.001)
Mean of the Dep. Var.	0.054	0.054	0.054	0.054
Std.Dev.	0.226	0.226	0.226	0.226
Observations	346,962	346,962	346,962	346,962

Notes - Data are drawn from the California and Florida Vital Statistics (1970–1985, 1989–2009). All estimates include the same set of socio-demographic controls used in column 4 of Table 2. Column 1 includes grand-mother fixed effects. Columns 2-4 include grand-mother level averages for education, adequacy of prenatal care, maternal age, number of children and zip-code level characteristics. Standard errors are clustered at the grand-mother level.