SSM -Population Health 2 (2016) 84–89

Contents lists available at ScienceDirect

SSM -Population Health

journal homepage: www.elsevier.com/locate/ssmph

Article

The Hispanic health paradox: New evidence from longitudinal data on second and third-generation birth outcomes

Osea Giuntella *

Blavatnik School of Government, University of Oxford

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form
15 February 2016
Accepted 17 February 2016

Keywords:
Healthy immigrant effect
Birth outcomes
Longitudinal data
Intergenerational

ABSTRACT

This study examines the birth weight of second and third-generation Hispanics born in California and Florida, two of the major immigrant destination states in the US. I exploit a unique dataset of linked birth records for two generations of children born in California and Florida (1970–2009) and linear probability models to investigate the generational decline in the birth outcomes of Hispanics in the US. The data allow using an extensive set of socio-demographic controls and breaking down the results by country of origin. Second-generation children of Mexican and Cuban origin have better birth outcomes than children of US-born white women. Children of Puerto Rican origin have instead worse birth outcomes. The advantage observed among second-generation Hispanics erodes substantially in the third generation but third-generation Mexicans retain some of it.

1. Introduction

Despite a lower socio-economic status, children of first-generation immigrants of Hispanic origin have better birth outcomes than children of US born white women (Acevedo-Garcia, Soobader, & Berkman, 2007). However, previous studies have shown that birth outcomes deteriorate in later generations (Acevedo-Garcia, Soobader, & Berkman, 2005), despite socio-economic assimilation (Duncan & Trejo, 2015; Teitler, Martinson, & Reichman, 2015). These facts are commonly referred to as the Hispanic Health Paradox which has been observed with respect to several health outcomes (Black, Devereux, & Salvanes, 2007; Markides & Coreil, 1986).

Poor health at birth has long-term consequences on adult health, socio-economic outcomes and it is associated with increased health care costs (Lewit, Baker, Corman, & Shiono, 1995). Second generation births have surpassed immigration as the main driver of the dynamic growth of the American population. Hispanics are by far the largest ethnic group in the US and children of Hispanic origin are a majority of newborns in many US counties (Passel, Cohn, & Lopez, 2011). Thus, understanding the health trajectories of Hispanics in the US is paramount to understand the health of next generation Americans and to address health disparities in the population (People, 2013).

* Correspondence address: University of Oxford, Blavatnik School of Government, Radcliffe Observatory Quarters, Woodstock Road, Oxford OX2 6GG. UK.
E-mail address: osea.giuntella@nuffield.ox.ac.uk

Most of the extant evidence on the Hispanic Paradox is based on cross-sectional data using synthetic cohorts or short panel surveys that do not allow a longitudinal analysis across generations (Antecol & Bedard, 2006; Hummer, Powers, Pullum, Gossman, & Frisbie, 2007; Riosmena, Wong, & Palloni, 2013; Shaw & Pickett, 2013). Furthermore, often the data does not allow analyzing the intergenerational health trajectories of immigrant descendants by country of origin. Hispanic ethnicity encompasses individuals coming from different backgrounds and migration histories. While several studies pointed out the need to analyze health trajectories of immigrants across generation (Jasso, Massey, Rosenzweig, & Smith, 2004), to the best of my knowledge there is no paper analyzing the Hispanic Health Paradox using individual linked data on two generation of immigrant descendants.

This study exploits a unique data drawn from administrative records of California and Florida Vital Statistics to fill this gap in the literature. Furthermore, the large sample size of immigrants allows me to conduct country specific analysis and rely on a broad set of control characteristics.

2. Methods

The primary data used in this study are drawn from the Birth Statistical Master File provided by the Office of Vital Records of the California Department of Health and the Birth Master Dataset provided by the Bureau of Vital Statistics of the Florida Department of Health. These data contain information extracted from the

http://dx.doi.org/10.1016/j.ssmph.2016.02.013

birth certificates of all children born in the years 1970–1985 (1970–1981 in California and 1971–1985 in Florida) and in the years 1989–2009. For expositional simplicity, I refer to all women giving birth between 1970 and 1985 as the first generation (grandmothers, G1); to all the children born between 1970 and 1985, as the second generation (G2); and to all the children born between 1989 and 2009 as the third generation (G3). I use this definition for both immigrant and natives. The socio-demographic information (e.g., age at delivery, education etc.) on second generation mothers (G2) is drawn from the third generation records (G3), while the socio-demographic information on the first generation grandmothers (G1) is drawn from the second generation (G2) birth records.

Information about the mother country and state of birth, the mother first and maiden name, the child full name, date of birth, gender, parity, race, birth weight, hospital of birth, and county of birth are available in both states for the full period considered. However, not all variables are available in each year and for each of the two states. For instance, the mother age is reported for the entire period in California but only after 1989 in Florida, whereas the mother education is reported for the entire period in Florida but only since 1989 in California. Information on birth weight is available for the entire period in both states, whereas other important measures of health at birth (e.g., Apgar score, gestational length, etc.) are unfortunately only available in more recent years. Though a few studies cast doubt on the notion that birth weight has a causal effect on mortality in particular and infant health more generally (Wilcox, 2001; Almond, Chay, & Lee, 2005), there is a general consensus that low birth weight (conventionally defined as a birth weight lower than 2500 g) is an important marker of health at birth and strongly associated with increased mortality and morbidity risk (Paneth, 1995; Conley & Bennett, 2000; Currie, 2011). Because this study does not analyze the effects of birth weight and birth weight is the only measure of birth outcomes available for the entire period, I will primarily focus on birth weight and the incidence of low birth weight as indicators of health at birth.

As in the previous literature employing administrative birth records (Fryer & Levitt, 2004; Currie & Moretti, 2007; Royer, 2009), I am able to link information available at a woman birth to that of her children, if the woman is born in California (Florida) and also gave birth in California (Florida).

One of the typical drawbacks of administrative vital statistics is the lack of information on individual income and occupation. However, the data contain certain information on parental education, hospital zip code, and the mother residential zip code. Data on zip code socio-demographic and economic characteristics are drawn from the U.S. Census (source: Social Explorer).

To construct the intergenerational sample, I linked the records of children (G2) born to first-generation (G1) mothers between 1970 and 1985 to the records of their own third-generation children (G3) born in California and Florida between 1989 and 2009. The matching is performed using the second-generation mother first and maiden names, date of birth, and state of birth.

I restrict the empirical analysis to children born between 1970 and 1985 to white mothers and Hispanic first-generation immigrant mothers coming from Cuba, Puerto Rico, and Mexico. I exclude children of Hispanic first-generation women born in countries besides Cuba, Puerto Rico, and Mexico as these are the only countries for which information on mother's country of birth is explicitly reported.

To verify the paradox within the longitudinal data, I estimate a 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. Formally, I consider the following model:

$$H_{izt,2} = \alpha + \beta Hisp_{izt,G_1,} + \gamma X_{izt,G_1} + \tau_{t,G_2} + \zeta_{z,G_2} + \varepsilon_{izt,G_2}$$

where the subscripts G1 and G2 represent the first and second generations, respectively.

The parameter $H_{izt,2}$ is the birth outcome (such as birth weight, incidence of low birth weight, etc.) of the second-generation child i (for both females and males), whose mother resided (or delivered) in zip code z at time t. The variable $Hisp_{izt,G_1,}$ is a dummy equal to one when the first-generation woman delivering between 1970 and 1985 was born in Cuba, Mexico, or Puerto Rico. The set of individual socio-demographic characteristics of the first-generation mothers is delineated in X_{izt,G_1}, including education (high school dropout, high school graduate, some college, and college or more), marital status, parity, race, age dummies (in Florida, the mother age is not available for the period 1970–1985), an index of the adequacy of prenatal care based on the month in which prenatal care began, father age (quadratic), father education (high school dropout, high school graduate, some college, and college or more), child gender, and type of birth (singleton vs.

Table 1
Descriptive statistics.

	2nd Generation (G2),1970–1985			3rd Generation (G3),1989–2009		
	Mean	*S.d*	*Observations*	*Mean*	*S.d.*	*Observations*
Female child	*0.49*	*0.50*	*4,704,571*	*0.49*	*0.50*	*2,076,487*
Marital status (apparent status)	*0.88*	*0.33*	*4,704,273*	*0.81*	*0.39*	*2,076,438*
Adequate prenatal care	*0.62*	*0.49*	*4,514,266*	*0.85*	*0.35*	*2,037,796*
Parity	*1.23*	*1.47*	*4,633,073*	*0.91*	*1.10*	*2,073,619*
Multiple birth	*0.02*	*0.13*	*4,679,958*	*0.03*	*0.16*	*2,076,487*
Maternal age	*24.92*	*5.26*	*3,312,788*	*24.72*	*5.06*	*2,076,340*
Paternal age	*27.81*	*6.32*	*3,229,460*	*27.55*	*6.09*	*1,924,394*
Maternal education						
Less than high-school	*0.26*	*0.44*	*1,287,632*	*0.22*	*0.41*	*2,050,522*
High-school degree	*0.43*	*0.49*	*1,287,632*	*0.36*	*0.48*	*2,050,522*
Some college	*0.19*	*0.40*	*1,287,632*	*0.24*	*0.42*	*2,050,522*
College	*0.12*	*0.33*	*1,287,632*	*0.18*	*0.38*	*2,050,522*
Paternal education						
Less than high-school	*0.21*	*0.40*	*1,193,534*	*0.21*	*0.41*	*1,804,608*
High-school degree	*0.39*	*0.49*	*1,193,534*	*0.42*	*0.49*	*1,804,608*
Some college	*0.20*	*0.40*	*1,193,534*	*0.19*	*0.39*	*1,804,608*
College	*0.21*	*0.40*	*1,193,534*	*0.18*	*0.39*	*1,804,608*

Notes: Data are drawn from the California and Florida Vital Statistics, (1970–1985, 1989–2009).