Disparate Disparities: Understanding Differences in Infant Mortality across Racial and Ethnic Groups*

Todd E. Elder, John H. Goddeeris, and Steven J. Haider

Michigan State University
Department of Economics
110 Marshall-Adams Hall
East Lansing, MI 48824

February 3, 2010

Abstract. We analyze the disparity between several race and ethnic groups and whites in a fundamental measure of population health: the rate at which infants die. Using micro-level Vital Statistics data from 2000 to 2004, we separate mortality disparities into three temporal components, and we assess the extent to which these components are predictable on the basis of observable background characteristics. The temporal patterns vary substantially across racial and ethnic groups: relative to whites, the high infant mortality rates of blacks and Puerto Ricans are primarily driven by disadvantages in fitness at birth, but the high infant mortality rate of Native Americans is driven by excess infant deaths during the post-neonatal period. Differences across races and ethnicities in conditional post-neonatal mortality are largely predictable from variation in background characteristics, particularly maternal marriage, education, and age In contrast, little of the fitness disadvantage among blacks and Puerto Ricans is predicted by background characteristics.

1. Introduction

* We thank Soren Anderson, Stacy Dickert-Conlin, Nigel Paneth, Gary Solon and seminar participants at Michigan State University, University of California-Davis, the University of Chicago, and the University of Michigan for helpful comments. Emails: telder@msu.edu, goddeeri@msu.edu, haider@msu.edu.

The infant mortality rate (IMR), the number of deaths in the first year of life per 1000 live births, is a widely cited indicator of population health. In 2006, the overall IMR for the United States was 6.68, but mortality rates varied widely across racial and ethnic groups (Matthews and MacDorman, 2010). Non-Hispanic blacks had the highest IMR at 13.35, compared to 5.58 among non-Hispanic whites. Among other race and ethnic groups, the IMRs among American Indians / Alaska Natives (8.28) and Puerto Ricans (8.01) were greater than that of non-Hispanic whites, while the IMRs for Mexicans (5.34), Central/South Americans (4.52), Cubans (5.08), and Asians / Pacific Islanders (4.55) were lower.[1]

Researchers and policymakers have devoted substantial attention to some of these IMR disparities, with a particular focus on the black-white gap. Studies have consistently found that most of the black-white gap is already apparent at birth, as evidenced by large black-white differences in the distributions of birth weight and gestational age. Very little of these differences in fitness at birth can be predicted by standard observable characteristics (for a more detailed discussion and literature review, see Elder, Goddeeris, and Haider, 2010a; hereafter EGH). Additionally, several studies have examined the relatively low IMRs of most Hispanic groups (e.g., Hummer, Powers et al., 2007). Because infant health is positively associated with education and economic status, the fact that most Hispanic groups experience lower infant mortality than non-Hispanic whites poses a real puzzle. Most of the other racial and ethnic IMR disparities, while noted in passing, are rarely studied in depth.

In this paper, we use U.S. micro-level Vital Statistics data from 2000 to 2004 to examine differences in infant mortality across a wide variety of racial and ethnic groups. We undertake this analysis for two primary reasons. First, the large IMR disparities have been remarkably

---

[1] The groups listed here are those that are identified in Vital Statistics data reported by all states between 2000 and 2004.

persistent over time, and researchers have made relatively little progress in understanding their underlying sources. A careful comparison across racial and ethnic groups, including those with relatively low IMRs, may shed light on the underlying sources of the disparities. Second, despite steady growth in the share of births to groups other than non-Hispanic whites and blacks, researchers have continued to largely focus on the black-white gap, suggesting a need for more attention to the infant health of Hispanics, Asians / Pacific Islanders, and American Indians / Alaska Natives.[2] Our analysis encompasses all of these groups. In the case of Hispanics, we consider separately the relatively high-mortality group Puerto Ricans and the relatively low-mortality group Mexicans.

To examine the differences in IMR among the various groups, we adopt the unified approach developed in EGH for studying the black-white IMR gap. This approach includes specifying a decomposition of the IMR gap into three distinct temporally-ordered components that are related to concepts frequently examined in previous studies: fitness at birth, conditional (on fitness) neonatal mortality, and conditional post-neonatal mortality. We then apply a reweighting procedure to examine how the IMR gap and each of its temporal components are related to an extensive set of background characteristics available on birth certificates.

Our analyses reveal several important findings. First, the temporal patterns in IMR disparities vary substantially across racial and ethnic groups. As is well-established for the black-white IMR gap, the Puerto Rican-white gap is largely driven by differences in fitness at birth. In contrast, the relatively high IMR among American Indian / Native Alaskans (hereafter, "Native Americans") primarily emerges during the post-neonatal period – white and Native

[2] Between 1996 and 2006, the share of births to non-Hispanic whites and non-Hispanic blacks fell from 60.6 to 54.1 percent and from 14.9 to 14.5 percent, respectively. In contrast, the share of births to Hispanics grew from 18.0 to 24.4 percent, the share to American Indians / Alaska Natives grew from are 1.0% and 1.1%, and the share to Asians grew from 4.3% and 5.7% (Martin, Hamilton et al., 2008).

American infants have similar distributions of birth weight and gestational age. The Mexican IMR advantage relative to whites is modest, and its composition depends on the measure of fitness used. The Asian IMR advantage develops during all three temporal periods, but it is especially driven by lower neonatal and post-neonatal mortality rates conditional on fitness at birth. Second, the predictability of mortality gaps based on background characteristics also varies substantially across groups. For example, blacks and Puerto Ricans have much higher IMRs than would be predicted based on background characteristics. The large fitness disadvantage of Puerto Ricans is essentially unpredicted. The overall Native American-white and Asian-white IMR gaps are predicted well by background characteristics, with the predicted temporal patterns roughly matching the actual temporal patterns. However, the predicted components do not match the actual mortality patterns of Mexicans at all. Specifically, the large differences between whites and Mexicans in characteristics such as education and marital status imply that whites should have large advantages in IMR in the post-neonatal period. In fact, Mexican have slightly *lower* conditional post-neonatal mortality rates than whites, suggesting that the “Hispanic paradox” (as noted by Hummer, Powers et al., 2007, among others) is primarily a post-neonatal phenomenon.

Finally, we have isolated the roles that individual observable characteristics play in the predicted gaps. Among the three groups that are most disadvantaged relative to whites – blacks, Native Americans, and Puerto Ricans – relatively low marriage rates and educational attainment among mothers are most responsible for producing high predicted IMRs, with an unfavorable age distribution among mothers playing a smaller but still significant role. The combined effect of all other background characteristics, including measures of prenatal care, plurality, birth order, and the state in which the birth occurred, is relatively small in all cases.

2. Background and Literature Review

In examining differences in infant mortality, either across groups or over time, researchers have found it useful to distinguish between differences in the fitness of infants at birth and differences in survival rates conditional on fitness. One reason for making such a distinction is that policy approaches to improving fitness might be quite different from approaches to reducing conditional mortality. Presumably the former involves factors that happen before birth, such as the environment, health and behavior of the mother, while the latter may have more to do with medical care during and after birth and the home environment of the infant.

Though recognized as imperfect, the most commonly used measures of fitness are birth weight and gestational age. For example, Lee, Paneth et al. (1980) found that the decline in neonatal mortality (mortality in the first 28 days of life) between 1950 and 1975 was entirely due to declines in mortality conditional on birth weight. Other studies using more recent data have reached similar conclusions, finding that reductions in mortality conditional on fitness have been overwhelmingly if not entirely responsible for declines in IMR over time (e.g., Carmichael and Iyasu (1998), Schempf, Branum et al. (2007), and Alexander et al. (2008)). EGH show that these conclusions are insensitive to whether gestational age or birth weight is used to proxy for fitness at birth.

Another analytic distinction common in the literature is between mortality during the neonatal period (the first 28 days) and the post-neonatal period (the remainder of the first year). In examining black-white IMR gaps, both Carmichael and Iyasu (1998) and Schempf, Branum, et al. (2007) decompose the gaps in neonatal mortality and post-neonatal mortality into fitness at birth components and mortality conditional on fitness at birth components. They find that fitness differences can more than fully account for black-white gaps in neonatal mortality but not post-neonatal mortality, which suggests that post-neonatal gaps are related to factors that occur after birth. Wise (2003) provides a useful conceptual discussion of black-white disparities in neonatal and post-neonatal mortality.

A number of researchers have also examined whether IMR differences across groups can be predicted based on differences in the background characteristics of group members. Examples include Eberstein, Nam, and Hummer (1990), Hummer, Biegler et al. (1999), and Frisbie, Song et al. (2004). These studies typically use logit models with micro data, with infant death as the outcome variable and controls for various background characteristics and for racial/ethnic group. One focus is on whether group effects are reduced or eliminated by the inclusion of background variables. Typically, group effects remain significant and a large black-white gap exists even when the background variables are included.[3]

EGH develop a reweighting method that integrates the decomposition of IMR differences into three temporal components (fitness at birth, neonatal mortality, and post-neonatal mortality) with an analysis of the role of background characteristics. They find that little of the black-white gap in fitness at birth can be predicted based on background characteristics, but most of the gap in the post-neonatal component can be predicted. They also find that the ability to predict the black-white IMR gap is declining over time, largely because the mortality gap among extremely low-fitness infants is increasingly unrelated to background characteristics.

A growing literature has examined the IMR gap between whites and Hispanics, consistently finding that Hispanics have similar (or slightly lower) infant mortality rates as non-Hispanic whites. Frisbie and Song (2003) analyze mortality and indicators for short gestational age and

[3] Sometimes researchers will control for birth weight or gestational age in models of mortality. In such models, black-white mortality gaps can be eliminated or reversed, but the question of why the differences in infant fitness exist remains open.

low birth weight, differentiating Hispanics by country of origin and birthplace of the mother. They find that although births to Hispanic mothers are uniformly more likely to be of short gestation than births to non-Hispanic white mothers, most Hispanic groups have lower IMR than whites, with particularly large advantages for foreign-born Mexican mothers. Hummer, Powers, et al. (2007) find a similar pattern, with the IMR among Mexican immigrants being 14 percent lower than that of native non-Hispanic whites, and they examine whether the relative advantage of Hispanics can be explained by selective out-migration. They reject this explanation on the grounds that much of the advantage develops within one day of birth, before much out-migration could occur, and instead point to the relative advantage as being additional evidence of the general finding that Hispanics are relatively advantaged along various dimensions of health.

Relatively little work has used population-based data to analyze the high infant mortality of Native Americans. Watson (2006) finds that federal investments in sanitation projects on Indian reservations played an important role in the convergence of infant mortality rates between Native Americans and whites between 1960 and 2001. Tomashek, Qin et al. (2006) compare infant mortality by birth weight between whites and Native Americans in 1989-91 and 1998-2000, finding that most of the higher Native American mortality can be traced to higher post-neonatal mortality in all birth weight groups.

3. Methods

3.1. Decomposing IMR Differences into Temporal Components

Studies of infant mortality rates routinely examine birth weight and gestational age distributions, neonatal morality, which captures deaths in the first 28 days of life, and post-neonatal mortality, which captures deaths during the remainder of the first year. We use a three-way decomposition of IMR differences that recognizes the temporal ordering of these three outcomes, following EGH.

For concreteness, we discuss a decomposition that uses birth weight to measure fitness at birth, but the same formulas apply if infants are categorized by gestational age. Let *g* designate a racial or ethnic group. Partitioning births into *K* mutually exclusive categories of birth weight, we may write the infant mortality rate of group *g* as

$$(1) \qquad IMR_g = \pi_g^{n}{}' s_g + \pi_g^{p}{}'[(1-\pi_g^n)\bullet s_g]$$

where s_g is a $K\times1$ vector of the shares of births in different birth weight categories, π_g^n is a $K\times1$ vector of birth weight-specific neonatal mortality rates, π_g^p is a $K\times1$ vector of birth weight-specific post neonatal mortality rates for those infants who survive the neonatal period, and the dot operator "$\bullet$" denotes element-by-element vector multiplication.[4] The vector $[(1-\pi_g^n)\bullet s_g]$ contains the fractions of births in each birth weight category who survive the neonatal period, so that π_g^p has a standard hazard formulation.

Based on (1) and standard decomposition techniques, the difference in the infant mortality rate between group *B* and group *A* (the reference group) may be written as

$$(2) \qquad IMR_B - IMR_A = \pi_B{}'(s_B - s_A) + (\pi_B^n - \pi_A^n)'[(1-\pi_B^p)\bullet s_A] + (\pi_B^p - \pi_A^p)'[(1-\pi_A^n)\bullet s_A]\ .$$

The first component in (2) isolates the role of differences in the birth weight distributions $(s_B - s_A)$, the second component isolates the role of differences in conditional neonatal mortality rates ($\pi_B^n - \pi_A^n$), and the third component isolates the role of differences in conditional post-

[4] Infant mortality rates are commonly reported per 1000 live births. We report them in that way as well, but in mathematical expressions we treat π's as probabilities per live birth.

neonatal mortality rates among those infants who survive the neonatal period ($\pi_B^p - \pi_A^p$). As is typical with these types of decompositions, alternative representations involving similar components exist. See EGH for alternative representations and empirical evidence that the different representations give quantitatively similar results for the black-white gap.

This three-way decomposition is useful because it separates the infant mortality rate into three temporally-ordered components that are frequently discussed in the literature: fitness differences apparent at birth, mortality differences during the first 28 days, and mortality differences during days 29 through 365. However, we stress that this decomposition reflects when evidence of fitness/mortality differences are observed, not when the underlying causes arise. In other words, neonatal mortality conditional on birth weight still reflects processes that began *in utero*, while post-neonatal mortality reflects processes that began *in utero* and during the first 28 days of life.

3.2. Assessing the Combined Effect of All Background Characteristics

The decomposition of IMR gaps into their temporal components provides important information about how gaps differ across groups, but it does not tell us why they differ. In this section, we explain our methods for examining how background characteristics affect infant mortality and its various components. We first consider the combined effect of all background characteristics.

Our approach (laid out in more detail in EGH) is based on a reweighting procedure pioneered by DiNardo, Fortin, and Lemiuex (1996; hereafter DFL).[5] The intuition for the method is straightforward: to measure the influence of differences across groups in the distributions of

[5] Several studies have assessed the asymptotic and finite-sample properties of reweighting methods, including Hirano, Imbens, and Ridder (2003), Wooldridge (2007), and Busso, DiNardo, and McCrary (2009), concluding that their performance in most circumstances is comparable to or better than regression-based and matching methods.

characteristics, we reweight the infants in group *A* so that their distribution of characteristics closely matches that of one of the other groups.

Formally, let $f(y \mid g)$ be the probability density of an outcome *y* for group *g* and let $F(x \mid g)$ be the cumulative distribution of background characteristics *x* for group *g*. We may write

$$(3) \qquad f(y \mid g) = \int_x f(y \mid g, x) dF(x \mid g) \equiv f(y; g_{y|x}, g_x),$$

expressing $f(y \mid g)$ as a density conditional on *x* integrated over the distribution of *x* of individuals who are in group *g*. This formulation highlights the potential for creating counterfactual densities by using the distribution of characteristics associated with different groups. To see this, define

$$(4) \qquad f(y; g_{y|x} = A, g_x = B) \equiv \int_x f(y \mid g = A, x) dF(x \mid g = B)$$

as the distribution of outcomes that would result if group *A* retained its own mapping from characteristics to outcomes, but had the group *B* distribution of characteristics.

The key insight of DFL is that (4) can be estimated as integral over a weighted function of the actual group A data, with weights that are simple to construct. Specifically,

$$(5) \qquad f(y; g_{y|x} = A, g_x = B) \equiv \int_x f(y \mid g = A, x) \psi_{A \to B}(x) dF(x \mid g = A),$$

where $\psi_{A \to B}(x)$ is defined as

$$(6) \qquad \psi_{A \to B}(x) \equiv \frac{dF(x \mid g = B)}{dF(x \mid g = A)} = \frac{\Pr(g = B \mid x)}{\Pr(g = A \mid x)} \times \frac{\Pr(g = A)}{\Pr(g = B)}.$$

The last equality in (6) follows from Bayes' Rule. The first fraction to the right of the equality can be estimated using a binary model (such as a logit or probit) of group membership as a function of covariates *x*, while the second fraction involves only the sample proportions of individuals in each group.

In practice, whites will play the role of group *A* and the other racial and ethnic groups will act as group *B*. For each of these other groups in turn, we pool its data with the data for whites and estimate a logit function to predict group membership as a function of *x*. We use the results to construct weights as in (6) for each observation in the white population. In these hypothetical, weighted populations (e.g., "white infants weighted to have the background characteristic distribution of blacks"), we can examine quantities such as IMR gaps relative to whites and the various components of these gaps. For example, the gap between the white IMR and the hypothetical IMR constructed by weighting whites to have the black distribution of characteristics is an estimate of how much of the black-white IMR gap is predictable based on differences in characteristics between these groups.[6]

3.3. Assessing the Role of Individual Background Characteristics

In addition to predicting differences across groups based on differences in the entire distribution of background characteristics, we would like understand the role of particular characteristics, such as mother's education or age. We rely on the reweighting methods developed in Elder, Goddeeris and Haider (2010b).[7] The method is a reweighting analog of

---

[6] Reweighting another group's population to match the characteristics of whites leads to a different estimate of the share of that group's IMR gap predicted by background characteristics. See EGH for examples of such calculations involving the black-white IMR gap. In this paper we present only estimates using whites weighted to match the characteristics of other groups. One reason for taking this approach is that we have found empirically that reweighting whites (the largest group) results in more precise estimates than reweighting smaller populations, such as blacks or Hispanics.

[7] A simpler approach would use weights that depend only on mother's education to create a counterfactual white population with, for example, the black distribution of education. This approach, which is analogous to estimating a

examining the effects of one regressor at a time using regression, answering questions like "What would be the white birth weight distribution if white mothers had the black distribution of education while retaining their own joint distribution of all other background characteristics?"

To apply the method, we partition the set of background characteristics *x*, into two parts, *z* and x_{-z}. The variable being switched from the group *A* distribution to the group *B* distribution is *z* (e.g., *z* is a categorical variable where each level of education is a different category), with all other background characteristics being in x_{-z}. We construct weights so that the reweighted (counterfactual) population has group *B*'s marginal distribution of *z* and group *A*'s distribution of x_{-z}. Specifically, the weights take the form

$$(7) \qquad \psi^{z}_{A \to B}(z, x_{-z}) = \frac{dF(z \mid j = B)}{dF(z \mid x_{-z}, j = A)}.$$

Intuitively, by replacing the distribution of *z* conditional on x_{-z} in group *A* with the unconditional distribution of *z* from group *B*, the reweighted group *A* population will keep its own distribution of x_{-z} but take on the unconditional distribution of *z* from group *B*. Each covariate successively plays the role of *z*.

We then calculate the role of *z* by performing the decomposition in (2), for example, with the counterfactual population taking the place of group *B*.[8] See EGH (2010b) for a detailed comparison of this method to other methods, including details about how the weights are constructed.

---

regression model including only education as an explanatory variable, has a serious drawback – the estimated role of education is not isolated from that of covariates that are correlated with education. Specifically, such an approach attributes to education both the effect of differences across groups in education and the effect of differences in other characteristics that are correlated with education.

[8] As discussed in EGH (2010b), some modifications may need to be made to expression (7) when interactions are important and when a support condition is not satisfied. Both are relevant to the case at hand. For results presented in this paper, we use the formulation that subtracts off the covariance term (see Sec. 4.1), setting the negative weights to zero, and adjusts violations of the support condition (see Sec. A3).

4. Data

Our data source is the linked birth/infant death cohort data compiled by the National Center for Health Statistics (NCHS).[9] We pool the five most recent publicly available annual datasets, from 2000 through 2004. The cohort data sets include information from the birth certificates of all live births occurring in the U.S. in the relevant calendar year, linked to death certificates for all infants who die within their first year of life. We limit our analysis to births that occur in the fifty U.S. states or the District of Columbia. NCHS is unable to match a small fraction of death certificates to birth certificates (about 1 percent in 2001); we ignore the unmatched deaths for our analysis.

As is standard in the literature, we classify births based on the race and ethnicity of the mother. From 2000 to 2004 all states classified births into at least four racial categories: White, Black, Native American, and Asian.[10] The data also distinguish between those who report Hispanic ethnicity and those who do not, defining five Hispanic groups by place of origin: Mexico, Puerto Rico, Cuba, Central or South America, and other or unknown origin.[11] Among Hispanics, we include mothers who report their place of origin as Mexico or Puerto Rico, excluding the other origins and those who report missing or unknown origin. Based on this information, we analyze six mutually exclusive categories of births: non-Hispanic White, non-Hispanic Black, Hispanic of Mexican origin, Hispanic of Puerto Rican origin, Asian, and Nativie Americans/Alaska Natives. For simplicity, we refer to these groups as whites, blacks,Mexicans,

---

[9] See information at http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm.
[10] Most states distinguish among at least several subcategories of Asian or Pacific Islander (NCHS, 2005) and infant mortality differs slightly across subgroups. Because not all states report in the same way, in this paper we consider only the aggregated category.
[11] "Origin" does not imply place of birth of the mother. All groups can, however, include infants born to mothers who were themselves born outside the U.S., including infants born in the U.S. to noncitizens.

Puerto Ricans, Asians, and Native Americans, respectively . When considering IMR disparities, we compare each of the latter five groups to the white population.

For sufficient statistical power for the smaller race/ethnic groups, we pool births from 2000 to 2004.[12] The smallest group, Native Americans, includes about 184,000 observations. For computational reasons, we use random samples for the largest racial/ethnic groups: 20 percent for whites, blacks and Mexicans, and 70 percent for Asians. This sampling scheme gives us the largest sample for whites (over 2.25 million) and about 600,000 observations each for blacks, Mexicans and Asians. We exclude observations with missing information on race or ethnicity, maternal education, prenatal care, birth order, and previous pregnancy loss.

We consider two measures of fitness at birth: birth weight and gestational age. When using birth weight, we divide births into cells by individual ounce, leading to 173 cells (the lowest being 9 ounces).[13] For comparability with other research, we express birth weight in grams when displaying or discussing our results. We disaggregate gestational age by integer values of completed weeks. We use the NCHS-edited gestational age variable provided in the public data files, which uses the mother's reported last menstrual period when it is both available and consistent with birth weight, and otherwise uses a clinical estimate. The small share of infants for whom birth weight or gestational age is not reported is omitted from the relevant figures but is included as a separate category in the analyses reported in tables.

---

[12] Appendix Figure A1 shows the year-specific race/ethnicity infant mortality rates from 2000 to 2004. The mortality rate is remarkably stable during these years, which motivates our use of the pooled data.

[13] While the research literature typically reports birth weight in grams (as do the cohort data sets), birth weight is commonly measured in pounds and ounces in the U.S. A close examination of the data reveals a pronounced heaping of observations at gram values that correspond to ounce values, so we round gram values to the nearest ounce when grouping by birth weight.

Background characteristics. Conceptually, we consider as background characteristics those observable attributes that are determined prior to information the mother might have received about the fitness of the fetus. Such predetermined characteristics can provide important insight into the factors that cause IMR disparities. In contrast, characteristics that are not predetermined may reflect differences in fetal/infant health that have already emerged. For example, information that a pregnancy is at high risk may lead to a greater number of prenatal visits, inducing a negative association between prenatal care and outcomes and obscuring any causal effect of prenatal care on fitness. Similarly, birth weight and gestational age do not satisfy our conceptual requirement of background characteristics because they measure underlying fitness, but tell us little about the processes that determine fitness.

It is important to recognize that associations between background characteristics and outcomes are only a starting point for understanding the causal mechanisms at work. For example, educational attainment may be associated with lower infant mortality because education imparts knowledge and income that aid in the production of a healthy infant, but the association might also reflect the influence of omitted maternal characteristics that lead to both more schooling and healthier infants.

Implementing our conceptual definition of predetermined characteristics is not always straightforward due to data limitations and our desire to connect to the previous literature. We include variables that are commonly used in previous studies and clearly predetermined to information on infant fitness: maternal education, maternal age, previous pregnancy loss (either elective or spontaneous), infant sex, live birth order, and plurality.[14] Two others that we

[14] We specify education with four indicator variables (<12 years, 12 years, 13-15 years, and >15 years), maternal age with six indicator variables (<20, 20-24, 25-29, 30-34, and >34), and live birth order with five indicator

examine, prenatal care and marital status, are less clearly predetermined but are often included in previous studies. In light of our concerns about prenatal care, we use only an indicator variable for whether it is begun in the first trimester. Marital status is measured at the time of birth, so it could potentially be affected by information on health of the fetus.[15]

One way in which this study differs from most previous studies is in including indicator variables that identify the state (or the District of Columbia) in which a birth occurs. Racial and ethnic groups are distributed very differently across states. Additionally, many important inputs for the production of healthy infants vary by geography, such as employment opportunities, social services, pollution, and health care access and quality. If some groups tend to live in states that are less conducive to the production of healthy infants than do others, these state indicators would predict some of the IMR gaps. Another geography-related variable that is available in the public-use Vital Statistics data in most years is whether a birth occurred in a county with a population greater than 250,000 people; we can only include this variable in some of our analysis because the information is not included in the 2002 public-use data set.

5. Racial/Ethnic Gaps in IMR and its Temporal Components

Table 1 presents basic descriptive statistics on measures of infant mortality and background characteristics in our analysis sample, with the infant mortality gaps graphed in Figure 1. The overall infant mortality rate of whites for the years 2000-2004 in our sample was 5.35 per 1000 live births. Roughly two-thirds of these deaths (3.51 / 5.35) occurred in the neonatal period, with

---

variables (1st, 2nd/3rd, >3rd). We use single indicators for whether the infant is male, whether the birth was plural, and whether the mother experienced a previous pregnancy loss (either elective or spontaneous).

[15] For example, Bachu (1999) reports that among first-time mothers unmarried at conception, 30.5 percent of non-Hispanic whites and 10.2 percent of non-Hispanic blacks were married at the time of birth in the 1990-94 period. It is unknown whether or to what extent the probability that an unmarried woman marries during pregnancy is influenced by information on the health of the fetus. See Cooksey (1990) and Akerlof, Yellen, and Katz (1996) for analyses of the decision of pregnant women to marry before they give birth.

the remaining occurring in the post-neonatal period. As is well-established, infant mortality rates among blacks are over twice as high as those among whites, overall and neonatally and post-neonatally. Native American mortality rates are substantially higher than those of whites but much lower than those of blacks, particularly in the neonatal period. In contrast, Mexicans experience lower mortality rates than whites, both neonatally and post-neonatally. The overall prevalence of mortality among Puerto Ricans is similar to that among Native Americans, but the timing of deaths differs dramatically – deaths are more common in the post-neonatal period than in the neonatal period among Native Americans, while Puerto Rican infants are more than twice as likely to die in the neonatal period as in the post-neonatal period. Finally, Asians have the lowest mortality rates of all six groups, only slightly more than one-third those of blacks.

The three-component decomposition of IMR differences given by (2) is shown in Table 2. The upper panel of Table 2 shows results that use birth weight to measure fitness at birth, while the lower panel results use gestational age. The two sets of estimates are similar, so we focus primarily on those based on birth weight. We focus here on the decomposition of the full group differences (the columns labeled “Act.”) leaving a discussion of the decomposition of the predicted differences (the columns labeled “Pred.”) for section 6.

The overall black-white IMR gap of 7.00 (with a standard error of 0.11) is overwhelmingly accounted for by differences between blacks and whites in fitness at birth: about 88 percent (6.15 / 7.00) of the black-white gap is due to differences in birth weight. Differences in conditional post-neonatal death rates account for about 16 percent (1.13 / 7.00) of the gap, while the conditional neonatal mortality component accounts for 4 percent (-0.27 / 7.00) fewer black deaths (conditional on fitness, a black infant is slightly *less* likely to die in the neonatal period than a white infant). By contrast, almost all of Native American-white IMR gap is due to differences in post-neonatal death rates. Only about 8 percent (0.24 / 2.96) of the Native American gap is accounted for by birth weight differences, while about 83 percent (2.46 / 2.96) is accounted for by the post-neonatal component. Thus, these two groups, which have the highest IMRs, show remarkably different patterns in the composition of their gaps relative to whites. Puerto Ricans, which have the next-highest IMR, show a similar decomposition as blacks – *all* of the Puerto Rican IMR gap is accounted for by an unfavorable birth weight distribution relative to whites.

Two groups have an IMR advantage when compared to whites: Mexicans and Asians. The composition of the small Mexican advantage depends on the measure of fitness. Mexicans appear to have a slight fitness advantage over whites when fitness is measured by birth weight, but a fitness disadvantage based on gestational age. Regardless of the measure of fitness used, the Asian advantage mainly emerges during the neonatal and post-neonatal period.

6. Are Group Differences Predictable Based on Background Characteristics?

Table 1 also shows group averages of background characteristics, illustrating the potential role of these characteristics in predicting variation across the groups in IMR. For example, Asians, who have the lowest IMR, also are advantaged on many indicators of socioeconomic status. Asian mothers are more likely than whites to have at least 16 years of education, they are less likely to be teenagers, and they are more likely to be married. However, a comparison across the columns of the table also highlights some puzzles. For example, black mothers are more than twice as likely as white mothers not to have completed high school (24 percent versus 12 percent) and to be less than 20 years of age at the time of the birth (18 percent versus 8 percent), hinting that maternal education and age might play important roles in producing IMR disparities. However, Native Americans, Mexicans, and Puerto Ricans all have roughly similar maternal age distributions to those of blacks and lower educational attainment, yet all three of these groups have substantially lower IMR than do blacks; moreover, the group with the lowest educational attainment of these three, Mexicans, has by far the lowest IMR.

For the group differences in background characteristics to contribute to IMR differences, the background differences must also be predictive of our fitness and mortality outcomes of interest. Table 3 shows regressions to demonstrate how predictive the background characteristics are within the white population for three outcomes: infant death, birth weight, and whether a birth is less than 1500g. All background characteristics are indicator variables, so each coefficient can be interpreted as a marginal effect relative to being in the omitted category. By far, the plural birth indicator has the largest marginal effect, but plural birth is a rare outcome in all groups, so absolute differences in its prevalence across groups are small. Education has large effects on the outcomes as well. Combining these large effects with the large group differences implies that educational differences are likely to be a large part of the story. Age and marital status appear to be the other potentially important variables.

This cursory analysis of group characteristics suggests that the most puzzling aspects of the IMR comparisons are why the black rate is so high relative to all others and why the Mexican rate is as low as it is. The relatively high (compared to whites) Native American and Puerto Rican IMRs and relatively low Asian IMR are less surprising. We now turn to a more formal analysis of the role of background characteristics using the methods of EGH.

6.1. The Combined Role of All Background Characteristics

We assess how much of the gaps and components of gaps are predictable by differences across groups in background characteristics using the reweighting methods described in section

3.2. Briefly, we reweight the population of white infants to create counterfactual populations that have the same distributions of characteristics as the other groups, while retaining the white mapping from characteristics to outcomes. IMR gaps that exist between whites and these counterfactual populations are "predicted" by the difference in the background characteristics.[16]

In the first "Predicted" column of Table 2, the top number of 2.53 indicates that, of the 7.00 excess black infant deaths per 1000 live births, 2.53 were predictable from differences in the distribution of background characteristics between blacks and whites. Put another way, if whites maintained their mapping between characteristics and mortality but were assigned the distribution of background characteristics found among blacks, we would predict that their IMR would increase by 2.53. The remaining 4.47 (7.00 – 2.53) of the black-white IMR gap is a residual in the sense that it cannot be predicted on the basis of differences in background characteristics. Of the predicted gap, 1.10 is due to predicted differences in fitness as measured by birth weight, 0.27 is due to predicted differences in neonatal mortality, and 1.15 is due to predicted differences in post-neonatal mortality.

A comparison of the predicted gaps across the five groups uncovers several interesting findings. First, only blacks are predicted to have a substantial gap in the birth weight component. The unfavorable birth weight distribution among Puerto Ricans is essentially unpredicted – although the unfavorable Puerto Rican birth weight distribution accounts for 2.30 excess deaths per 1000 live births, only 0.37 of these 2.30 excess deaths were predicted. In total, only 44 percent (1.00 / 2.28) of the Puerto Rican IMR gap is predicted. In contrast, over two-thirds of the Native American IMR gap is predicted (2.04 / 2.96), and all of the Asian IMR gap is

[16] Appendix Table A1 shows summary statistics for the populations of blacks, Native Americans, and Puerto Ricans, along with statistics for whites reweighted to match the distributions of characteristics of these three groups. As the table shows, the reweighting procedure works well in terms of producing close matches in the distribution of characteristics.

predicted (-1.01 / -1.17). The predicted gap of Mexicans is roughly similar in magnitude to that of Native Americans and Puerto Ricans, in spite of the fact that Mexicans have much lower actual IMRs than these groups. Put another way, Mexicans, Native Americans, and Puerto Ricans have background characteristics that are associated with high IMR among whites, but only Native Americans and Puerto Ricans actually have high IMR (note that the groups are shown in decreasing order of their predicted IMR gaps from left to right in the table). Moreover, the components of the predicted gaps for Mexicans are similar to those for Native Americans; for example, the predicted post-neonatal mortality gap components are 1.14 and 1.28 for Native Americans and Mexicans, respectively. These predicted components match the actual gap components relatively well for Native Americans, but not at all for Mexicans. These findings further underscore the puzzle of low mortality rates among Mexicans.

Figure 2 shows the patterns in Table 2 graphically. For each group, the figure plots the magnitude of the actual (in blue) and predicted (in red) gaps associated with each component. The figure further highlights the differences across groups, particularly the differences in the sources of the Native American and Puerto Rican gaps and the lack of an actual gap for Mexicans. The strikingly large black birth weight component is also evident.

The top panel of Figure 3 shows cumulative excess infant deaths per 1000 live births for the three groups with the highest mortality rates, blacks, Native Americans, and Puerto Ricans, across the distribution of birth weights.[17] The rightmost endpoint of each curve measures the excess IMR for each of these groups, i.e., the IMR gap relative to whites, and the curve itself shows where in the birth weight distribution infant deaths arise. For example, the rightmost

[17] For example, the curve labeled “Blacks” is constructed by calculating for each birth weight cell the number of black infant deaths at birth weights less than or equal to that weight per 1000 total black births, and subtracting from it the corresponding number for whites.

endpoint of the curve labeled “Blacks” in the top panel has a height of 7.0, matching the overall black-white IMR gap. The height of the curve at 1000 grams is roughly 4.5, meaning that roughly 64 percent (4.5 / 7.0) of the black-white gap is accounted for by deaths among infants weighing less than 1000 grams. Similarly, nearly the entire Puerto Rican IMR gap emerges by 1000 grams. In contrast, very little of the Native American gap emerges at birth weights below 2000 grams. The comparison of the Native American and Puerto Rican curves is especially striking. Both groups have similar mortality rates, but Puerto Ricans have high IMRs almost exclusively because of very low birth weight infants, while very little of the excess deaths among Native Americans occur in the left tail of the birth weight distribution.

The bottom panel of the figure shows predicted excess infant deaths, based on white infants reweighted to have the characteristics of the other groups.[18] Comparing the right endpoints of the corresponding curves in the two panels replicates the findings from Table 2: relatively little of the full black-white gap is predictable, much of the Native American-white gap is predictable, and the Puerto Rican-white gap represents an intermediate case. The shape of the curves also reveal some interesting patterns. For example, predicted black excess deaths are concentrated in the left tail of the birth weight distribution, although not to the same extent as actual excess deaths.[19] About one-third of predicted black excess deaths occur at 1000 grams or less, and nearly half occur at 2000 grams or less, even though only about 0.5 percent and 2.6 percent of white births occur in these birth weight ranges, respectively. In contrast, only about 18 percent of predicted Native American excess deaths occur at birth weights of 2000 grams or less, and predicted excess deaths at birth weights greater than 2000 grams are larger for Native Americans

[18] The curves are constructed just as in the top panel, except that instead of, for example, actual cumulative black deaths per 1000 births, we use the cumulative distribution of deaths per 1000 births in the hypothetical population of whites with the black distribution of characteristics.

[19] The fact that the black predicted excess death curve is the highest of the three at the lowest birth weights is closely related to the Table 2 finding that blacks have the largest birth weight component among the predicted IMR gaps.

than for blacks, which is also true of actual excess deaths. The differences in the shapes of the predicted excess death curves across the groups are similar to the differences in the shapes of the actual excess death curves.

6.2. The Role of Individual Characteristics

We next turn to the roles of individual background characteristics in predicting IMR differences across groups, focusing again on comparing whites to the three groups with the highest mortality rates: blacks, Native Americans, and Puerto Ricans. We use the reweighting methods described in Section 3.3 to address the following questions: (1) Which characteristics are most responsible for the predicted gaps in IMR, in total and by component, as compared with whites? (2) Why do blacks have a larger predicted birth weight component (as seen in Figure 2) than the other groups? (3) Why do Puerto Ricans have a smaller predicted post-neonatal component than either blacks or Native Americans?

Figure 4 displays our main results graphically. The figure is broken into three panels, one for each of the high-IMR groups. Each panel contains nine sets of four bars, one set for each of the background variables that we include.[20] Each set of bars shows the excess deaths per 1000 live births as compared with whites, predicted from differences in the distribution of the variable in question. The four bars in each set show the full predicted gap and the predicted gaps based on each of the three temporal components. To illustrate, consider the set of bars labeled "edu" in the first panel of Figure 4. They show that according to our methods, if white mothers had the distribution of education of black mothers while retaining their own distribution of all other

[20] In order to conserve on space, maternal education is labeled in the figure as "edu", maternal age as "age", marital status as "mar", prenatal care in the first trimester as "pref", previous loss as "prev", whether the infant is male as "male", whether the birth was plural as "plur", birth order as "ord", and the set of indicators for the state in which the birth occurred as "state".

characteristics and their own mapping from characteristics to outcomes, there would be roughly 0.63 more deaths per 1000 live births among whites, with most of the effect resulting from a change in the distribution of birth weights.

To use Figure 4 to address our first question, it is clear that the three characteristics that most predict higher infant mortality are the same for the three groups: mother's education, age, and marital status. Of these, differences in education and marital status have the largest predicted effects. In both cases, more than half of the predicted effect comes through the birth weight component. For marital status, the post-neonatal mortality component is also important, accounting for about one-third of the predicted effect. In contrast, to the extent that differences in maternal age predict higher mortality, the effect comes almost entirely through the post-neonatal mortality component.

Turning to the second question, "why is the predicted birth weight component largest for blacks?" we note first that because Native Americans and Puerto Ricans are on average less educated than blacks (and because low education is associated with low birth weight among whites), education differences by themselves predict a larger birth weight component for Native Americans and Puerto Ricans than for blacks. However, this factor is more than offset by several others. First, black mothers are the least likely to be married, and being unmarried is associated with low birth weight among whites. Second, plural births are more common among blacks than among Native Americans and Puerto Ricans, and birth weights and survival rates are substantially lower for plural births than for singletons. Whites also are more likely to have plural births than are Native Americans and Puerto Ricans, so the predicted birth weight components associated with plurality are negative for both of these groups.

Finally, the third factor, which is especially important for Puerto Ricans, involves the states in which births occur. Births among blacks tend to occur in states in which the birth weight distribution for whites is relatively unfavorable relative to the overall birth weight distribution among whites, and this factor predicts a small amount of the black-white birth weight gap. For Native Americans, this "state of birth" effect is small and is of the opposite sign. For Puerto Ricans, there is a strong effect, also in the opposite direction: Puerto Rican births tend to occur in states in which the birth weight distribution for whites is relatively favorable. To illustrate this point, note that nearly 51 percent of Puerto Rican births occur in three states – New York, Florida and New Jersey – and the share of white infants weighing less than 1500 grams at birth in those states is roughly 8 percent lower than in all other states.[21]

Regarding the third question, "Why do Puerto Ricans have a smaller predicted post-neonatal component than either blacks or Native Americans?" the most important contributing factor is again the distribution of births across states. The states in which Puerto Rican births are concentrated have relatively low post-neonatal mortality rates conditional on birth weight. In contrast, black and Native American babies are born in states with modestly unfavorable post-neonatal mortality rates.

6.3. Other Background Factors

WORK IN PROGRESS. We are currently exploring the role of other background factors that are not easily accommodated in the framework above, including SES and immigrant status differences. In Appendix Table A3, we present evidence on specific socioeconomic measures for our six groups, taken from various external sources. Broadly speaking, blacks, Puerto Ricans,

[21] In New York, Florida, and New Jersey, the share of white babies born at less than 1500 grams is 1.06 percent, as compared to 1.15 percent in all other states. For whites, these very low birth weight babies are roughly 77 times more likely to die in the first year than babies weighing more than 1500 grams.

Mexicans, and Native Americans are similarly disadvantaged when compared to whites, although some potentially important differences exist. For example, blacks have the highest unemployment rate and lowest median income, while Hispanics and Native Americans are much more likely to be uninsured. Mexicans are most likely to have less than a high school diploma, which is consistent with the results shown in Table 1. Asians have an even higher median household income than whites, the same unemployment rate, and they are second to whites on the other measures.

7. Discussion and Conclusions

We have used micro-level Vital Statistics data from 2000 to 2004 to examine differences in infant mortality across several racial and ethnic groups in the U.S. Based on the approach of Elder, Goddeeris and Haider (2010a), we decomposed mortality disparities into three temporal components – fitness at birth, conditional neonatal mortality, and conditional post-neonatal mortality – and estimated the extent to which these components are predictable based on differences across race and ethnicity in the distribution of background characteristics.

Our analyses revealed several substantive findings. First, the timing of IMR disparities varies widely across racial and ethnic groups. The relatively high IMRs of blacks and Puerto Ricans are primarily driven by disadvantages in fitness at birth. In contrast, nearly all of the Native American IMR gap is driven by excess infant deaths during the post-neonatal period, with differences in fitness between white and Native American infants playing only a minor role. Mortality among Mexican infants is slightly lower than that among whites, with each temporal component playing a modest role. Asians have substantially lower mortality than whites, primarily because of lower neonatal and post-neonatal mortality rates conditional on fitness.

Second, the relationship between background characteristics and group IMR disparities varies widely across groups. For example, Puerto Ricans' substantial disadvantage in birth weight relative to whites is essentially unpredicted by the full set of background characteristics–although Puerto Ricans are disadvantaged on several measures of socioeconomic status, these disadvantages do not predict the large Puerto Rican-white gap in birth weight when other factors, including the geographic distribution of births, are taken into account. In contrast, much of the Native American IMR disadvantage and the entire Asian IMR advantage are predicted based on background characteristics. The predicted gap of Mexicans is similar in magnitude to those of Native Americans and Puerto Ricans, so it is remarkable that the actual Mexican IMR is lower than that of whites.

Despite the large differences in the composition and predictability of IMR gaps, some consistent patterns emerge across groups. First, differences in neonatal mortality conditional on fitness at birth do not play a large role in any of the IMR gaps. Second, fitness at birth varies widely across groups, but the large fitness disadvantages of blacks and Puerto Ricans are not well-predicted based on background characteristics. As a result, most of the variation across race and ethnicity in the number of deaths occurring at birth weights below 1000 grams remains a mystery. Third, in contrast to the unpredictability of fitness gaps, variation across groups in conditional post-neonatal mortality is closely related to background characteristics.

Finally, we have isolated the roles that individual observable characteristics play in the predicted gaps. Among the three groups that are most disadvantaged relative to whites – blacks, Native Americans, and Puerto Ricans – relatively low marriage rates and maternal education are most responsible for producing high predicted IMRs, with an unfavorable maternal age distribution playing a smaller but still significant role. The combined effect of all other background characteristics, including measures of prenatal care, plurality, birth order, and the state in which the birth occurred, is relatively small in all cases. As a concrete example, if white mothers had the same distributions of educational attainment, age at birth, and marital status as do Native American mothers, our estimates predict that nearly two-thirds of the Native American-white IMR gap would disappear. Differences between whites and Native Americans in all other characteristics are predicted to account for less than 5 percent of the gap in total.

By shedding light on the sources and temporal patterns of several racial and ethnic IMR gaps, these results suggest where future research and policy efforts would have the most impact. For example, because differences in conditional neonatal mortality rates play no significant role in any of the racial or ethnic IMR gaps, disparities in access to neonatal medical care probably have very little effect on the magnitudes of existing disparities. In contrast, we find that conditional post-neonatal death rates, which are closely related to observable characteristics, are important components of the relatively high IMRs of blacks and Native Americans (and the relatively low IMR of Asians). These patterns suggest that efforts to reduce post-neonatal mortality among disadvantaged populations may be an effective means of reducing IMR disparities.

References

Akerlof GA, Yellen JL, Katz ML (1996). "An Analysis of Out-of-Wedlock Childbearing in the United States." Quarterly Journal of Economics 111(2):277-317.

Alexander G. R., M. S. Wingate, D. Bader and M. D. Kogan (2008). "The Increasing Racial Disparity in Infant Mortality Rates: Composition and Contributors to Recent US Trends." American Journal of Obstetrics and Gynecology; 198:51.e1-51.e9.

Bachu, Amara (1999). Trends in Premarital Childbearing: 1930 to 1994. Current Population Reports P23-197. Washington, DC: U.S. Census Bureau.

Busso, M., J. DiNardo, and J. McCrary (2009). "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators." IZA Discussion Paper No. 3998.

Carmichael, S. L. and S. Iyasu (1998). "Changes in the black-white infant mortality gap from 1983 to 1991 in the United States." Am J Prev Med 15(3): 220-7.

Cooksey, Elizabeth (1990). "Factors in the Resolution of Adolescent Premarital Pregnancies." Demography 27(2): 207-18.

DiNardo, J., N. M. Fortin and T. Lemieux, (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." Econometrica 64(5): 1001-44.

Eberstein IW, Nam CB, Hummer RA. (1990) "Infant Mortality by Cause of Death: Main and Interaction Effects." Demography 27(3):413-430.

Elder, Todd E., John H. Goddeeris and Steven J. Haider (2010a). "A Deadly Disparity: A Comprehensive Assessment of the Black-White Infant Mortality Gap." Working Paper, Michigan State University.

Elder, Todd E., John H. Goddeeris and Steven J. Haider (2010b). "Isolating the Role of Individual Covariates in Reweighting Estimation." Working Paper, Michigan State University.

Frisbie, W. P., S.-E. Song, et al. (2004). "The Increasing Racial Disparity in Infant Mortality: Respiratory Distress Syndrome and Other Causes." Demography 41(4): 773-800.

Hirano, K., G. W. Imbens, and G. Ridder (2003). "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." Econometrica 71(4): 1161-1189.

Hummer, R. A., M. Biegler, et al. (1999). "Race/Ethnicity, Nativity, and Infant Mortality in the United States." Social Forces 77(3): 1083-118.

Lee, K. S., N. Paneth, et al. (1980). "Neonatal mortality: an analysis of the recent improvement in the United States." Am J Public Health 70(1): 15-21.

Martin JA, Hamilton BE, Sutton PD, et al. (2009). Births: Final data for 2006. National vital statistics reports; vol 57, no 7. Hyattsville, MD: National Center for Health Statistics.

Mathews TJ, MacDorman MF (2010). Infant mortality statistics from the 2006 period linked birth/infant death data set. National vital statistics reports; vol 58 no 17. Hyattsville, MD: National Center for Health Statistics.

National Center for Health Statistics (2005). Technical Appendix from Vital Statistics of the United States 2003, Natality. Hyattsville, MD.

National Center for Health Statistics (2010). Health, United States, 2009: With Special Feature on Medical Technology. Hyattsville, MD. 2010.

Schempf, A. H., A. M. Branum, et al. (2007). "The contribution of preterm birth to the Black-White infant mortality gap, 1990 and 2000." Am J Public Health 97(7): 1255-60.

Tomashek KM, Qin C, Hsia J, Lyasu S, Barfield WD, Flowers LM. (2006) "Infant mortality trends and differences between American Indian/Alaska native infants and white infants in the United States, 1989-1991 and 1998-2000." Am J Public Health 96(12):2222-27.

Watson T. (2006) "Public health investments and the infant mortality gap: Evidence from federal sanitation interventions on US Indian reservations." Journal of Public Economics 90(8-9):1537-60.

Wise, P. H. (2003). "The Anatomy of a Disparity in Infant Mortality." Annual Review of Public Health 24:341–62.

Wooldridge, J. (2007). "Inverse probability weighted estimation for general missing data problems." Journal of Econometrics 141(2): 1281-130.

Figure 1: Actual and Predicted IMR Gaps by Race/Ethnic Group

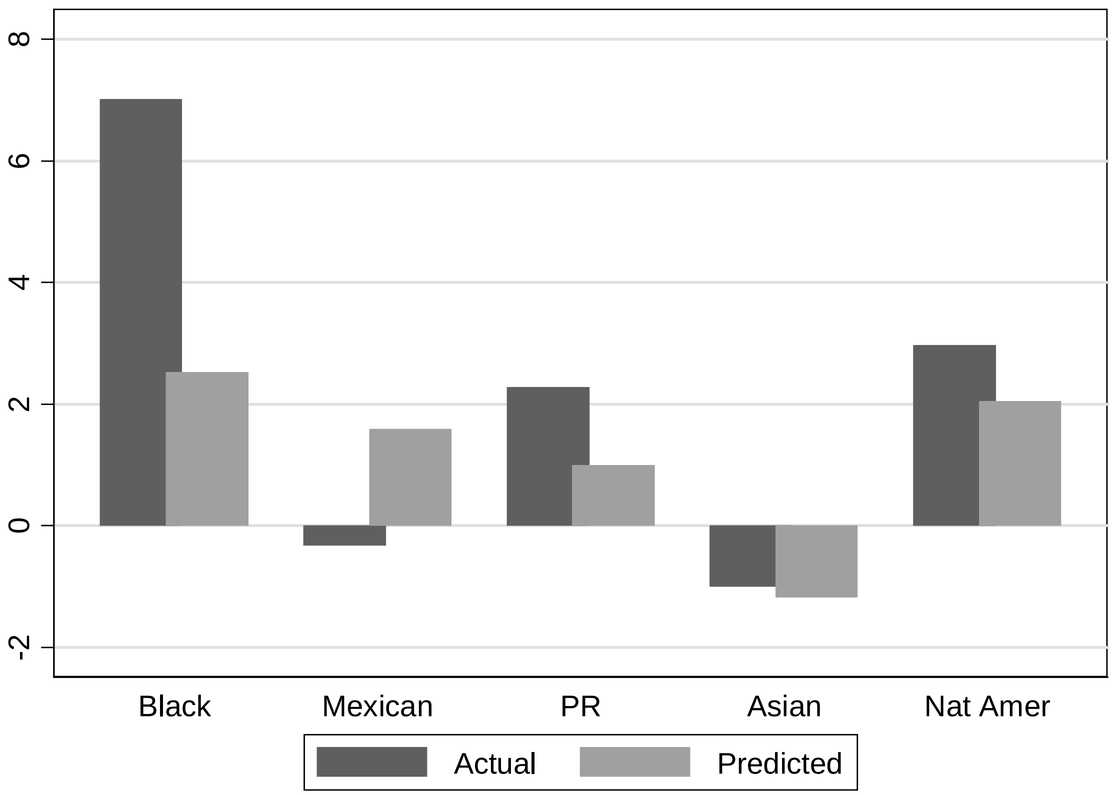


Figure 2: Actual and Predicted IMR Gaps Relative to White IMR

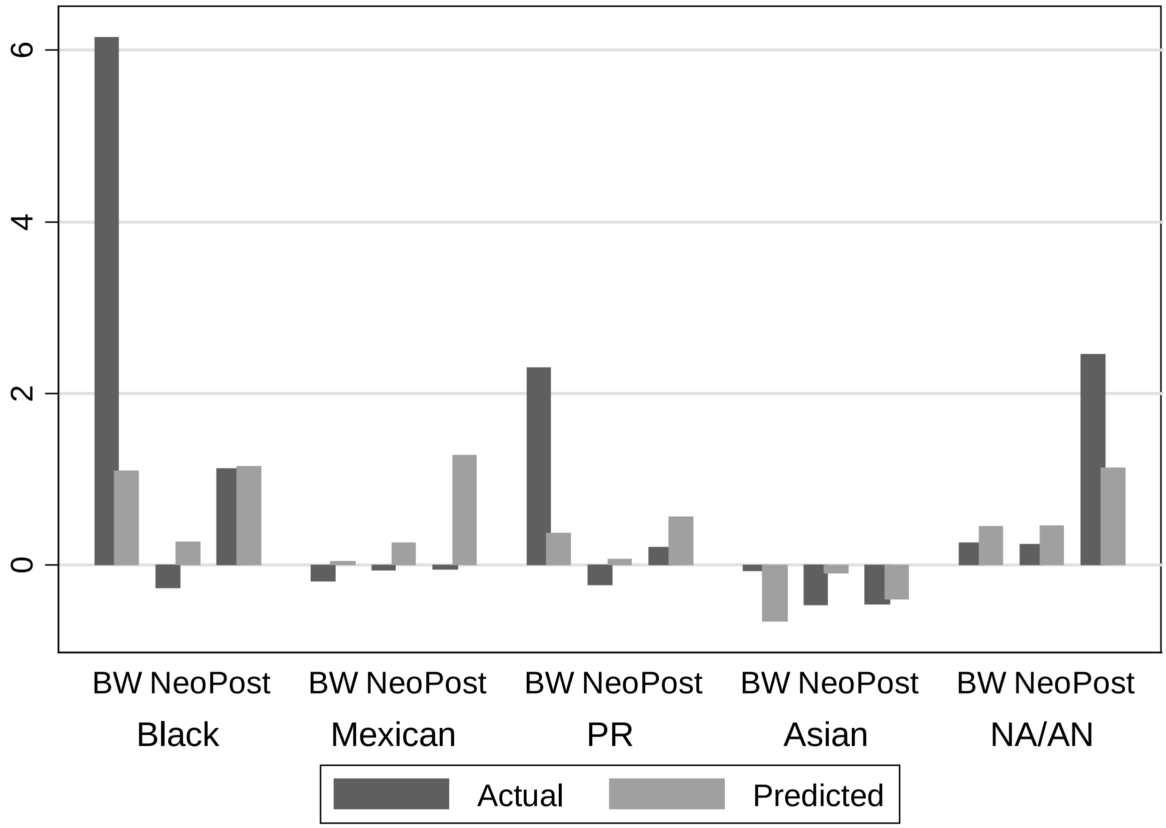


Figure 3: Actual and Predicted Excess Infant Deaths Relative to Whites

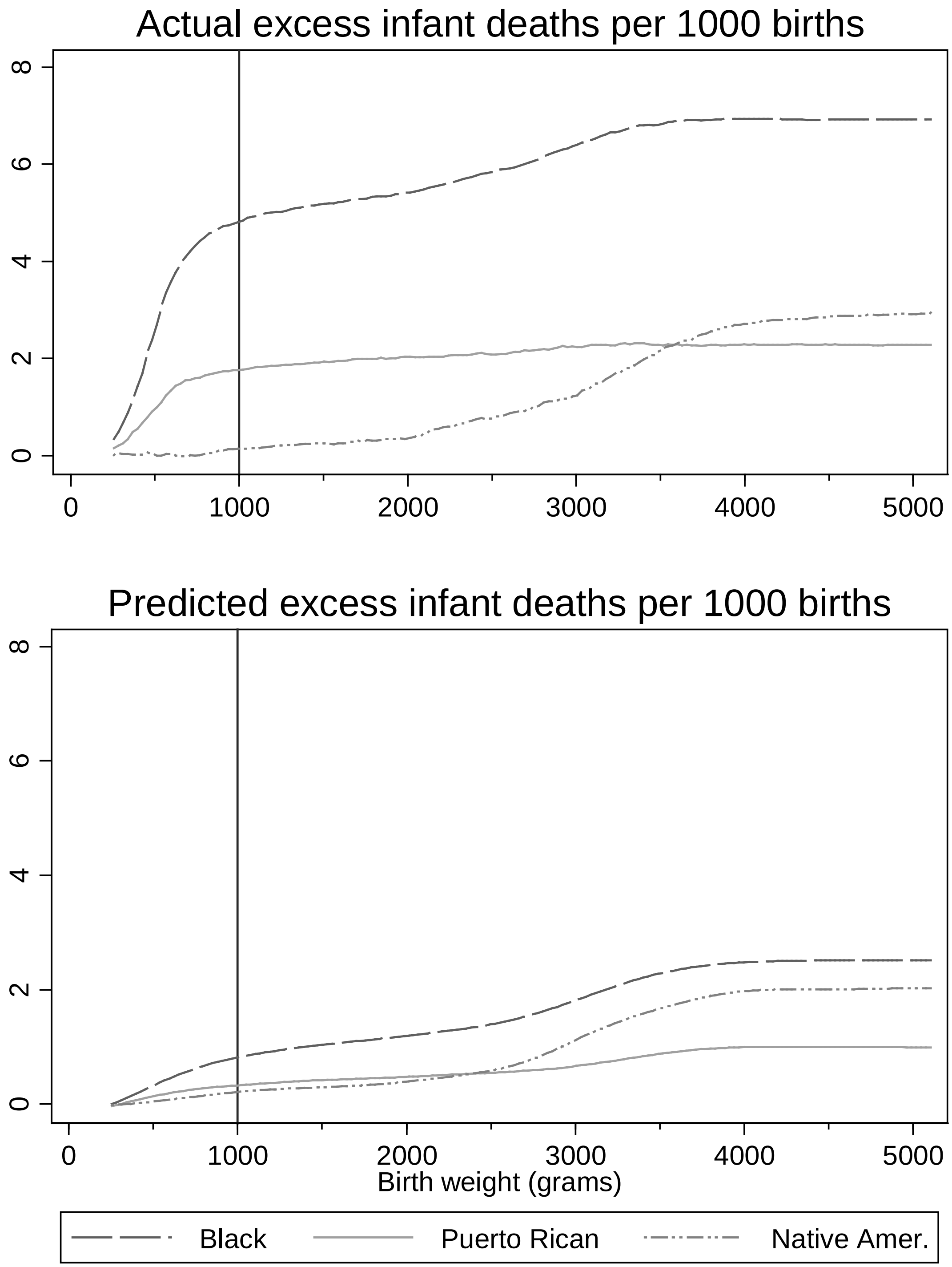
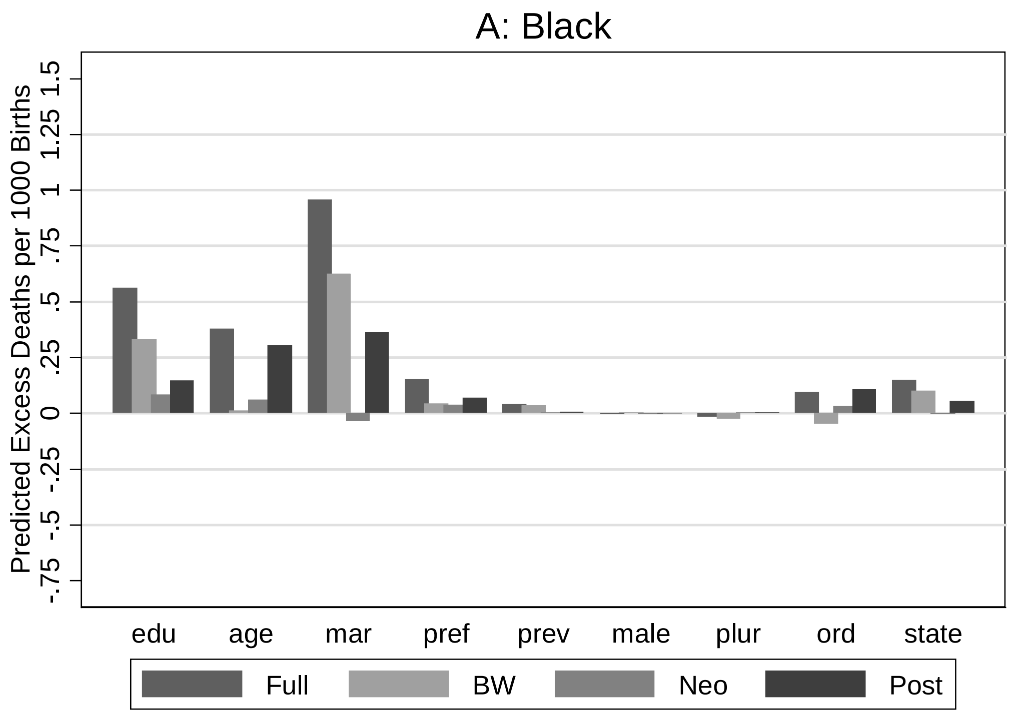
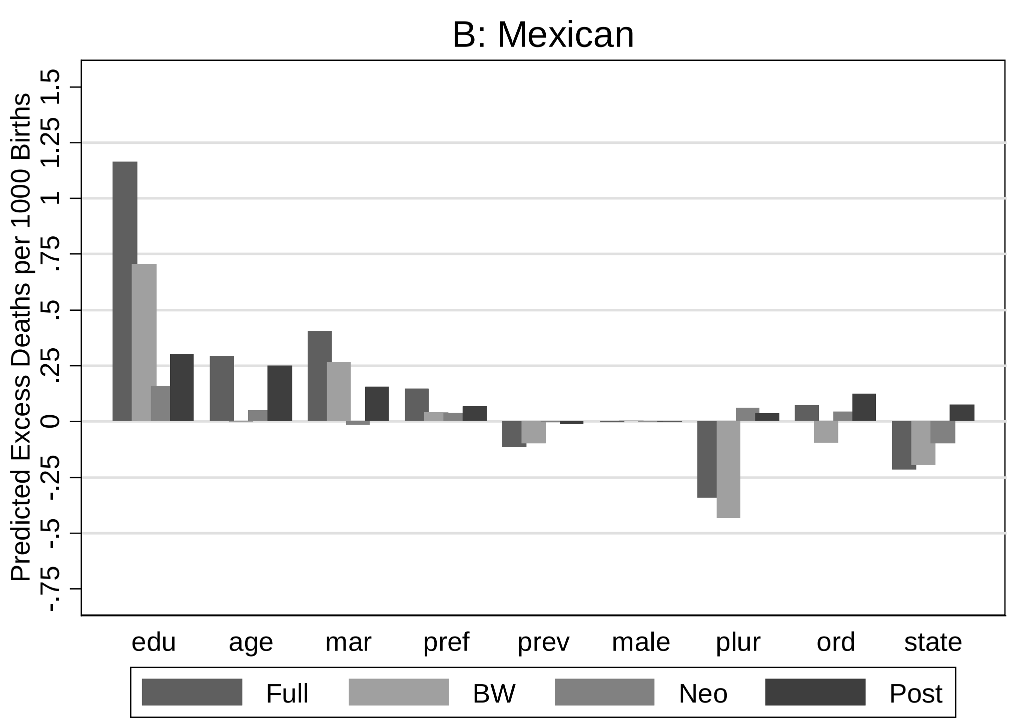
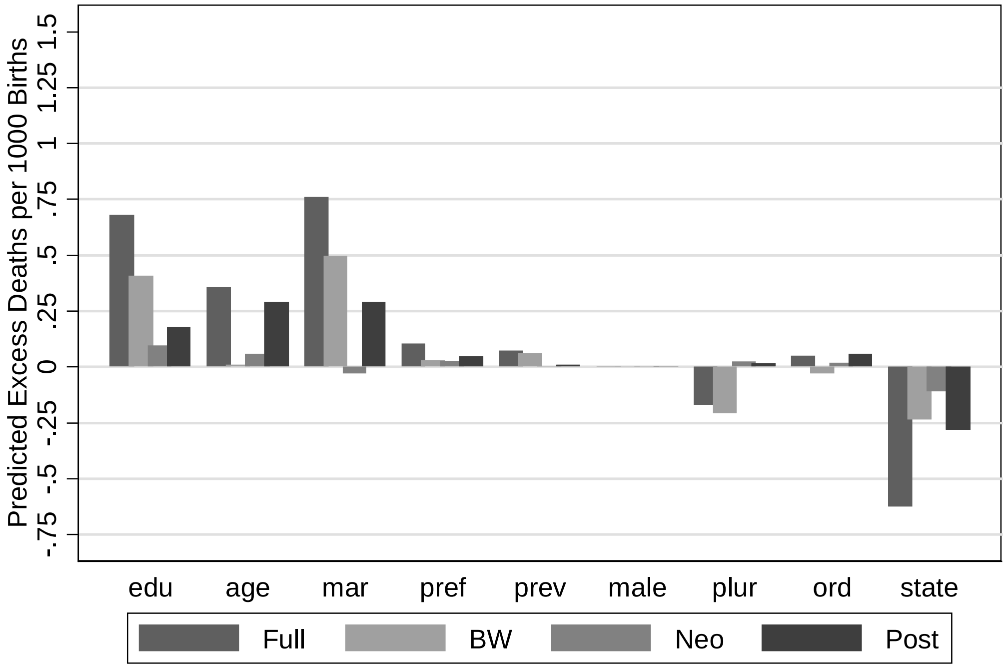


Figure 4: Predicted Excess Deaths by Background Characteristic Relative to Whites

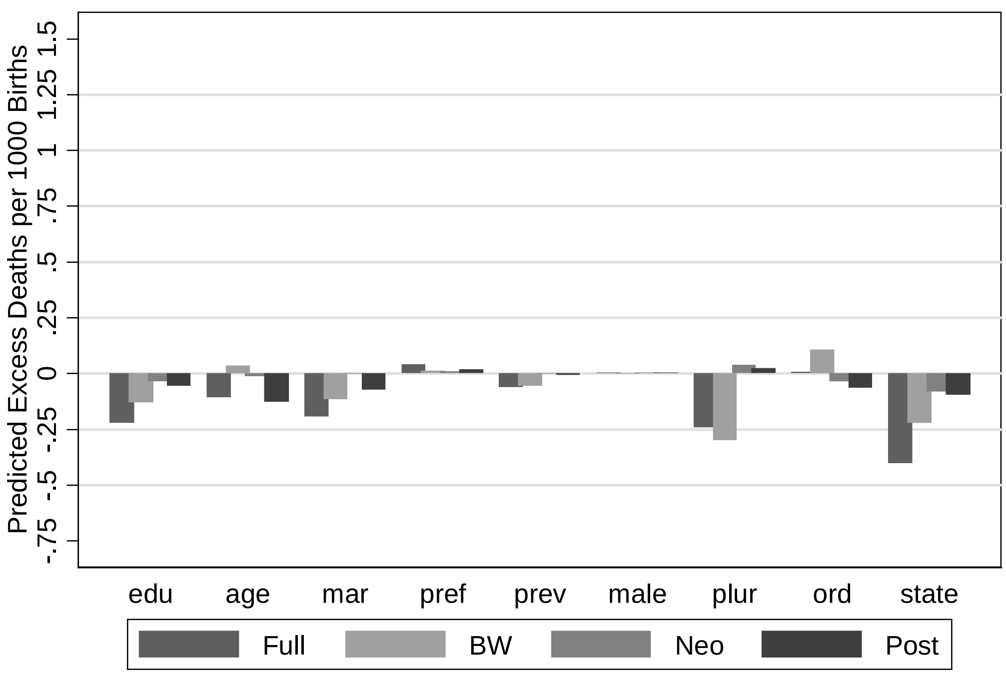




C: Puerto Rican



D: Asian



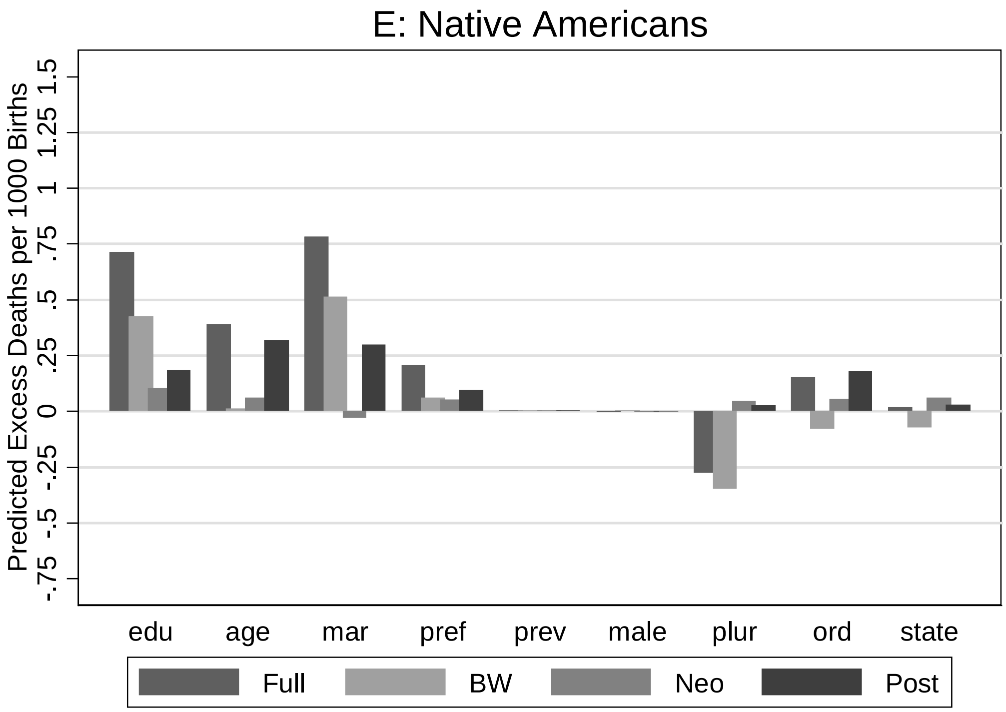

E: Native Americans
Predicted Excess Deaths per 1000 Births
-.75
-.5
-.25
0
.25
.5
.75
1
1.25
1.5
edu
age
mar
pref
prev
male
plur
ord
state
Full
BW
Neo
Post

Table 1: Mortality, Fitness and Background Characteristics by Group

	Whites	Blacks	Mexicans	Puerto Ricans	Asians	Native Amer.
Observations	2,253,861	555,376	604,207	277,401	684,234	184,366
Mortality rates						
Infant	5.35	12.35	5.03	7.63	4.23	8.31
Neonatal	3.51	8.12	3.28	5.24	2.92	3.95
Post-neonatal	1.84	4.23	1.75	2.41	1.43	4.36
Fitness measures						
Birth weight (g)	3356	3099	3323	3216	3214	3351
Missing BW	<.001	<.001	<.001	<.001	<.001	<.001
Gestational age (w)	38.8	38.2	38.8	38.6	38.8	38.7
Missing GA	.004	.005	.022	.003	.014	.006
Background info.						
Maternal ed. (years)						
<12	0.12	0.24	0.54	0.32	0.10	0.30
12	0.30	0.39	0.29	0.34	0.23	0.40
13-15	0.24	0.24	0.11	0.23	0.21	0.22
16 +	0.34	0.13	0.06	0.12	0.47	0.08
Maternal age						
<20	0.08	0.18	0.16	0.18	0.04	0.18
20-24	0.22	0.33	0.31	0.32	0.13	0.34
25-29	0.27	0.23	0.27	0.24	0.29	0.24
30-34	0.27	0.16	0.17	0.17	0.34	0.15
35 +	0.17	0.10	0.09	0.10	0.20	0.09
Mother married	0.77	0.31	0.58	0.41	0.86	0.40
First trimester care	0.88	0.75	0.75	0.79	0.85	0.70
Previous loss	0.25	0.29	0.16	0.31	0.21	0.26
Male	0.51	0.51	0.51	0.51	0.52	0.51
Plural birth	0.036	0.036	0.020	0.029	0.025	0.024
Live birth order						
1^{st}	0.41	0.38	0.35	0.39	0.47	0.35
2^{nd} - 3^{rd}	0.50	0.47	0.50	0.49	0.47	0.46
4^{th} +	0.09	0.15	0.15	0.12	0.06	0.19
Census region						
Northeast	0.18	0.16	0.03	0.59	0.21	0.04
Midwest	0.28	0.19	0.11	0.10	0.13	0.20
South	0.35	0.57	0.35	0.24	0.21	0.27
West	0.19	0.08	0.52	0.07	0.46	0.49

Notes:

Table 2: Actual and Predicted IMR Gaps with Whites, 2000-2004

	Black		Mexican		Puerto Rican		Asian		Native Amer.	
Gap Type	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
Overall	7.00	2.53	-0.32	1.59	2.28	1.00	-1.01	-1.17	2.96	2.04
	(.11)	(.10)	(.11)	(.17)	(.17)	(.14)	(.11)	(.10)	(.21)	(.20)
BW										
Fitness	6.15	1.10	-0.19	0.04	2.30	0.37	-0.08	-0.66	0.26	0.45
	(.10)	(.06)	(.05)	(.07)	(.12)	(.07)	(.06)	(.06)	(.09)	(.10)
Neonatal	-0.27	0.27	-0.07	0.26	-0.23	0.07	-0.47	-0.11	0.24	0.46
	(.04)	(.05)	(.06)	(.10)	(.08)	(.08)	(.07)	(.06)	(.12)	(.15)
Post-neonatal	1.13	1.15	-0.06	1.28	0.21	0.56	-0.46	-0.40	2.46	1.14
	(.07)	(.06)	(.05)	(.12)	(.10)	(.08)	(.08)	(.07)	(.14)	(.13)
GA										
Fitness	5.60	1.08	0.21	0.22	2.17	0.42	-0.24	-0.53	0.82	0.58
	(.09)	(.11)	(.05)	(.07)	(.11)	(.11)	(.06)	(.07)	(.09)	(.11)
Neonatal	-0.08	0.23	-0.40	0.07	-0.19	-0.01	-0.39	-0.19	-0.16	0.31
	(.05)	(.06)	(.06)	(.10)	(.12)	(.12)	(.08)	(.09)	(.11)	(.14)
Post-neonatal	1.48	1.22	-0.13	1.30	0.30	0.58	-0.38	-0.46	2.31	1.15
	(.08)	(.07)	(.06)	(.11)	(.12)	(.12)	(.08)	(.08)	(.14)	(.14)

Notes: These three-component decomposition follows equation (2). “Predicted” refers to gaps or components of gaps that result when whites are compared with the white population weighted to have the background characteristics of the relevant group. Standard errors (in parentheses) are calculated from 50 bootstrapped replications.

Table 3: Infant Death and Birth Weight OLS Regressions for Whites

	Infant death (x1000)	Birth weight (grams)	Low birth weight (<1500g, x1000)
Maternal ed. (<12 excluded)			
12	-2.01	82.7	-2.2
	(0.18)	(1.3)	(0.3)
13-15	-3.24	129.3	-4.0
	(0.20)	(1.5)	(0.3)
16 +	-4.01	162.1	-6.4
	(0.20)	(1.5)	(0.3)
Maternal age (<20 excluded)			
20-24	-0.72	-20.7	-0.2
	(0.22)	(1.6)	(0.3)
25-29	-1.05	-14.6	1.0
	(0.24)	(1.8)	(0.3)
30-34	-1.25	-11.0	1.9
	(0.25)	(1.9)	(0.4)
35 +	-0.79	-32.2	3.5
	(0.26)	(2.0)	(0.4)
Mother married	-1.92	72.1	-3.5
	(0.14)	(1.0)	(0.2)
First trimester prenatal care	-1.66	27.8	-0.8
	(0.16)	(1.2)	(0.2)
Previous loss	1.25	-22.3	2.9
	(0.11)	(0.9)	(0.2)
Male	1.14	118.2	-0.1
	(0.10)	(0.7)	(0.1)
Plural birth	20.42	-1082.5	95.6
	(0.26)	(2.0)	(0.4)
Live birth order (1 excluded)			
$2^{nd} – 3^{rd}$	-0.46	93.0	-6.1
	(0.11)	(0.8)	(0.2)
4^{th} +	0.58	116.3	-6.4
	(0.20)	(1.8)	(0.3)
State indicators	Yes	Yes	Yes
Dependent variable mean	5.35	3356	11.3
R^2	.004	.147	0.030
Observations	2,253,861	2,253,679	2,253,679

Notes:

Table A1: Mortality and Background Characteristics for Reweighted Whites

		Whites Reweighted to Look Like:				
	Whites	Blacks	Mexicans	Puerto Ricans	Asians	Native Amer.
Mortality rates						
Infant	5.35	7.91	6.94	6.35	4.18	7.40
Neonatal	3.51	4.66	3.70	3.88	2.84	4.30
Post-neonatal	1.84	3.25	3.23	2.47	1.34	3.09
Background info.						
Maternal ed. (years)						
<12	.12	.25	.55	.30	.09	.30
12	.30	.39	.28	.34	.22	.39
13-15	.24	.24	.11	.24	.20	.22
16 +	.34	.13	.06	.12	.49	.09
Maternal age						
<20	.08	.18	.18	.17	.03	.18
20-24	.22	.32	.35	.32	.13	.35
25-29	.27	.23	.26	.24	.29	.24
30-34	.27	.16	.15	.17	.34	.14
35 +	.17	.11	.07	.10	.21	.08
Mother married	.77	.32	.52	.43	.87	.41
First trimester care	.88	.74	.71	.79	.85	.70
Previous loss	.25	.30	.17	.31	.20	.26
Male	.51	.51	.51	.51	.52	.51
Plural birth	.036	.034	.019	.029	.025	.024
Live birth order						
1st	.41	.38	.36	.40	.48	.36
2^{nd} / 3^{rd}	.50	.47	.49	.48	.46	.46
4^{th} +	.09	.15	.15	.11	.06	.18
Census region*						
Northeast	.18	.15	.02	.54	.20	.04
Midwest	.28	.19	.10	.10	.13	.19
South	.35	.58	.38	.28	.20	.30
West	.19	.08	.50	.08	.47	.47

Notes: *

Table A2: Estimates and Standard Errors for Role of Covariates Graphed in Figure 4

	Edu	Age	Mar	Prenat.	Prev.	Male	Plur	Order	State
<u>Black</u>									
Full	.56	.38	.96	.15	.04	-.01	-.02	.09	.15
Birth Weight	.33	.01	.62	.04	.04	.00	-.02	-.04	.10
Neonatal	.08	.06	-.03	.04	.01	-.00	.00	.03	-.00
Post-neonatal	.15	.31	.36	.07	.01	-.00	.00	.11	.05
<u>Mexican</u>									
Full	1.16	.29	.40	.15	-.12	-.00	-.34	.07	-.22
Birth Weight	.70	-.01	.26	.04	-.10	.00	-.44	-.09	-.19
Neonatal	.16	.05	-.02	.04	-.00	-.00	.06	.04	-.10
Post-neonatal	.30	.25	.16	.07	-.01	-.00	.04	.12	.07
<u>Puerto Rican</u>									
Full	.68	.36	.76	.10	.07	-.00	-.17	.05	-.62
Birth Weight	.41	.01	.50	.03	.06	.01	-.21	-.03	-.24
Neonatal	.09	.06	-.03	.03	.00	-.00	.03	.02	-.11
Post-neonatal	.18	.29	.29	.05	.01	-.00	.02	.06	-.28
<u>Asian</u>									
Full	-.22	-.10	-.19	.04	-.06	.00	-.24	.01	-.40
Birth Weight	-.13	.04	-.11	.01	-.05	-.00	-.30	.10	-.22
Neonatal	-.03	-.01	-.00	.01	-.00	.00	.04	-.03	-.08
Post-neonatal	-.06	-.13	-.07	.02	-.01	.00	.02	-.07	-.09
<u>Native Amer.</u>									
Full	.71	.39	.78	.21	.00	-.00	-.28	.15	.02
Birth Weight	.42	.01	.51	.06	.00	.00	-.35	-.08	-.07
Neonatal	.10	.06	-.02	.05	.00	-.00	.05	.06	.06
Post-neonatal	.18	.32	.30	.09	.00	-.00	.03	.18	.03

Notes:

Table A3: Select Population Attributes by Race/Ethnicity

	U.S.	White	Black	Puerto Rican	Mexican	Native American	Asian
Less than HS diploma among 25+ (%) *	16.0	11.1	20.6	28.4	46.5	21.0	14.6
Unemployment among 16+ (%) *	4.2	3.4	8.0	6.3	5.2	7.3	3.4
Median household income ($) *	50,007	54,189	33,546	37,152	38,823	37,229	65,429
Poverty rate (%) *	9.8	6.1	21.7	21.8	20.8	18.6	8.4
Not in labor force among 20-64 (%) **	24.1	22.1	30.9	32.6	32.6	31.8	28.3
Uninsured (%) ***	15.4	10.6	19.6	32.8	32.8	32.1	16.5

Notes:

* Data for these rows were obtained from the U.S. Census Bureau's American FactFinder, "Selected Population Profile in the United States (S0201)", based on the 2005-2007 American Community Surveys. The white column is for individuals who report being white alone and not Hispanic or Latino. The black column is for individuals who report black or African American alone or in combination with one or more races. The Native American column is for individuals who report being American Indian or Alaska Native alone or in combination with one or more other races.

** Data for this row was obtained from the Census 2000 Brief "Employment Status: 2000" (C2KBR-18), Table 2, based on the 2000 Census. Puerto Ricans and Mexicans are not listed separately in this publication; the rate given in the table is that listed for "Hispanic or Latino (of any race)".

*** Data for this row was obtained from the Current Population Reports "Income, Poverty, and Health Insurance Coverage in the United States: 2007" (P60 report), Table 7, based on the 2005-2007 Current Population Surveys. Puerto Ricans and Mexicans are not listed separately in this publication; the rate given in the table is that listed for "Hispanics".

Figure A1: Infant Deaths per 1000 Live Births for Various Racial and Ethnic Groups by Year, 2000-2004