Education, birth weight, and infant mortality

Timothy B. Gage[1,2], Fu Fang[1], Erin O'Neill[1], Howard Stratton[2]

1. Department of Anthropology, University at Albany, State University of New York.
2. Department of Epidemiology and Biostatistics, School of Public Health, University at Albany, State University of New York

Corresponding Author: Timothy B. Gage
Department of Anthropology
AS 114
University at Albany – SUNY
Albany, NY 12222, USA
Telephone: 518-442-4704
Fax: 518-442-5710
Email: tbg97@albany.edu

This project is supported by NIEHS R01 HD037405

ABSTRACT

This research determines if the declines in infant mortality with increased education are due to “indirect” effects that operate through improved birth weight and potentially causal or to “direct” effects, independent of birth weight. The data used are the 2001 US national linked birth-death files, African Mexican, and European American cohorts by sex. Education is dichotomized as low (i.e. high school and less) versus high (i.e. college and above) education. The analysis is conducted using Covariate Density Defined mixture of logistic regression, which explores “normal” and “compromised” births separately. Among “normal” births, mean birth weight increases significantly with high education, 27 to 108 grams. Mortality declines significantly due to “direct” effects of high education (by a factor of 0.4-to 0.96). Only 3 (of six possible) “normal” “indirect” effects are significant, and all increase infant mortality with higher education despite improvements in birth weight! European American births display the largest and Mexican American the smallest effects of education on birth outcomes. Among “compromised” births education has small and inconsistent effects. These results are consistent with the view that interventions targeting birth weight may not result in lower infant mortality rates.

INTRODUCTION

Measures of socioeconomic status, such as maternal education, are thought to be important correlates of body size and mortality (Floud, Wachter and Gregory 1990; Waaler 1984). What is not clear is whether socioeconomic status influences mortality "directly", independent of body size, or "indirectly" through its influence on the anthropometric measures. If the effects are "indirect", then body size may be a part of the causal mechanism. If the effects are "direct", then variation in body size is not causal. One well-established example is the relationship of socioeconomic level with birth weight and infant mortality. The correlation between birth weight and infant mortality is empirically so well documented that the current US policy to reduce infant mortality is to improve birth weight outcomes (Buehler et al. 1987; Institute-of-Medicine 1985; Mc Cormick 1985; US-DOHHS 2000). However, many theoreticians have argued that, at least among "normal" births, birth weight is not on the causal pathway to infant mortality (Mosely and Chen 1984; Wilcox and Russell 1990; Wise 2003). Recently a statistical method of determining if birth weight is or is not on the "causal" pathway to infant mortality has been developed (Gage et al. 2004), based on the Wilcox-Russell definition of "causality" (Wilcox and Russell 1990). The aim of this research is to determine if maternal education, a measure of socioeconomic level, affects infant mortality "indirectly" though birth weight, "directly", or both. The analyses are conducted on three populations by sex, African, European, and Mexican Americans from the 2001 US birth cohort.

BACKGROUND

Wilcox and his colleagues argue (Wilcox and Russell 1990) that among "normal" births when the birth weight distribution shifts right or left due to a stressor (e.g. smoking during pregnancy; fetal development at altitude) that the birth weight specific mortality curve shifts in

the same direction a similar amount so that there is no net change in infant mortality due to the shift in the birth weight distribution (Figure 1a). If so, birth weight cannot be on the "causal" pathway to infant mortality. In addition, they argue that there may be a consistent increase (or decrease) in infant mortality at all birth weights (independent of birth weight) due to a stressor (e.g., smoking during pregnancy; but not fetal development at altitude) (Figure 1b). Based on this definition it only need be shown that the shift in the birth weight distribution among "normal" births is not matched by a similar shift in the birth weight specific infant mortality curve to demonstrate that birth weight could be on the "causal" pathway to infant mortality (Figure 1c). This will be referred to here as an "indirect" effect of the stressor through birth weight on infant mortality. The birth weight independent effect will be called a "direct" effect of the stressor. The Wilcox-Russell hypothesis does not discuss whether "compromised" births behave in the same manner as "normal" births. Nevertheless, to fully examine the Wilcox-Russell hypothesis requires a method that distinguishes between "normal" and "compromised" births. Covariate Density Defined mixture of logistic regressions (CDDmlr) distinguishes between "normal" and "compromised" births and can estimate the significance of "indirect" as well as "direct" effects.

Figure 1 about here

The analysis presented below uses Covariate Density Defined mixture of logistic regressions (Gage et al. 2004). Conceptually, this method fits a multi-component Gaussian finite mixture model to the birth weight distribution, which divides the population into several (two in this case) latent subpopulations, and simultaneously fits a separate logistic regression on infant mortality by birth weight to each latent subpopulation. This procedure is useful when the latent subpopulations account for otherwise unobserved heterogeneity with respect to mortality. For the purposes of identification the subpopulation accounting for the majority of individuals is labeled

the primary (p) subpopulation and the remaining minority component is labeled the secondary (s) subpopulation. The logistic regressions on infant mortality are parameterized as second-degree polynomials of birth weight to account for the reverse-J shaped relationship of birth weight and infant mortality. Previous applications indicate that a) the primary subpopulation represents births undergoing "normal" fetal development given its location in the center of the birth weight distribution; b) the secondary subpopulation accounts for most low birth weight and macrosomic births and is consequently interpreted as accounting for births undergoing "compromised" fetal development (Figure 2a); and c) the two components identified by the finite Gaussian mixture model are heterogeneous with respect to infant mortality with the "compromised" subpopulation consistently displaying lower birth weight specific mortality but higher overall mortality due to the less favorable distribution of birth weight (Figure 2b) (Gage et al. 2004).

Figure 2 about here

Here we expand this basic model by adding exogenous covariates to the density submodel and to the logistic regressions. In the case presented here the covariate is education: low education versus high education. In the density submodel, all parameters, i.e. the mixing proportion, means and variances are each defined as a function of education. In addition, education is added to the logistic regressions, as a covariate on the constant, linear and squared terms of a second-degree polynomial of birth weight. Finally we define birth weight within each logistic regression as the Z-score based upon the Gaussian subpopulation that it represents. The result is a model that can test the Wilcox-Russell definition (Wilcox and Russell 1990) of "direct" and "indirect" effects (Figure 1). After standard transformation of the quadratic covariates to account for correlation inherent in this specification, the education interaction

covariate on the constant of the birth weight polynomial can be interpreted as a "direct" effect (constant at all birth weights). The education interaction covariate on the linear birth weight term accounts for the shift, left or right, in the birth weight specific mortality curve with respect to mean birth weight. Due to the use of Z-scores of birth weight, an insignificant interaction on the linear birth weight interaction term indicates that the birth weight specific mortality curve shifts in concert with birth weight density, i.e. no "indirect" effect of education as argued by the Wilcox-Russell hypothesis (Wilcox and Russell 1990). The education interaction on the squared term accounts for changes in shape of the birth weight specific infant mortality curve relative to the Z-scored birth weight distribution. This is not considered in the Wilcox-Russell theory, but represents a second way that birth weight specific infant mortality could be uncoupled from the Z-scored birth weight density. Again an insignificant interaction term indicates no "indirect" effect, due to the use of Z-scored birth weight. Here we examine shift and shape effects together, as potential "causal" effects, since they both represent an uncoupling of the birth weight density and birth weight specific mortality. A formal definition of this model is provided in the methods.

DATA AND METHODS

Data Source

The data for these analyses were obtained from the 2001 national linked birth-death files. Race and ethnic origin is based on mother's reported race and ethnic origin. Births with missing information or gestational age <20 weeks or birth weight <500 grams are excluded. Summary statistics for the birth cohorts of interest are presented in Table 1.

Table 1 about here

Statistical Model

Formally the model is defined as follows. The probability of death (y) is a product of a) the density distribution of the birth weight (x) given the exogenous dichotomous covariate (z), and b) the conditional mortality given x and z :

$$f(y,x \mid z;\beta,\theta) = f_2(y \mid x,z;\beta,\theta) \cdot f_1(x \mid z;\theta) \quad \text{(Eq. 1)}$$

In the case of two Gaussian subpopulations (labeled as p and s), $f_1(x|z;\theta)$ is expressed as:

$$f_1(x \mid z;\theta = (\pi_s, \mu_s, \sigma_s, \mu_p, \sigma_p))$$

$$= \pi_s(z) \cdot \tilde{N}(x;\mu_s(z),\sigma_s(z)) + [1-\pi_s(z)] \cdot \tilde{N}(x;\mu_p(z),\sigma_p(z)) \quad \text{(Eq. 2)}$$

$$\pi_s(z) = logit(\eta_s(z)) = logit(\alpha_0 + z \cdot \alpha_1) \quad \text{(Eq. 3)}$$

$$\mu_i(z) = \gamma_{i,0} + z \cdot \gamma_{i,1} \quad \text{(Eq. 4)}$$

$$\sigma_i(z) = \lambda_{i,0} + z \cdot \lambda_{i,1} \quad \text{(Eq. 5)}$$

where π_s is the proportion of births belonging to the subpopulation s. For $i = p$ and s, $\tilde{N}$ represents the Gaussian density, truncated at 500 grams, with mean μ_i and variance σ_i. And the probability of death conditioned on x and z is given by:

$$f_2(y = 1 \mid x,z;\beta = (\beta_s, \beta_p),\theta)$$

$$= q_s(x \mid z,\theta) \cdot P_s(x_i^* \mid z,\beta_s) + [1-q_s(x \mid z,\theta)] \cdot P_p(x_i^* \mid z,\beta_p) \quad \text{(Eq. 6)}$$

where x_i^* is the standardized x using the respective μ_i and σ_i, P_i is the probability of death for an infant with birth weight (x) and covariate z in subpopulation i given by a quadratic logistic form:

$$P_i(x_i^* \mid z,\beta_i = (A_i, B_i, C_i)) = \frac{e^{A_i(z) + x_i^* \cdot B_i(z) + (x_i^*)^2 \cdot C_i(z)}}{1 + e^{A_i(z) + x_i^* \cdot B_i(z) + (x_i^*)^2 \cdot C_i(z)}} \quad \text{(Eq. 7)}$$

$$A_i(z) = A_{i,0} + z \cdot A_{i,1} \quad \text{(Eq. 8)}$$

$$B_i(z) = B_{i,0} + z \cdot B_{i,1} \quad \text{(Eq. 9)}$$

$$C_i(z) = C_{i,0} + z \cdot C_{i,1} \qquad \text{(Eq. 10)}$$

and q_s is the conditional probability of that infant belonging to the subpopulation s. The density submodel $f_1(x|z;\theta)$ determines that

$$q_s(x \mid z;\theta) = \frac{\pi_s(z) \cdot \tilde{N}(x \mid z;\theta_s)}{f_1(x \mid z;\theta)} \qquad \text{(Eq. 11)}$$

Model Fitting

The model was fitted to individual data by the method of maximum likelihood (*ms* routine in the SPLUS statistical library). The likelihood functions, as defined by Eqs. 1-11, were used except that the second-degree polynomial of subpopulation specific mortality curve was fitted in its respective linear form, and then transformed to the non-linear form after fitting. Overall, there are 22 parameters for the model with a dichotomous covariate z and only 11 for the model without z. Bias-adjusted 95% confidence intervals for the parameters and estimates were estimated from 200 bootstrap samples.

Decomposition of the Maternal Education Effect on Infant Mortality

Decomposition of the maternal education effect was carried out in two steps. First, the total absolute maternal education effect on infant mortality was decomposed into deaths attributable to differences in the mixing proportion and rate effects for "normal" (i.e. subpopulation p) and "compromised" births (i.e. subpopulation s) using the standard Kitagawa decomposition method (Gupta 1978).

Then the maternal education effect on the overall infant death (integrated across all birth weights) in each subpopulation was further decomposed into two multiplicative components by factoring the respective subpopulation specific relative risk (RR_i, for $i = p$ and s) of overall infant death for mothers receiving higher education (i.e. college and above, $z = 1$) as compared

to that for mothers receiving lower education (i.e. high school and below, , $z = 0$) into a "direct" factor ($F_{i,1}$) and an "indirect" factor ($F_{i,2}$) of maternal education:

$$RR_i = \frac{\overline{P_i(y = 1 \mid z = 1, \beta_i)}}{\overline{P_i(y = 1 \mid z = 0, \beta_i)}} = F_{i,1} \cdot F_{i,2} \qquad \text{(Eq. 12)}$$

$$F_{i,1} = e^{A_{i,1}} \qquad \text{(Eq. 13)}$$

$$F_{i,2} = \frac{\sum \left\{ \tilde{N}(x_i \mid z = 1; \theta) \cdot e^{A_{i,0} + (C_{i,0} + C_{i,1}) \cdot \left[x_i^* + (B_{i,0} + B_{i,1})\right]^2} \cdot \left[1 - P_i\left(y = 1 \mid z = 1, x_i^*; \beta_i\right)\right] \right\}}{\sum \tilde{N}(x_i \mid z = 1; \theta) \cdot \overline{P_i(y = 1 \mid z = 0, \beta_i)}} \qquad \text{(Eq. 14)}$$

$F_{i,1}$ is a constant and therefore it is independent of birth weight. $F_{i,2}$ represents the combined effects of all birth weight related factors on the mortality disparity between infants born to mothers with different levels of education in subpopulation i. In particular, birth weight related factors include differences in shape and horizontal shift of the reverse-J-shaped standardized birth weight specific mortality curve, as well as non-linear transformation between the probability and the logit of infant death at any standardized birth weight.

RESULTS

Education influences the birth weight distribution through changes in the mean and standard deviation of both the primary and secondary birth weight distributions (Figure 3, Table 3). The mean of the primary subpopulation increases significantly with education in all birth cohorts examined. The increase is largest in European Americans (about 100 grams) and smallest in Mexican Americans (about 30 grams). The standard deviation of the primary subpopulation birth weight density declines in the European (about 20 grams) and Mexican (perhaps 4 grams) American cohorts, but these declines are significant only for European Americans. Among African American births, the primary standard deviation increases, in both sexes (perhaps 9

grams), but is significant only in males. Overall these shifts in the birth weight distribution represent improved birth outcomes using the standard metrics, such as mean birth weight, or the estimated low birth weight rate (proportion of births less then 2500 grams) (Figure 3a). In particular, the low birth weight rate declines in the primary subpopulation of all populations (Table 2).

Figure 3 about here

Table 2 about here

The mean birth weight of the secondary subpopulation increases significantly in the European American cohorts (227 to 253 grams) (Figure 3b, Table 2). The standard deviation of secondary birth weight declines significantly with education among the European American cohorts (by 51 to 66 grams), but increases significantly in the African American cohorts (by 114 to 152 grams). Overall, the increase in mean birth weight and decline in the standard deviation clearly improve European American birth weight densities. The decline in the secondary low birth weight rate in these populations exceeds 8% (Table 2). The increase in the standard deviation among secondary African American births has little effect on the low birth weight rate (Table 2), since the mean of secondary birth weight is below 2500 grams and the density is truncated at 500 grams. Consequently, among African and Mexican American birth cohorts there is little or no change in the secondary birth weight densities associated with educational level.

Finally, higher education is associated with an increase in the proportion of primary births. Although this is typically less than 1%, and is only significant in European and African American male births, nevertheless it represents an improvement in birth outcomes using the standard metrics of mean birth weight and the low birth weight rate since the primary subpopulations have a higher mean and smaller standard deviation, and hence a lower low birth

weight rate. Consequently, considering all birth weight density factors, overall birth outcomes improve significantly with higher education in all populations.

The logistic regression results indicate that mortality generally declines with increasing education (Figure 4). Kitagawa decomposition attributes the majority of the absolute decline in mortality with education to the primary subpopulation in all birth cohorts accept Mexican American females (Table 3, Figure 4a). Changes in secondary subpopulation mortality are smaller and are significant only in the European American birth cohorts (Table 3, Figure 4b). Finally the mixing proportion effect reduces infant mortality in most birth cohorts, significantly in African and European males. However, this latter effect is small.

Figure 4 about here

Table 3 about here

Further decomposition into “direct” (independent of birth weight) and “indirect” (associated with changes in birth weight) effects indicates a strong “direct” effect in the primary subpopulation, which reduces infant mortality (Table 4). This direct effect is significant in all birth cohorts accept Mexican American males. There is also an “indirect” effect in the primary subpopulation, which tends to increase mortality with higher education. It is significant in all three female cohorts but no male birth cohorts. That the ”indirect” effects increase mortality with higher education is surprising given that the primary birth weight distribution improves with increased education as described above. The results for the secondary subpopulation are less consistent. In general a “direct” effect in the secondary subpopulation reduces mortality but is significant only among European and African American males and Mexican American females. The “indirect” effect generally increases mortality with education but is only significant in European and African American males. Again, this “indirect” effect is not consistent with the

improvements in the birth weight distribution, observed in European American male secondary birth weight densities described above. African American male secondary birth weight densities are largely unaffected by education. Perhaps the lack of consistently significant results is due to the relatively small absolute effects associated with the secondary subpopulation (Table 4).

Table 4 about here

Overall, effects of education on mortality are shown in Figure 4c. Note that at low birth weights low education is associated with lower mortality, while at normal birth weights low education is associated with higher mortality, and higher mortality overall (Table 2) compared to the high education subpopulation.

DISCUSSION

There are limitations to the validity of comparisons across race and ethnicity. In particular, these analyses are not controlled for such variables as maternal age, parity, etc. which might mediate the effects. Never the less a number of differences stand out. First, education appears to have little effect on Mexican American birth outcomes compared to the other birth cohorts. The effect of higher education on the Mexican American birth weight distributions tends to be small. Primary mean birth weight does increase but the effect is only about 1/3 of the effect observed in European and African American birth cohorts. The impact of education on mortality is smaller as well. In fact an education effect is only significant for the male cohort. It is not clear why this population does not respond to the higher levels of education like the other birth cohorts. However, Mexican American birth cohorts have the lowest observed infant mortality rate of any of the populations examined (Table 1). In addition many Mexican American births are to Mexican born mothers. Perhaps the differences in response to education noted above are a

part of the “nativity complex” that is thought to influence infant mortality in recent migrant populations (Hummer et al. 1999),

Second, education increases the African/European American racial disparity in infant mortality. That is the high education population has a larger racial differential both in absolute difference in the death rates and relative risk! In particular, the low education group has a racial disparity (relative risk) of 1.64 and 1.59 respectively for females and males, while the high education group has a racial disparity of 2.5 and 2.6 for females and males respectively. The fact that racial disparities increase with increasing education levels has been observed in a number of settings (Sngh, and Yu 1995; Din-Dzietham and Hertz-Picciotto 1998). Our decomposition attributes the majority of this increase in the racial disparity to the “compromised” subpopulation. The mixing proportion favors African Americans although this is significant only in males. The “normal” subpopulation contributes slightly to the racial differential in that mortality declines more with increasing education among European Americans, but this difference is not significant for either sex. The major contributor is that mortality improves significantly among “compromised” European American births and not among “compromised” African American births with increased education. Decomposition into “direct” and “indirect” effects is not consistent across the sexes in the “compromised” subpopulation, so it is unclear whether these are birth weight dependent or birth weight independent effects. Additional analysis will be necessary to determine if these racial and ethnic differences are consistent, as well as, how these differences are influenced by confounders such as maternal age and parity etc. On the other hand, comparison across education within populations has fewer limitations and is much more consistent.

It is well documented that birth weight increases and infant mortality declines with increased maternal education (Gortmaker 1979; Cramer 1987; Sngh, and Yu 1995; Din-Dzietham and Hertz-Picciotto 1998), although infant mortality has been reported to increase again with very high levels of education (Shoham-Yakubovich and Barel 1988; Haglund 2008). The results presented above are consistent with these general findings. The division of education levels by less then 12 completed years of education versus greater than 12 years of education does not allow an examination of birth weight and infant mortality at very high educational levels. Overall, birth weight increases with educational level, as indicated by increased mean birth weight, declines in the standard deviation of birth weight and hence a decline in the low birth weight rate. These are largely driven by the "normal" subpopulation of births, the dominant group occurring predominately in the center of the birth weight distribution. The "compromised" births, which accounts for the majority of births in the tails of the birth weight distribution, but also accounts for births at all birth weights, is more heterogeneous. European American "compromised" births weights improve, but African and Mexican American "compromised" births remain more or less constant.

The decline in infant mortality with higher education is largely due to the "normal" subpopulation. These beneficial effects are entirely independent of the changes in the birth weight distribution. All of the birth weight dependent effects tend to increase infant mortality despite general improvements (or little change) in the birth weight density! These birth weight dependent effects are overwhelmed by the "direct" effects of education, so overall infant mortality declines with increasing education. However, at low birth weights, less maternal education is associated with lower infant mortality. This is the "educational paradox" which has been reported by others (Gisselmann 2005). Why should the socially disadvantaged births have

lower infant mortality at low birth weights, that is those birth weights where social advantage and aggressive medical intervention might have the most beneficial effect? Further, these results indicate that the current U.S. National policy of reducing infant mortality by reducing the low birth weight rate (US-DOHHS, 2000} may be ineffective or even detrimental. The policy could increase mortality, if the interventions chosen to improve birth weight do not also happen to carry along the birth weight independent beneficial effects. Finally, education could provide a more beneficial effect if the birth weight dependent effects could be avoided.

Close examination of the "indirect" effects in the "normal" subpopulation indicates that the birth weight specific increase in infant mortality is due to the birth weight specific mortality curve becoming increasingly (more severely) reverse J-shaped with increasing education. Mortality improves at the optimum birth weight due to the "direct" effects, but increases more rapidly at relative (Z-scored) birth weights away from the optimum. This occurs regardless of the changes in standard deviation of primary birth weight, i.e. in European American females the standard deviation declines with education, among African American females it increases and among Mexican American females it remains the same. It is possible that this increased reverse J-shape, is a result of relaxed fetal loss among the high education populations, resulting in lower average vitality at live birth particularly at the relatively low and high birth weights. If this is correct, then the detrimental "indirect" effects might also be viewed as beneficial, a result of improved in utero survival. This would resolve some of the issues presented above. For example, it resolves the educational paradox, which is no longer a paradox but the result of benefits (in this case improved in utero survival) to the advantaged population. It implies that the birth weight dependent effects of education cannot be avoided since they are due to an unobserved part of the beneficial effects of education. And finally this interpretation implies that current US policy (US-

DOHHS, 2000) would not be detrimental. On the other hand, if birth weight is not on the causal pathway to infant mortality then improving birth weight may have little effect on infant mortality.

The results and interpretations presented above provide statistical support to the hypothesis of Wilcox and others (Wilcox and Russell 1990) that birth weight is not on the "causal" pathway to infant mortality. In particular, the statistical evidence indicates strong associations of birth weight with education, but that the improvements in infant mortality are a result of birth weight independent effects. The facts that a) significant changes in the birth weight density are not always accompanied by a change in infant mortality, b) the "indirect" and potential "causal" effects which are significant are in the wrong direction, and c) alternative explanations of increasing infant mortality with increased birth weight based on unobserved heterogeneity (fetal loss) all support the hypothesis that birth weight is not on the "causal" pathway to infant mortality. Similar results have been found using the same methods and maternal age as an instrument instead of maternal education (Gage et al. 2009). In this case, however, no convincing evidence for "indirect" effects was observed at all in either "normal" or "compromised" births.

While CDDmlr appears to provide a reasonable test of Wilcox's theory, there are some differences and limitations. In particular, Wilcox's theory implicitly assumes that shape effects which depend upon the standard deviation of birth weight and the shape of birth weight specific mortality curve are constant among "normal" births. CDDmlr relaxes this assumption and allows the standard deviation to change and the shape of the birth weight specific mortality curve to change. We also considered this potential uncoupling of birth weight and mortality to be potential "causal" effects in addition to the shift effects described by Wilcox. All of the indirect

effects reported here are driven by changes in shape and not a differential shift in the birth weight density and the infant mortality curve. A limitation of Wilcox's original theory, and CDDmlr as defined here, is that they do not completely account for all of the potential influence of birth weight on infant mortality. In particular, Wilcox's original theory assumes that the reverse J-shaped birth weight specific mortality curve is constant (Wilcox and Russell 1990). It is possible that birth weight is responsible for the reverse J-shape of the infant mortality curve. Recently, Basso (Basso, Wilcox and Weinberg 2006) has provided an extension of the original theory that attributes the reverse J-shape to confounding. CDD mixture of logistic regression could be used to explore this possibility as well. However this is beyond the scope of the present paper.

The advantage of CDDmlr is that it can be used to exquisitely define changes in the birth weight density and the infant mortality curve, and distinguish between "direct" and "indirect" (through birth weight) effects of a covariate while controlling for unobserved heterogeneity, which we have interpreted in the birth weight case as "normal" and "compromised" fetal development. This allows the identification and decomposition of trends in birth weight and infant mortality into a number of components, which do not necessarily reflect the overall trends. Here this has allowed us to identify "indirect" effects of education, which are in the opposite direction of the overall trends. The analysis above, as well as, another application using maternal age (Gage et al. 2009) all indicate that the association between birth weight and infant mortality is not necessarily "causal". It would be useful to repeat these analyses using gestational age in place of birth weight. While CDDmlr was designed specifically to examine the issue of causality in birth outcomes it is not limited to this application. Compared to conventional regressions CDDmlr is useful wherever the density of a potentially mediator, like birth weight is well

described by a finite mixture model, and this mixture accounts for some unmeasured heterogeneity in the ultimate dependent variable, such as infant mortality. If so then the introduction of additional covariates can test if the mediator could be causal or not. For example, in a case similar to birth weight, another biomarker, body mass index, a measure of obesity, is closely associated with mortality in adults (Waaler 1984) again in a reverse J-shaped pattern. Recently, it has been shown that the relationship between body mass index and mortality is not fixed but relative to the mean of body mass index (Su 2005). This is very similar to the dynamics of birth weight and infant mortality, and the original basis for arguing that birth weight was not on the causal pathway to infant mortality (Wilcox and Russell 1990).

CONCLUSIONS

Education, a surrogate measure of socioeconomic status, is associated with significant changes in birth outcomes including infant mortality.

1. Overall birth outcomes, such as mean birth weight, standard deviation in birth weight and low birth weight rate improve with higher education.
2. Mortality declines with higher educations. This is entirely independent of the changes in birth outcomes, and due to “direct” effects.
3. All significant “indirect” effects (potentially causal through birth weight), tend to increase infant mortality even though birth outcomes generally improve. These increases in mortality are overwhelmed by the beneficial “direct” effects. This counter intuitive “indirect” effect is due to increased severity of the reverse J-shaped birth weight specific mortality curve. As a result infant mortality among the less educated tends to be lower at low birth weights, i.e. the education paradox. This could be the result of relaxed fetal

selection particularly at the low and high birth weights among the higher educated population. Thus these indirect effects could be the result of uncontrolled heterogeneity, rather then “causal” forces operating through birth weight.

4. Birth outcomes and infant mortality do not appear to be as strong among the Mexican American population, compared to African and European American births.
5. African /European disparities increase with educational level due largely to improvements in mortality among “compromised” births among European Americans but not African American births.

REFERENCES

Basso, O., A. Wilcox, and C. Weinberg. 2006. "Birth weight and mortality: Causality or confounding?" *Am. J. of Epidemiol.* 164:303-311.

Buehler, J.W., J.C. Kleinman, C.J.R. Hogue, L.T. Strauss, and J.C. Smith. 1987. "Birth weight-specific infant mortality, United States, 1960 and 1980." *Public Health Reports* 102(2):151-161.

Cramer, J.C. 1987. “Social Factors and Infant Mortality: Identifying High-Risk Groups and Proximate Causes,” Demography 24(3): 299-322.

Din-Dzietham, R. and I Hertz-Picciotto, 1998 “Infant mortality differences between whites and African Americans: the effect of maternal education.,” Am J Public Health 88(4): 651-656.

Floud, R., K. Wachter, and A. Gregory. 1990. *Height, Health and History*. Cambridge: Cambridge University Press.

Gage, T.B., M.J. Bauer, N. Heffner, and H. Stratton. 2004. "The pediatric paradox: Heterogeneity in the birth cohort." *Human Biology* 76:327-342.

Gage, T.B., F. Fang, E. O'Neill, and H. Stratton. 2009. "Maternal age and infant mortality: A test of the Wilcox-Russell hypothesis." *American Journal of Epidemiology* 169(3):294-303.

Gisselmann, M.D., 2005. “Education, infant mortality, and low birth weight in Sweden 1973--1990: Emergence of the low birth weight paradox,” Scand J Public Health 33(1): 65-71.

Gortmaker, S.L. 1979 “Poverty and Infant Mortality in the United States,” American Sociological Review 44(2): 280-297.

Gupta, P.D. 1978. "A General Method of Decomposing a Difference Between Two Rates into Several Components." *Demography* 15(1):99-112.

Haglund, B., S. Cnattingius, and M.-L. Nordström, 1993 “Social differences in late fetal death and infant mortality in Sweden 1985-86,” Paediatric and Perinatal Epidemiology 7(1): 33-44.

Hummer, R.A., M. Biegler, P.B. DeTurk, D. Forbes, W.P. Frisbie, Y. Hong, and S.G. Pullum. 1999. "Race/ethnicity, nativity, and infant mortality in the United States." *Social Forces* 77(3):1083-1118.

Institute-of-Medicine. 1985. *Preventing Low Birthweight*. Washington: National Academy Press.

Mc Cormick, M.C. 1985. "The contribution of low birth weight to infant mortality and childhood morbidity." 312(2):82-90.

Mosely, W.H.and L.C. Chen. 1984. "An analytical framework for the study of child survival in developing countries." in *Population and Development Review Supplement to Vol. 10.*, edited by W.H. Mosely and L.C. Chen. Cambridge: Cambridge University Press.

US-DOHHS. 2000. "Healthy People 2010: Understanding and improving health." Pp. 91-95. Washington, DC: US Department of Health and Human Services.

Shoham-Yakubovich, I. and V. BAREL, 1988 “Maternal Education as a Modifier of the Association between Low Birthweight and Infant Mortality,” Int. J. Epidemiol. 17(2): 370-377.

Singh, G.K. and S M Yu, 1995, “Infant mortality in the United States: trends, differentials, and projections, 1950 through 2010.,” Am J Public Health 85(7): 957-964.

Su, D. 2005. "Body mass index and old-age survival: A comparative study between the Union Army Records and the NHANES-I Epidemiological Follow-Up Sample." *American Journal of Human Biology* 17(3):341-354.

Waaler, H. 1984. *Height, Weight and Mortality: The Norwegian Experience*. Oslo: Gruppe for Helsetjenesteforskning.

Wilcox, A.and I. Russell. 1990. "Why small black infants have a lower mortality rate than small white infants: The case for population-specific standards for birth weight." *J. of Ped.* 116:7-10.

Wise, P.H. 2003. "The anatomy of a disparity in infant mortality." *Annual Review of Public Health* 24:341-362.

FIGURE CAPTIONS

Figure 1. Graphical representation of the Wilcox-Russell definition of "causality". Panel a represents a shift in birth weight that is accompanied with a shift in the birth weight specific mortality curve so that no change in mortality occurs (birth weight is not "causal"). Panel b represents a birth weight independent change in infant mortality, (direct effect, birth weight is not "causal"). Pane c represents a shift in birth weight that is not accompanied by an identical shift in the birth weight specific mortality curve so that mortality does change.

Figure 2. Graphical representation of Covariate Density Defined mixture of logistic regressions as applied to birth weight. The results presented are 2001 European American males based on analyses developed below in this paper. Totals are solid lines. Short dotted lines represent the primary subpopulation, and longer dashed lines the secondary subpopulation. Corresponding fine lines are 95% confidence limits (sometimes they overlap the predicted line and are not visible). Panel a represents the density of total birth weight as the sum of two Gaussian densities, primary and secondary. Secondary is considered to represent "compromised" births because it accounts for the most low birth weight and most macrosomic infants. Panel b represents characteristic total, as well as, primary and secondary birth weight specific mortality curves. Results for the other 5 populations are similar.

Figure 3. Shift in birth weight density due to higher education: European American males. The solid line represents low education, and the dashed line high education. Corresponding fine lines are 95% confidence limits (sometimes they overlap the predicted line and are not visible). Panel a represents the primary subpopulation, panel b the secondary subpopulation and panel c the total population. African American males and females, and European American females are similar. The shift for Mexican Americans is much smaller.

Figure 4. Shifts in birth weight specific due to higher education: European American males. The solid line represents low education, and the dashed line high education. Corresponding fine lines are 95% confidence limits (sometimes they overlap the predicted line and are not visible). Panel a represents the primary subpopulation, panel b the secondary subpopulation and panel c the total population. African American males and females, and European American females are similar. The shift for Mexican Americans is much smaller.

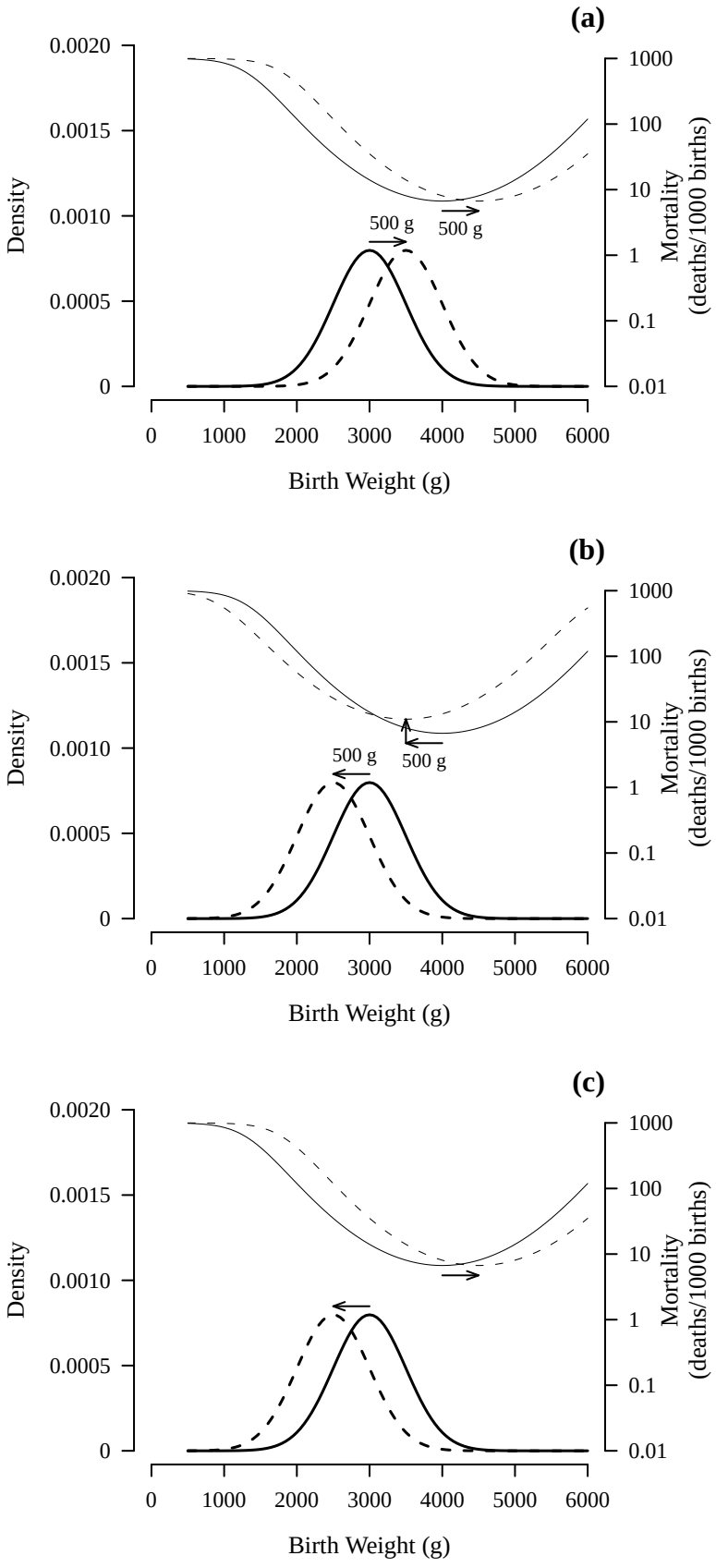
(a)
0.0020
0.0015
0.0010
0.0005
0
Density
1000
100
10
1
0.1
0.01
Mortality
(deaths/1000 births)
500 g
500 g
0
1000
2000
3000
4000
5000
6000
Birth Weight (g)
(b)
0.0020
0.0015
0.0010
0.0005
0
Density
1000
100
10
1
0.1
0.01
Mortality
(deaths/1000 births)
500 g
500 g
0
1000
2000
3000
4000
5000
6000
Birth Weight (g)
(c)
0.0020
0.0015
0.0010
0.0005
0
Density
1000
100
10
1
0.1
0.01
Mortality
(deaths/1000 births)
0
1000
2000
3000
4000
5000
6000
Birth Weight (g)

Figure 1

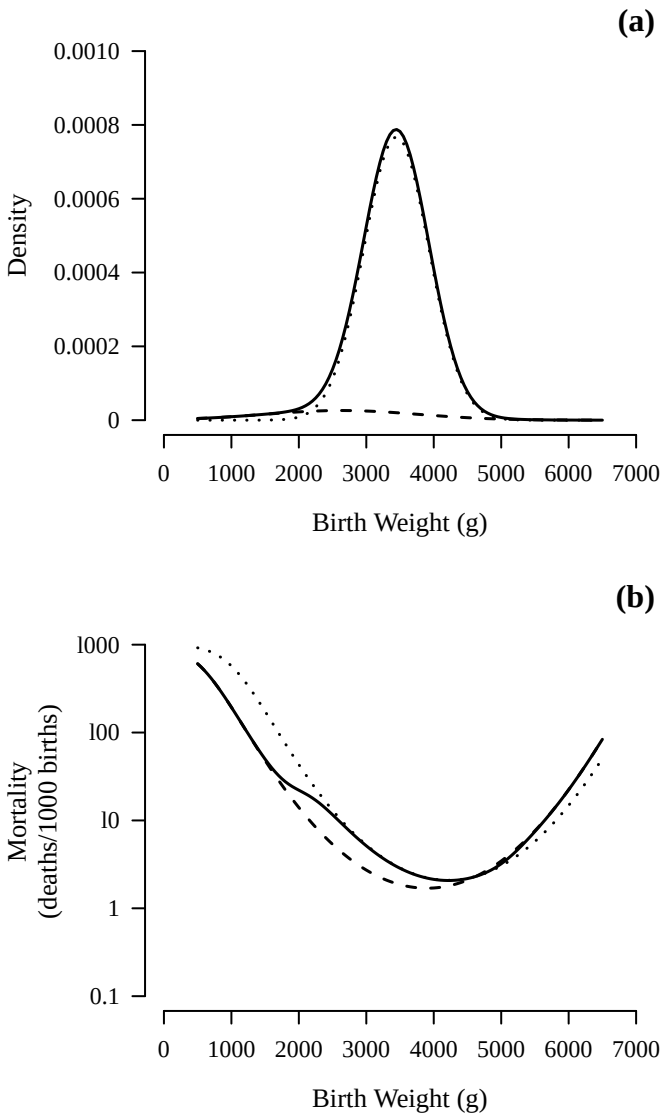
(a)
0.0010
0.0008
0.0006
0.0004
0.0002
0
Density
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)
(b)
1000
100
10
1
0.1
Mortality
(deaths/1000 births)
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)

Figure 2

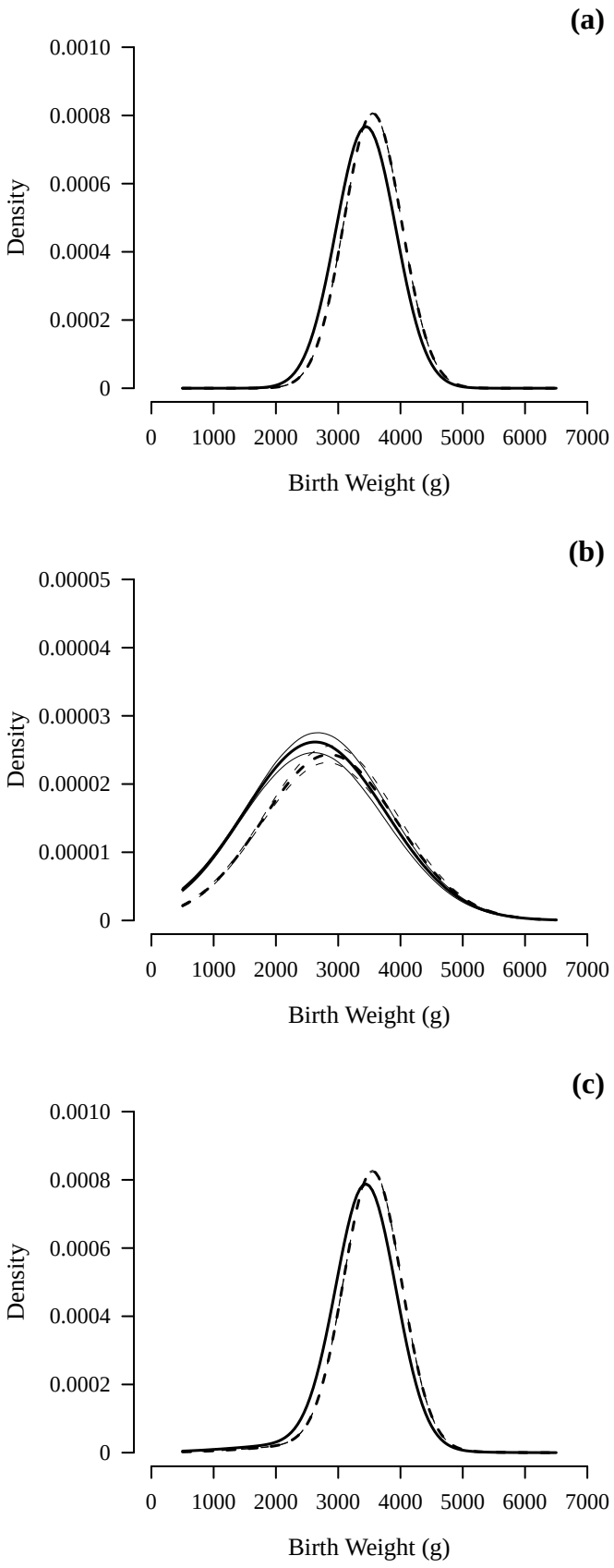
(a)
0.0010
0.0008
0.0006
0.0004
0.0002
0
Density
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)
(b)
0.00005
0.00004
0.00003
0.00002
0.00001
0
Density
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)
(c)
0.0010
0.0008
0.0006
0.0004
0.0002
0
Density
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)

Figure 3

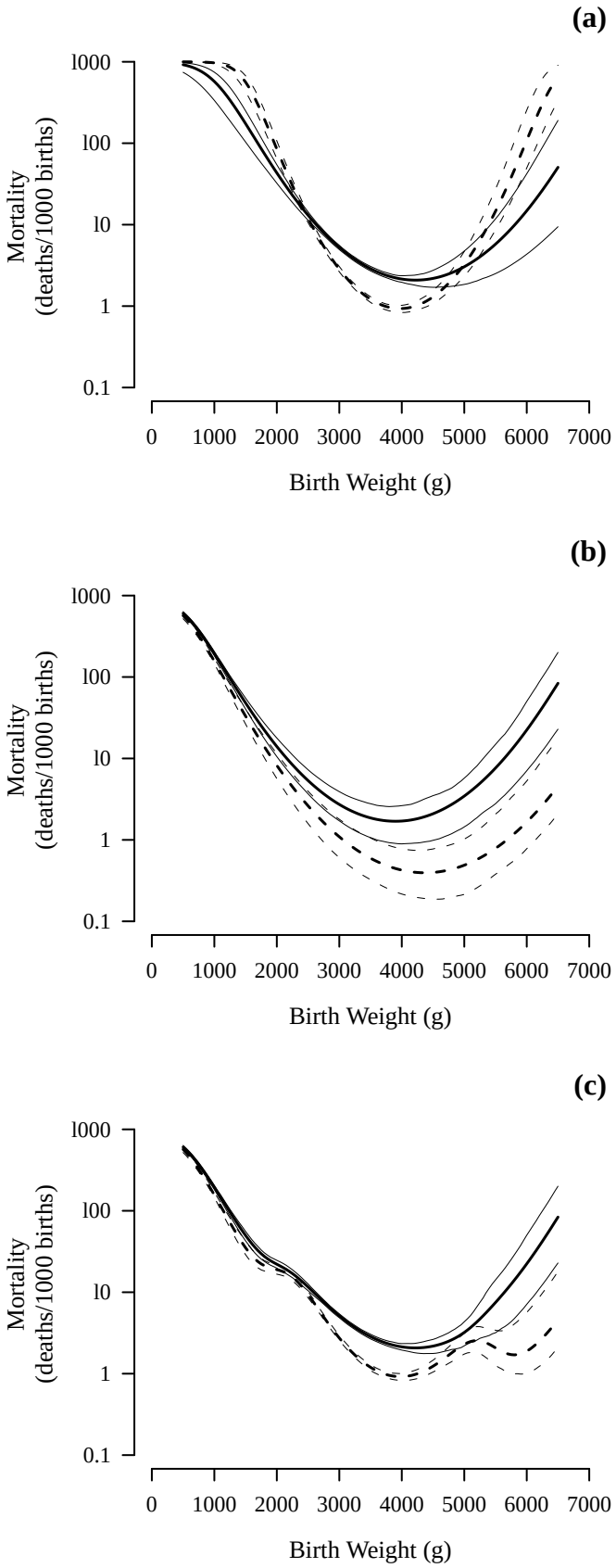
(a)
1000
100
10
1
0.1
Mortality
(deaths/1000 births)
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)
(b)
1000
100
10
1
0.1
Mortality
(deaths/1000 births)
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)
(c)
1000
100
10
1
0.1
Mortality
(deaths/1000 births)
0
1000
2000
3000
4000
5000
6000
7000
Birth Weight (g)

Figure 4

Table 1 Descriptive statistics for the sample populations

Birth Cohort	high school and below (z=0)		Birth Weight Distribution (g) for z=0 birth cohort					
	# of Births	CDR	mean	5%	25%	median	75%	95%
non-His. Eur. Am. F.	432,034	4.73	3278	2410	2977	3290	3628	4111
non-His. Eur. Am. M.	454,249	6.36	3393	2460	3090	3430	3751	4253
non-His. Af. Am. F.	162,394	7.72	3061	2055	2778	3095	3420	3912
non-His. Af. Am. M.	167,571	10.22	3165	2097	2863	3210	3544	4038
Mex. Am. F.	222,480	3.57	3299	2495	3005	3315	3610	4090
Mex. Am. M.	230,989	4.45	3393	2523	3090	3402	3735	4224
Birth Cohort	college and above (z=1)		Birth Weight Distribution (g) for z=1 birth cohort					
	# of Births	CDR	mean	5%	25%	median	75%	95%
non-His. Eur. Am. F.	583,889	2.45	3390	2580	3090	3402	3714	4180
non-His. Eur. Am. M.	614,327	3.05	3512	2637	3204	3540	3856	4338
non-His. Af. Am. F.	89,290	6.04	3150	2154	2863	3203	3515	4005
non-His. Af. Am. M.	92,281	7.86	3266	2211	2977	3317	3655	4155
Mex. Am. F.	43,746	2.97	3322	2523	3033	3335	3629	4111
Mex. Am. M.	45,275	3.31	3424	2551	3135	3446	3761	4252

non-His. = non-Hispanic; Eur. = European; Af. = African; Mex. = Mexican; Am. = American; F. = females; M. = males

CDR = Crude death rate (deaths per 1000 births)

Table 2 Model-estimated change in infant birth weight distribution and mortality characteristics with bias-adjusted 95% confidence intervals: infants born to mothers with college and above education vs. infants born to mothers with high school and below education

Birth Cohort	Females			Males		
	Non-His. Eur. Am.	Non-His. Af. Am.	Mex. Am.	Non-His. Eur. Am.	Non-His. Af. Am.	Mex. Am.
μ_p (g)	103 (101 ; 105)*	88 (83 ; 93)*	27 (22 ; 31)*	108 (106 ; 110)*	95 (91 ; 99)*	35 (30 ; 41)*
σ_p (g)	-20 (-22 ; -18)*	7 (1 ; 11)*	-3 (-10 ; 3)	-19 (-22 ; -17)*	9 (3 ; 14)*	-4 (-12 ; 2)
Pri. LBW	-2.0 (-2.1 ; -1.9)*	-2.0 (-2.3 ; -1.7)*	-0.5 (-0.7 ; -0.3)*	-1.4 (-1.5 ; -1.3)*	-1.5 (-1.8 ; -1.3)*	-0.4 (-0.5 ; -0.2)*
Pri. DR	-1.8 (-2.0 ; -1.6)*	-1.6 (-2.2 ; -1.1)*	-0.1 (-0.6 ; 0.5)	-2.2 (-2.4 ; -1.9)*	-2.2 (-2.7 ; -1.5)*	-0.8 (-1.4 ; -0.3)*
π_s (%)	-0.4 (-0.9 ; 0.0)	-0.8 (-1.6 ; 0.1)	-0.3 (-1.5 ; 0.9)	-0.8 (-1.2 ; -0.4)*	-1.6 (-2.4 ; -0.8)*	0.2 (-1.0 ; 1.7)
μ_s (g)	253 (202 ; 304)*	-65 (-247 ; 97)	-36 (-183 ; 73)	227 (169 ; 273)*	-136 (-317 ; 20)	4 (-119 ; 126)
σ_s (g)	-51 (-80 ; -28)*	114 (41 ; 175)*	-24 (-101 ; 47)	-66 (-87 ; -42)*	152 (93 ; 217)*	-29 (-95 ; 29)
Sec. LBW	-8.3 (-9.9 -6.5)*	-0.5 (-3.4 ; 2.7)	1.3 (-2.7 ; 6.3)	-7.6 (-9.1 ; -5.5)*	0.4 (-2.4 ; 3.5)	-0.2 (-4.5 ; 4.3)
Sec. DR	-8.7 (-13.1 ; -5.2)*	0.2 (-6.1 ; 6.2)	-6.3 (-13.0 ; 1.7)	-16.0 (-19.5 ; -11.8)*	4.1 (-3.2 ; 12.7)	-5.1 (-11.0 ; 2.9)
Tot. LBW	-2.6 (-2.6 -2.4)*	-2.3 (-2.5 ; -2.0)*	-0.5 (-0.7 ; -0.3)*	-2.1 (-2.2 ; -2.1)*	-2.2 (-2.4 ; -2.0)*	-0.3 (-0.5 ; -0.1)*
Tot. DR	-2.3 (-2.5 ; -2.0)*	-1.7 (-2.3 ; -1.0)*	-0.6 (-1.1 ; 0.0)	-3.3 (-3.6 ; -3.0)*	-2.4 (-3.0 ; -1.6)*	-1.1 (-1.7 ; -0.5)*

non-His. = non-Hispanic; Eur. = European; Af. = African; Mex. = Mexican; Am. = American; F. = females; M. = males

DR = death rate (deaths per 1000 births)

Pri. = Primary Subpopulation; Sec. = Secondary Subpopulation; Tot. = Total Population

LBW = % of births less than 2500 g

* Parameter is significantly different from 0.

Table 3 Kitagawa decomposition of infant mortality (deaths/1000 births) disparity with bias-adjusted 95% confidence intervals due to maternal education effect [#]

Birth Cohort	Mixing Proportion Effect	Rate Effect		Total Disparity
		Secondary	Primary	
non-His. Eur. Am. F.	-0.08 (-0.16 ; 0.00)	-0.52 (-0.79 ; -0.33)*	-1.68 (-1.90 ; -1.47)*	-2.28 (-2.54 ; -2.06)*
non-His. Eur. Am. M.	-0.19 (-0.28 ; -0.09)*	-1.09 (-1.34 ; -0.78)*	-2.03 (-2.26 ; -1.78)*	-3.31 (-3.58 ; -3.03)*
non-His. Af. Am. F.	-0.23 (-0.51 ; 0.02)	0.02 (-0.58 ; 0.59)	-1.50 (-2.04 ; -1.05)*	-1.71 (-2.34 ; -0.97)*
non-His. Af. Am. M.	-0.75 (-1.16 ; -0.34)*	0.38 (-0.30 ; 1.20)	-2.00 (-2.48 ; -1.31)*	-2.37 (-3.07 ; -1.57)*
Mex. Am. F.	-0.05 (-0.26 ; 0.14)	-0.42 (-0.89 ; 0.12)	-0.12 (-0.59 ; 0.49)	-0.60 (-1.11 ; 0.01)
Mex. Am. M.	0.04 (-0.28 ; 0.34)	-0.41 (-0.91 ; 0.29)	-0.77 (-1.26 ; -0.25)*	-1.14 (-1.72 ; -0.51)*

[#]: Based on estimated death rates in Table 2

non-His. = non-Hispanic; Eur. = European; Af. = African; Mex. = Mexican; Am. = American; F. = females; M. = males

*: Estimate is significantly different from 0

Table 4 Subpopulation specific relative risk of infant mortality due to higher maternal education decomposed into direct and indirect multiplicative factors with bias-adjusted 95% confidence intervals [#]

Birth Cohort	Primary Subpopulation		
	Relative Risk	Direct Factor	Indirect Factor
non-His. Eur. Am. F.	0.47 (0.42 ; 0.52)*	0.40 (0.35 ; 0.47)*	1.16 (1.03 ; 1.36)*
non-His. Eur. Am. M.	0.48 (0.44 ; 0.52)*	0.45 (0.37 ; 0.56)*	1.07 (0.87 ; 1.21)
non-His. Af. Am. F.	0.66 (0.57 ; 0.76)*	0.46 (0.37 ; 0.57)*	1.45 (1.22 ; 1.76)*
non-His. Af. Am. M.	0.63 (0.55 ; 0.74)*	0.58 (0.47 ; 0.72)*	1.07 (0.86 ; 1.30)
Mex. Am. F.	0.94 (0.74 ; 1.25)	0.55 (0.10 ; 0.90)*	1.69 (1.03 ; 8.65)*
Mex. Am. M.	0.69 (0.49 ; 0.89)*	0.96 (0.65 ; 1.33)	0.72 (0.50 ; 1.02)
Birth Cohort	Secondary Subpopulation		
	Relative Risk	Direct Factor	Indirect Factor
non-His. Eur. Am. F.	0.66 (0.55 ; 0.77)*	1.33 (0.50 ; 3.10)	0.50 (0.18 ; 1.55)
non-His. Eur. Am. M.	0.54 (0.47 ; 0.63)*	0.23 (0.07 ; 0.57)*	2.31 (1.27 ; 4.71)*
non-His. Af. Am. F.	1.01 (0.83 ; 1.20)	0.53 (0.07 ; 3.34)	1.88 (0.25 ; 15.57)
non-His. Af. Am. M.	1.08 (0.94 ; 1.28)	0.14 (-0.01 ; 0.74)*	7.60 (2.93 ; 40.89)*
Mex. Am. F.	0.71 (0.46 ; 1.10)	0.22 (0.07 ; 0.96)*	3.23 (0.30 ; 23.04)
Mex. Am. M.	0.80 (0.59 ; 1.16)	0.47 (0.12 ; 1.74)	1.68 (0.54 ; 4.07)

[#]: Based on estimated death rates in Table 2

non-His. = non-Hispanic; Eur. = European; Af. = African; Mex. = Mexican; Am. = American; F. = females; M. = males

*: Estimate is significantly different from 1