FAMILY BACKGROUND AND ADOLESCENT WEIGHT: AN EXAMINATION OF SOCIAL AND GENETIC INFLUENCES

Molly A. Martin, Columbia University

Gary D. Sandefur, University of Wisconsin-Madison

Abstract

This research investigates the family-level social factors associated with adolescent weight after accounting for genetic influences by using sibling resemblance models and the genetic sample of the National Longitudinal Study of Adolescent Health. By conducting sibling model analyses in a structural equation framework, we account for unmeasured family background characteristics and determine the proportion of the variation between families and that within families. We estimate models for all sibling pairs and separate models for each sibling-pair type (i.e., monozygotic twins, dizygotic twins, full siblings, half siblings) and test for differences in the influence of various family background factors across sibling-pair types. Family status characteristics, as well as family behaviors related to physical activity, inactivity and mealtime behavior, are associated with adolescent weight. The models of the different pair types bound the estimates of these family background effects and reveal that social, as well as genetic factors, contribute to an intergenerational similarity in weight.

FAMILY BACKGROUND AND ADOLESCENT WEIGHT:
AN EXAMINATION OF SOCIAL AND GENETIC INFLUENCES

Parental contributions to a child's well-being are both social and genetic, but social scientists focus on parents' social contributions. The consideration of genetics comes to the fore, however, when the outcome in question has a clear biological basis, such as childhood obesity. Due to both genetic heritability and socialization, a child is likely to have a body composition similar to that of their parents. Research investigating parental influences on child obesity, and other outcomes with a genetic component, should appropriately account for the multiplicity of parental influences. Models of individual-level data are sorely inadequate because they not only do they miss the genetic similarities within families, but these models generally underestimate the role of even social family-level factors. To arrive at better estimates of the social conditions of families important for adolescent weight, we use the genetic sample within the National Longitudinal Study of Adolescent Health (Add Health) to estimate sibling resemblance models across all types of possible sibling pairs. To estimate these effects net of the genetic similarities within pairs, we then estimate separate sibling resemblance models for each sibling-pair type (i.e., monozygotic twin, dizygotic twin, full sibling, half sibling) to arrive at better estimates of the influence of the family's social characteristics on adolescent weight.

Exploring adolescent weight is important for a number of reasons. Research demonstrates that there is a strong relationship between adolescent and adult obesity (Barlow & Dietz, 1998; Guo et al, 1994; Mossberg, 1989). In addition, obese adolescents are likely to suffer from significant physical and mental health problems as a consequence. First, obese adolescents are at greater risk of coronary heart disease morbidity and mortality, diabetes, orthopedic disorders, sleep disorders, increased cholesterol levels, and gall bladder disease

(Barlow & Dietz, 1998; Dietz, 1998). Second, negative stereotyping could lead to discrimination against obese adolescents (Dietz, 1998; Gortmaker, et al., 1993). Third, the psychological consequences of being overweight are significant because adolescents' perceptions about their weight are frequently worse than their actual measurements and a physician's assessment (Levinson, Powell, & Steelman, 1986). Finally, the psychological consequences are often compounded among adolescents who mature relatively early. Early maturing adolescents are likely to experience increased body fatness, decreased self-esteem, and an incongruity between their actual emotional maturity and the level of maturity adults perceive in them as a result of the teen's larger size. This, in turn, could result in socialization problems and feelings of frustration (Dietz, 1998). Together, the social and medical consequences of being an overweight adolescent are substantial, especially when considering the implications for their later social, psychological, and physical health.

The prevalence of adolescent obesity has increased dramatically since the 1960s, especially during the 1980s and early 1990s (Troiano & Flegal, 1998). In the early 1990s, approximately 11 percent of adolescents were overweight (Troiano & Flegal, 1998). Several physicians consider the rapid increase in obesity prevalence to be a manifestation of societal changes, primarily increases in television viewing (Troiano & Flegal, 1998; Dietz, 1990).

The proposed research seeks to better analyze the social factors associated with adolescent weight after controlling for genetic influences. By taking into account both family- and individual-level characteristics, we model adolescents' weight relative to their height using the genetic sample of the Add Health data. We assert that families play an important role in facilitating or inhibiting unhealthy weight. As such, we analyze sibling resemblance models with the family as the unit of analysis. By conducting sibling model analyses in a structural

equation framework, we can account for unmeasured family background characteristics and determine what proportion of the variation in adolescent weight is due to variation between families versus variation within families. We then estimate a multiple groups model for each sibling-pair type to allow for differential associations between family background characteristics and adolescent weight by sibling-pair type. In the following section we review the previous literature on the social and familial factors associated with adolescent weight. Then, we discuss our overall analytic strategy, as well as our data. Next we detail the results for the sibling analysis of all pair types and then discuss the results for the multiple group sibling models. Finally, we conclude with a discussion of the results and our proposals for future research.

PREVIOUS RESEARCH

An individual's risk of obesity increases if a family member is also obese (for review, see Comuzzie & Allison, 1998). Whitaker and colleagues (1997) matched medical records of parents' and children residing in Washington state and found that the risk of adult obesity was significantly greater if either the mother or father was obese at every age interval. What factors explain this intergenerational association? Previous research highlights both genetic and environmental factors.

Through twin, adoption, and family studies, previous researchers have estimated that between 40 to 70 percent of the variation in weight in humans is heritable (Jacobson & Rowe, 1998; for review, see Comuzzie & Allison, 1998). Regardless of the actual estimates of heritability, genetic factors play a large role in determining weight. On the one hand, future research should account for genetic contributions to weight even if no attempts are made to estimate its heritability. On the other hand, the rapid increase in the prevalence of childhood obesity cannot be attributed to purely genetic factors. Therefore, it is important to consider the

behavioral factors that influence obesity.

Parents and children not only share genetic endowments, they also share environmental contexts and behavioral patterns that could contribute to an intergenerational similarity in weight. Several key family background characteristics could explain part of the association in parents' and children's body compositions. First, previous research has documented racial and ethnic differences in adult and childhood obesity. Blacks and Mexican Americans are more likely to be overweight than are non-Hispanic whites (Popkin & Udry, 1998; Gordon-Larsen et al, 1999; Troiano & Flegal, 1998; Winkleby, et al., 1999). Using data from the Add Health survey, Popkin and Udry (1998) find that all racial and ethnic minority groups, with the exception of Chinese and Filipinos, have higher levels of obesity than non-Hispanic whites. Racial differences in obesity are much less pronounced among young men than among young women. For example, young black women are almost twice as likely to be obese as young white women, but prevalence differences between young black men and young white men are much smaller (Troiano & Flegal, 1998).

Second, socioeconomic status could influence both parental and child weight. Among adult women, research has consistently found an inverse relationship between socioeconomic status and overweight, but this pattern has not been consistently established among children and adolescents (Troiano & Flegal, 1998). Using the third National Health and Nutrition Evaluation Survey (NHANES) data, Troiano and Flegal (1998) find that overweight prevalence for non-Hispanic whites is inversely related to family income, but no such relationship is found for Mexican American and non-Hispanic black youths. In addition, they find no discernable pattern in the bivariate association of adolescent overweight and parental education (Troiano & Flegal, 1998). In contrast, Winkleby et al. (1999) finds socioeconomic status, measured as either the

household head's educational attainment or as the ratio of the household's income to needs, is a significant predictor of adolescent weight.

Third, parents and children share neighborhood contexts. As such, the family's access to affordable and healthy grocery stores, a community recreational center, and outdoor recreational sites, such as a park or bike trail, could be important for family members' body composition. In addition, the family's residential location will influence their transportation patterns and amount of daily walking. Finally, the safety of the family's neighborhood could influence how frequently they venture outside. Using the Add Health data, Gordon-Larsen and her colleagues (2000) find that environmental factors are associated with physical activity. The use of a community recreation center increases the likelihood of engaging in moderate to vigorous physical activity, whereas living in a neighborhood with a high level of serious crime decreased the likelihood of moderate to vigorous physical activity (Gordon-Larsen, McMurray, Popkin, 2000).

Finally, parents and children could have similar nutritional and physical activity patterns. Furthermore, because parents can influence their child's daily nutrition intake and activity levels, this is an important avenue for intervention. With regard to nutrition, children's food preferences, food intake, and energy regulation are important for childhood and adolescent obesity. Parents can play an important role in controlling children's intake of high-fat, energy-dense foods, encouraging children to eat a variety of foods, and helping children respond to internal cues of hunger and satiety (Birch & Fisher, 1998). Finally, eating is a social occasion and other family members, including parents, can model food selection preferences and orientations to dietary restraint (for a review, see Birch & Fisher, 1998). With regards to energy expenditure, both activity and inactivity are important for adolescent obesity. Boys and girls who watch more than

four hours of television each day have greater body fat and greater body-mass index (Andersen et al., 1998). Vigorous activity levels are lower among females and minority males than for non-Hispanic white males (Gordon-Larsen et al., 1999; Andersen et al., 1998).

To summarize, previous research suggests a large and important role of genetics for determining an adolescent's body composition, but the relatively recent increase in adolescent obesity cannot be explained by changes in genetic factors. Therefore, we need to better understand the behavioral and environmental factors influencing adolescent weight while adequately controlling for genetic influences. We do so by estimating sibling models. Our approach is notably different than that of behavioral geneticists for the following reasons: (1) we do not try to estimate the proportion of the variation in siblings' weight due to genetic versus environmental factors and (2) we directly estimate the effects of various environmental and behavioral characteristics on adolescent weight.

Our research also draws from the life course perspective (Elder, 1994). Not only are children's lives linked to their parents' but their own early life conditions and the accumulation of various risks or protective factors over the life course are important for their experiences in adolescence. As such, we also incorporate indicators of birth weight and breastfeeding duration to map their present physiology back to their early health status and nutritional intake.

METHODS

We estimate sibling resemblance structural equation models using data from two or more siblings to control for unobserved family background characteristics (Hauser, 1991). A sibling model decomposes a basic regression into (1) a between-family regression of common family factors (or characteristics similar across siblings) and (2) a pair of within-family regressions of factors varying across siblings (Hauser, 1988).

Sibling models offer several advantages over individual-level models. First, sibling models allow researchers to provide a global indication of the strength of family background characteristics on different outcomes. Generally, the effects of family background variables are underestimated in individual-level models (Hauser, Sheridan, and Warren, 1999). Second, sibling models offer a criterion for proposed theoretical explanations of familial and individual differences in various outcomes. By knowing what proportion of the variance in siblings' weight is explained by measured, family background variables, one can then ask what unobserved, but shared family variables explain the rest of the sibling similarity (Hauser, Sheridan, & Warren, 1999). Third, sibling models control for the global effects of family background, and therefore give better estimates of the effects that vary between and within families (Hauser, Sheridan, & Warren, 1999). In individual-level models, the within- and between-family variances operate as variation across people, so the sources of this variation cannot be identified. This third advantage is especially important for the current research. By parceling the variation in adolescents' weight into their between- and within-family components, we can arrive at better estimates of the effects of key variables on adolescent weight.

Structural equation modeling can better represent and analyze causal models of effects. Our research uses LISREL 8.5 to model family processes for adolescent weight. The model is expressed in LIRSEL notation (see Sörbom and Jöreskog, 1981) by the following three equations:

Structural model: $\eta = \beta\eta + \gamma\xi + \zeta$ (1)

Measurement models: $X = \lambda_x\xi + \delta$ (2)

$Y = \lambda_y\eta + \varepsilon$ (3)

Equation 1 represents the structural model in which η is a vector of latent endogenous

variables, β is a matrix of the effects of η on η, ξ is a vector of latent exogenous variables, γ is a matrix of the effects of ξ on η, and ζ is a vector of structural disturbances with a covariance matrix Ψ. The structural disturbances in the endogenous latent variables, ζs, are specified as uncorrelated with one another and uncorrelated with the exogenous latent variables. The exogenous latent variables are freely correlated with one another in a matrix Φ.

Equations 2 and 3 are the measurement models for the latent variables. X is a vector containing measured variables related to the exogenous latent variables. λ_x is a vector of factor loadings of X on ξ, and δ is a vector of measurement errors in X, with a variance-covariance matrix of θ_δ. Y is a vector of measured individual weight variables. λ_y is a vector of factor loadings of Y on η, and ε is a vector of measurement errors in Y, with a variance-covariance matrix θ_ε.

To elucidate the differences between models of individuals versus models of siblings, Figure 1 displays an individual-level model of the variables in our analysis and Figure 2 displays our baseline sibling model. To help reading these figures, those elements contained in squares represent measured variables (Y's and X's) and those elements in the labeled circles are the latent constructs in the structural model (η's and ξ's). The small, unlabelled circles affecting the Y and X variables indicate measurement error in their respective variables (ε's and δ's). The arrows represent causal paths from latent constructs to other latent constructs (β's and γ's) and from latent constructs to their measures (λ_y's and λ_x's). Curved, double-sided arrows represent unanalyzed correlation in measurement error.

The top of Figure 2 displays the exogenous family-level latent constructs and their measures that influence a common family factor, labeled "Family BMI" (η_1). The path coefficients from the family-level latent constructs to the common family factor derive from

differences across families. The exogenous family-level constructs measured by continuous indicators have a measurement model, but those measured by categorical variables do not. The small oval labeled "Z1" is ζ_1 and represents all other causes of the variation across families not accounted for in the present model. The common family factor, then affects each sibling's latent BMI. Each sibling's BMI is measured with three indicators.

Several individual-level variables are allowed to affect only the adolescent's own BMI. These individual-level constructs appear on the lower left and right sides of the figure. The left side of the figure represents the causal paths from these characteristics for Sibling 1's BMI (η_2), while the right side of the figure models the characteristics of Sibling 2 (η_3). Since there is only one variable per each individual-level construct, the paths to the measured variables all equal 1.0 and there is no estimated measurement error. The path coefficients from the individual-level constructs to each sibling's latent BMI derive from variation within the family. The latent constructs labeled "Z2" and "Z3" captures each sibling's characteristics that are (1) not derived from the family and (2) are different from the other sibling.

For a multiple groups model, we estimate a sibling model separately for each pair type (i.e., monozygotic twin, dizygotic twin, full sibling, half sibling) to arrive at better estimates of the influence of the family's social characteristics on adolescent weight. In essence, we will estimate a model like that portrayed in Figure 2 for each sibling pair type. Then, we test what parameters are equal across the different groups. Given the genetic influences on weight, one would expect monozygotic twins to have a lower within-family variation in BMI, followed by full siblings and dizygotic twins, and then half-siblings. The multiple groups model accounts for the expected differences in the within-family variation.

DATA

We estimate these sibling resemblance models using the genetic sample of the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a nationally representative sample of adolescents in grades 7-12 in the United States in 1995 (Wave 1) and 1996 (Wave 2). All students were surveyed in school and a sample of students was also interviewed at home. Add Health is well suited for sibling resemblance models because a substantial number of sibling or sibling-like pairs were identified and interviewed separately as part of a genetic supplement to the in-home survey. Siblings were identified in the in-school survey. Any student who identified him or herself as a sibling to another adolescent in grades 7-12 on the in-school questionnaire was included in the in-home genetic supplement. In addition, previously unreported twins in grades 7-12 who were discovered during the in-home interview were added at that time. These survey procedures lead to a sample of 2,631 unique sibling pairs from 2,213 families. The genetic supplement to Add Health identifies monozygotic and dizygotic twins, other siblings of twins, other full siblings, half-siblings, and non-related sibling-like pairs (i.e., adolescents living in the same household who did not share the same biological mother or father).

We restrict our sample to only contain pairs comprised of monozygotic and dizygotic twins, full siblings, and half-siblings. In addition, we remove pairs where one or both of the siblings reported outlier values for BMI (< 15 or > 39) or reported a pregnancy anytime in 1994 and 1995. The final sample is also restricted to those pairs with complete information on the model variables. In the sibling model of all pairs, we include twins whose zygosity is undetermined, but these twins are omitted from the multiple groups models. In the final sample, there are 149 unique pairs of monozygotic twins, 240 dizygotic twin pairs, 593 full sibling pairs,

158 half sibling pairs, and 44 twins of undetermined zygosity. Because our analysis is not concerned with the particular ordering of siblings within a family or the pair, we double-enter the data to make the data for Sibling 1 and Sibling 2 symmetric. As a result of this data symmetry, the latent family BMI will be constrained to load equally onto each sibling's latent BMI. In calculating the variance-covariance matrix needed for the modeling exercise, we weight the sibling pair data so that each family has a weight of 1 to insure that the results are not influenced by family size.

Measures

Table 1 provides the descriptions of the measured variables used in our analysis. The dependent variable in our analysis, weight, is measured as the continuous measure of body mass index (BMI). BMI is defined as weight in kilograms divided by the square of height in meters (kg/m^2). Researchers prefer BMI as the standard means to measure adolescent weight (see Cole 1991, Himes & Dietz, 1994, WHO, 1995). Despite the emphasis on qualitative categories of "overweight" and "obese," we treat BMI as a continuous measure because we are interested in the full range of weight.

In both Wave 1 and Wave 2 of Add Health, respondents were asked to report their height and weight, which we use to calculate BMI. In Wave 2, the Add Health study also took physical measures of respondents' height and weight. Adolescents were weighed with clothes but not shoes on a spring scale brought by the interviewer. The interviewer measured adolescents' height in feet and inches using a seamstress-type tape measure. Interviewers were trained in the methods of obtaining both height and weight according to the protocol developed by Add Health staff (Goodman, Hinden & Khandelwal, 2000). Recent research by Goodman and colleagues (2000) using the Add Health survey indicates that while reported weight and height is not a

perfect indicator of the true values, it is generally quite close to their measured height and weight. In the models, the sibling's latent BMI loads onto their Wave 2 physical measurements.

We transform the three measures of BMI for the models because our modeling strategy requires that the dependent variables have a normal distribution. The transformation results in a variable whose distribution is not significantly different than a normal curve using the Shapiro-Wilk *W* test (Shapiro & Wilk, 1965). Then, to avoid confounding differences in BMI with differences in physical maturation, we regress BMI on the Add Health's four measures of physical maturity obtained in Wave 1 separately for boys and girls. For boys, the measures of pubertal development are based on four questions regarding facial hair, underarm hair, lowered voice, and general physical development. For girls, these developmental measures are based on questions about breast development, body curvature, onset of menses, and general physical development. Our sibling models use the residual BMI's from these regressions as our indicators of adolescent weight. It is interesting to note that after residualizing BMI on these developmental measures, the bivariate association of gender and BMI is not statistically significant in our data.

We include several family background characteristics in our analysis that we hypothesize are related to adolescent weight. For the latent family background constructs measured with continuous indicators, we use both siblings' reports to capture the underlying, true value for these variables, but for those constructs measured with dichotomous categories, we use Sibling 1's report of the variable. Given the symmetry in the data, all siblings contribute information on these variables for their family. The family background variables can be categorized as reflecting six domains: (1) social status, (2) the parents' own weight, (3) the child's early life

conditions, (4) mealtime behavior, (5) physical activity and inactivity, and (6) neighborhood conditions related to physical activity.

To capture the family's status, we include three latent constructs. First, the race and ethnicity of the family is categorized into non-Hispanic white, non-Hispanic black, Hispanic, non-Hispanic "other" race (including Asians and Native Americans), and pairs where siblings do not share the same racial or ethnic identity. In the models, the omitted category is non-Hispanic white pairs. Second, we include the parents' educational attainment from the parental report of their highest level of schooling. For adolescents with reports for both maternal and paternal education, we average these values to arrive at our final measure of parental education, but for those with valid data for only one parent, we use that one report for parents' education. We transformed the final value of parental education so that its distribution across the individuals in the final sample would approximate a normal curve. Third, we incorporate a measure of economic vulnerability based on the parent's response to the following question: "Do you have enough money to pay your bills?" Those responding "no" are coded as equal to 1.

To reflect the intergenerational similarity of weight, we include measures of parental obesity. The parental respondent was asked if the biological mother has obesity and if the biological father has obesity. We created two categorical indicators from these questions: (1) both parents are obese and (2) one parent is obese. The omitted category is neither parent is reported as obese.

Next, to account for the child's early life experiences that could contribute to their body composition during adolescence, we include two indicators - birthweight and duration of breastfeeding. The parent respondent provided the information for both variables. For birthweight, the parent reported the child's weight at birth in pounds and ounces. The models use

birthweight in pounds with the ounces converted to decimals. For breastfeeding, the parent could indicate the number of months the child breastfed from among seven choices (never, < 3 months, 3-5 months, 6-8 months, 9-11 months, 12-23 months, and ≥ 24 months). We have recoded the bracketed categories to the average value and use 26 for the value of the final category.

We attempt to account for the child's energy intake by including measures of their mealtime behavior. In Wave 1 of Add Health, the survey asked adolescents: "What do you usually have for breakfast on a weekday morning?" The respondents could check mark a number of items, including that they typically ate nothing or snack foods for breakfast. We include these two indicators as categorical variables if either of these responses were marked. In addition, we include the adolescent's response to the Wave 2 survey question: "In the last seven days, on how many days did you eat breakfast?" Finally, we include the adolescent's response to the Wave 1 survey question: "On how many of the past 7 days was at least one of your parents in the room with you while you ate your evening meal?" We acknowledge that these mealtime behaviors do not map onto the nutritional quality of their food, but these are family-level social processes that influence the adolescent's nutritional intake and, thereby, their weight.[1]

Energy expenditure is also important for weight (Andersen, et al., 1998; Dietz & Gortmaker, 1985) and family-level processes could be important for establishing children's levels of activity and inactivity. To capture moderate to vigorous physical activity, we have used information from Wave 1 survey questions about the total number of times during the past week

[1] In initial analyses, we tried to include more direct measures of nutritional intake by using Wave 1 responses to questions about the number of servings of meat, dairy foods, fruit, vegetables, starches, and sweets eaten yesterday, but these factors did not work well in the models. Also, we were advised that the Wave 2 nutrition variables were of poor quality (Barry Popkin, personal communication, 2/5/04).

the adolescent (1) played "an active sport, such as baseball, softball, basketball, soccer, swimming, or football, " (2) went "roller-blading, roller-skating, skate-boarding, or bicycling," or (3) did "exercise, such as jogging, walking, karate, jumping rope, gymnastics, or dancing" (Gordon-Larsen, McMurray, & Popkin, 2000). The available response options were four categories, which we recoded as follows: "not at all" recoded as 0, "1 or 2 times" recoded as 1, "3 or 4 times" recoded as 2, and "5 or more times" recoded as 3. To arrive at our final measure of moderate to vigorous physical activity, we summed the recoded values across the three survey items. This recoding is problematic, but increases in the final measure of activity should lead to reductions in fatness and, possibly, increases in muscle mass. Therefore, the anticipated effect on BMI of this poorly measured variable is unclear. To capture inactivity, we use responses to survey items in Wave 1 that elicited the number of hours a week spent watching television, watching videos, and playing video or computer games. Fortunately the responses are on a continuous scale and our final measure totals the number of hours across these three activities. Increases in inactivity are expected to increase weight (Gordon-Larsen, et al., 2000). Finally, we include a categorical indicator for whether the adolescent reported at Wave 1 that they played a sport with her mother or father during the last four weeks.

Neighborhood features of the family's environment could also have important consequences for adolescent weight. We include two indicators of relevant neighborhood features. First, we include a categorical variable, equal to 1, if either the adolescent or the interviewer reports that the neighborhood is not safe in Wave 1.[2] We expect safer neighborhoods to foster more physical activity outside the home. Second, we include a categorical variable,

[2] The adolescent was asked the following question: "Do you usually feel safe in your neighborhood?" The interviewer responded to the question: "On your way to the respondent's

equal to 1, if the adolescent respondents affirmatively to the question: "Do you use a physical fitness or recreational center in your neighborhood?" Access and use of a neighborhood recreational center, like activity, could increase muscle mass, but decrease fatness, leading to an ambiguous expectation for adolescent weight.

All of the aforementioned variables are expected to capture family-level processes that vary across families, but not significantly within families, whereas the following variables are expected to vary within a family or sibling pair. First, we include variables indicating the child's gender, equal to 1 if female. Second, we include the respondent's age at Wave 1, divided by 10 to rescale the variable for reasonable parameter estimates. Third, we include a categorical variable based on whether the respondent identifies herself as a "regular smoker," defined in the survey item as smoking one cigarette a day for at least 30 days in a row. Fourth, we include an indicator of disability if the adolescent or her parent identifies her as disabled in the Wave 1 at-home survey. Finally, we include a final measure of food-related behavior based on the Wave 2 survey question asking respondents the following: "In the last seven days, on how many days did you eat at a fast food type place - McDonalds, Kentucky Fried Chicken, Pizza Hut, Taco Bell, etc?" We decided to include this variable as an individual-level variable because we assume that eating fast food during adolescence is more related to peer group activities than family-specific processes, but we acknowledge that parental preference and parental control could influence their eating fast food and other peer-related activities.[3] Together, this final set of variables contributes to within-family differences in BMI.

---

home, did you feel concerned for your safety?" We use information from both respondents to create the indicator for neighborhood safety.

[3] The intersection of peer group influences and family influences is an especially interesting one during the period of adolescence. Future research could theoretically consider and appropriately model the dynamics between these domains for children's well-being.

RESULTS

Models for all sibling pair types

To help demonstrate the relative advantage of estimating structural equation sibling models, we first estimate the model depicted in Figure 1 and labeled Model 0 in Table 2. It is an individual-level MIMIC model for individuals in the final data set. In this model, the independent variables influence a latent construct of BMI, which is measured by the three, residualized measures of BMI. Using BIC as the measure of model fit (Raftery, 1995), we find that the BIC for this model is –276.8.[4]

Next, we estimate a sibling model, Model 1, that is similar to that pictured in Figure 2, but the only family-level factors included in the model are the racial and ethnic categories. All of the within-family variables are included in Model 1. Even with these limited predictors of BMI, Model 1 is a significantly better fitting model, in part because this model accounts for the unmeasured characteristics of families and individuals significant for adolescent weight. By accounting for the within-family resemblance in BMI, the sibling model is a better model than the individual-level model. Model 2 includes all of the variables in the sibling model, as displayed in Figure 2. The addition of the other family-level variables leads to a dramatic improvement in model fit. The family-level processes captured through these family characteristics are, collectively, important for understanding adolescent weight.

Models 3, 4, and 5 test for improvements in model fit as a result of changes in the measurement portion of the model. In Model 3, we allow all Wave 2 measurement errors for BMI to be freely correlated as an occurrence-specific correlation in error. This additional specification leads to a significant improvement in model fit for the additional 6 degrees of

freedom used. Model 4 tests whether the sibling's latent BMI loads equally onto both Wave 2 measures. Setting these measurement paths to be equal actually leads to an improvement, not a loss, in model fit. Finally, Model 5 tests whether the family-level constructs for parental education, birthweight, breastfeeding, and inactivity load equally onto each sibling's measures of these latent variables. In the previous models, the measurement paths to the sibling 2 indicators were close to 1.0. This equality also does not result in a loss in model fit. Therefore, Model 5 is our final and preferred model.

Table 3 presents selected parameters for our final model. The top portion of the table provides the estimates of the effects of the exogenous family-level variables, while the bottom portion provides the coefficients for the exogenous individual-level variables. Beginning with the family-level variables, the parameters for social status prove significant. First, blacks and Hispanics are heavier than whites after controlling for observed and unobserved family characteristics and net of differences in pubertal development. On average, a black adolescent has a BMI that is 1.07 greater than a white adolescent or heavier by a factor of 0.018 standard deviations in BMI.[5] A Hispanic adolescent has a BMI that is on average 1.11 points greater than whites, or 0.021 standard deviations in BMI. Adolescents with other racial identities and those who live in mixed race families are not significantly different than non-Hispanic whites. Second, increases in parental education are associated with declines in the adolescent's BMI. One standard deviation increase in parental education is associated with a 0.08 standard deviation

---

[4] In this formulation for BIC, lower values are better and differences greater than 10 are very significant.

[5] The exponentiated parameter is in the original metric of BMI because the model measure of BMI is logged. In addition, the standardized parameter estimate allows for the comparison of estimates across exogenous variables with different metrics. The "completely standardized solution" in LISREL standardizes both the observed and unobserved variables so the

decrease in the child's BMI.[6] Despite this small effect, the significance of parental education indicates an important behavioral or informational component of family resources for adolescent weight. Finally, economic vulnerability, measured by whether or not the parents can pay their bills, is associated with elevated BMI in children. Economically vulnerable adolescents have BMIs that are 1.05 points, or 0.09 standard deviations, greater than those with more economic security. The difference by economic vulnerability, or poverty, is similar in magnitude to the estimated differences between whites and blacks.

The evidence for the effects of early life conditions are mixed. On the one hand, differences in birthweight prove significant. Heavier babies are likely to become heavier adolescents. A standard deviation increase in birthweight is associated with a 0.13 standard deviation increase in BMI during adolescence. On the other hand, breastfeeding duration is not associated with adolescent weight.

As one might expect, having one or both parents with obesity is associated with higher values of BMI. These coefficients are very large, indicating that having one obese parent is associated with a 1.17 point increase in BMI and having two obese parents is associated with a 1.23 point increase. These parameters are two of the largest estimated effects in the model. The interpretation of parental obesity, however, is not clear in this model because it captures both genetic and behavioral characteristics of the parents, as well as other shared features of the family's environment.

---

standardized estimate relates a standard deviation change in the exogenous variable to a standard deviation in BMI.

[6] Recall that for the analysis, BMI is transformed using natural logs. Therefore, the relationship between the continuous independent variables, i.e. parental education, and BMI is not perfectly linear.

The family-level constructs related to mealtime behavior suggest that forgoing meals at home leads to lower values of BMI. The constructs based on Wave 1 categorical measures of weekday breakfast behavior are not significant, but the quantitative measures of the number of days the adolescent ate breakfast (Wave 2) and the number of days they ate dinner with their parents (Wave 1) are significant. In fact, the strength of the coefficient for the number of days the adolescent ate breakfast at home is one of the strongest predictors in the model. Initially, this substantive conclusion is illogical, but given that BMI is measured on a continuous scale, eating meals at home may not lead to overweight in adolescents but protect against unhealthy low weight.

The adolescent's level of inactivity is an important predictor of BMI. As the number of hours spent watching television, watching a video, or playing computer or video games increases, the adolescent's BMI also increases. This finding confirms previous research. Moderate to vigorous activity, as it is measured in this model, is not significantly related to BMI. On the other hand, those adolescents who play sports with their parents have weights that are lower than their sidelined peers. This suggests a direct mechanism and potential intervention for helping reduce obesity prevalence. Increasing the family's level of activity can help the child, and probably the parent, eliminate excess pounds. It is important to note, however, that this construct could be tapping into extra-familial physical activity levels among the adolescents. Children who play sports with their parents during the high school years are likely to play sports with their peers or at school as well. And, in fact, those who play sports with their parents could be the most active youth. Although the present model does not account for this confounding, the findings suggest a statistically significant role for physical activity for weight management.

Finally, the two measures of neighborhood characteristics are not statistically significant for adolescent weight. Although this could suggest a relatively weak role of neighborhood environment for adolescent weight, this could also result from measurement error or a misspecification of neighborhood features significant for weight.

Most of the individual-level factors are significant for BMI, but gender is not. This finding, however, is purely a result of residualizing BMI on pubertal development. Recall that even in the bivariate association, gender was not significant. Age, on the other hand, is significant. Because weight has been residualized on Wave 1 indicators of pubertal development, age captures both trends in an individual's BMI, but also age-graded social processes that influence weight. For each additional year, adolescent BMI increases by approximately 0.14 points.[7] After controlling for other factors, being a smoker is not significantly associated with BMI. Being disabled is associated with a lower BMI, but the magnitude of this parameter is quite small. Finally, contrary to expectations, BMI purportedly decreases as the number of days eating fast food increases. Upon closer inspection, however, this finding reflects a U-shaped curve in the association of BMI and days of fast food where adolescents reportedly eating fast food between 2 and 5 times a week have the highest BMIs and those at the bottom and top of the distribution have lower BMIs. Future research should determine whether this finding is an artifact of the data, a function of model misspecification, or measurement error because it is contrary to popular wisdom and nutrition expert findings. Recently published research indicates that as fast food intake increases, nutritional quality deteriorates and, presumably, BMI increases (Bowman, et al., 2004).

[7] Recall that age has been divided by 10 and BMI has been logged in this analysis. Our interpretation uses the original metrics of both variables, but the modeled relationship between age and BMI is not perfectly linear.

Given the specification of this model, we cannot identify measurement error in the individual-level portion of the model. But we can investigate the measurement properties of the family-level variables with continuous indicators. The reliabilities for the measured variables in the final model, Model 5, are reported in Table 4. The BMI indicators are very good indicators, as are the measures of parental education. The other constructs based on parental reports, birthweight and breastfeeding duration, are also measured well. The meal behavior variables (days ate breakfast and days ate dinner with parents) and the measures of inactivity are weak. For these measures over half of the latent constructs' effects are based on measurement error. As we anticipated, the measures of moderate to vigorous physical activity are miserable. We need to find better ways to measure physical activity because the present formulation is unacceptable. Measurement error alone is contributing to the lack of a significant association between BMI and physical activity.

Table 5 details the between- and within-family components of the total variance in BMI. The between-family variance of BMI is 0.024, while the within-family variance is 0.057. Thus, the within-family variance in BMI is 70% [0.057/(0.057+0.024)] of the total variance and only 30% of the variance in BMI lies between families. Some of the within-family variance can be accounted for by the genetic similarity of the siblings. Therefore, our analysis by sibling pair type will help explain some of the variation within families.

The key advantage of our analysis is our ability to control for observed and unobserved differences between families that could account for observed differences in BMI. Our model explains about 29% of the variation between families. This is a 12-point increase in the percent of variation in BMI explained across families over our previous sibling model research (Martin, Sandefur, & Meier, 2002) because we added the family-level measures for early life conditions,

mealtime behavior, activity and inactivity, and neighborhood characteristics. In comparison, however, our individual-level variables explain approximately 45% of the variation within families. Even without accounting for the relatedness of different sibling pairs, we do a better job at modeling the individual-level factors important for weight. Together these results bolster our support for the sibling model framework. Although we cannot explicitly account for over half of the variation within-families and the majority of the between-family variation, this unexplained variation is included in the estimation of the parameters for the family- and individual-level factors. In addition, these estimates of unexplained variation encourage us to further develop our theoretical conceptualizations about the relevant factors and processes important for adolescent weight.

Models by sibling pair type

To estimate separate sibling models by pair type, we have to make several adjustments to our model. First and foremost, we have to reduce the number of constructs we investigate to have fewer parameters than we have degrees of freedom in the smallest pair type, the monozygotic twins. Second, because the monozygotic twins frequently have the same values for various measures, we cannot estimate measurement models for the family-level constructs in the multiple group models and compare the parameters across all pair types.[8] Given the results in Table 4, the omission of a measurement model will be especially problematic for the parameter estimates for mealtime behavior, physical activity, and inactivity. Our multiple groups baseline model is restricted to that which is portrayed in Figure 3.

Table 6 presents the fit statistics for the multiple group models. By comparing Model 0 in Table 6 with Model 5 in Table 2, we find that even with the reduction in latent constructs and

the elimination of the measurement error, Model 0 is a good model. Merely accounting for the relatedness of the sibling pairs improves the model fit over our final model of sibling pairs. In Model 1 of Table 6, we restrict the loadings of each sibling's latent BMI so that it loads equally onto both Wave 2 measures. This results in an improvement in fit.

Our primary interest in the multiple group model is to test for the equality of parameters across pair types. But before we can test whether the structural parameters are equal, we must ensure that the measurement properties, where included in the model, are equal across groups. Model 2 restricts the factor loadings of each sibling's latent BMI to be equal across groups. This restriction actually improves model fit. Next, Model 3 adds the restriction that the measurement error in the observed BMI variables is equal across groups. This, too, leads to an improvement in fit. Therefore, the measurement properties are equal across groups.

The first test of the structural parameters, found in Model 4, estimates whether the covariance of the exogenous latent variables, ξ's, in the matrix Φ are equal across groups. This restriction is soundly rejected. Therefore, the covariance of these latent variables will be estimated freely in the remaining models. Model 5 tests whether the effects of the latent exogenous variables are equal across pair types. In essence, this model tests whether the family characteristics included in the model work differently across sibling-pair types. For example, does parental education have a similar association with adolescent BMI among monozygotic twins as it does for full siblings? Restricting these paths to be invariant across groups does lead to a worse fitting model; instead, model fit improves. Therefore, these family processes, as they are measured, work similarly across the pair types included.

---

[8] In structural equation terms, this problem results in a nonpositive definite matrix because some of the covariances in the data matrix are linear transformations of each other.

The remaining model fitting tests involve the variance of the latent endogenous variables. Given the differences in genetic similarity across the pair types, we would expect that the within-family variance to be the smallest for the monzygotic twins and the largest for the half-siblings. This genetic relatedness, however, should not lead to between family differences. Instead, if the pair types are in a random sample of families, then one would generally expect that the between family variance is invariant across pair type. The first model constraining the pattern of within and between variances across pair types tests the highly implausible restriction that the within and between family variances are equal across all pair types. As expected, this model, Model 6, is rejected. Model 7 tests whether the between family variance is equal across all pair types. Surprisingly, this model is rejected. Upon examination of the variance in η_1 across the pairs, one finds that the between-family variation for monozygotic twins is much larger than the between-family variation in the other pair types. We will return to a discussion of this difference later, but given this observation, the next model tests that the between family variance is equal for the dizygotic twins and the full siblings. This equality constraint is upheld. Model 9 then tests whether the between family variance is also equal for the half siblings. Again, this equality constraint is upheld.

The next tests examine the differences in variability within family types. Here, the relatedness of the pairs should lead to significant differences in the within-family variances. Behavioral geneticists would predict that the within family variation in BMI would be the same for dizygotic twins and full siblings, but smaller for monozygotic twins and larger for half siblings. But if families and other social actors treat twins more similarly than they do non-twins, then one could expect the within family variation for dizygotic twins to lie somewhere between that for monozygotic twins and full siblings. Therefore, by examining the unexplained

variation within families across pair types, we can provide a test of behavioral geneticist assumptions with regard to BMI. Model 10 tests whether the within family variance parameters are equal for dizygotic twins and full siblings. These restrictions are not rejected; instead, the behavioral genetic assumption holds for the within family differences for these two pair types. In the final model, Model 11, we test whether the within family variance of half siblings is equal to the parallel within family variance parameters for full siblings and dizygotic twins. Traditional behavioral geneticist assumptions would expect this test to fail, but we are unable to detect a significant difference in these within family variance estimates. Model 11 is our best-fitting model for the multiple groups case.

Table 7 presents the parameter estimates for the best-fitting model. And although these estimates are confounded with measurement error, these parameter estimates are typically smaller than those estimated across all sibling pair types. Therefore, in a model of all sibling pairs, we arrived at upwardly biased estimates for the true effects of the family background characteristics because we did not account for the genetic resemblance of the siblings. There are a two notable exceptions, however, to this general pattern. The parameters for parental obesity are actually larger in the multiple groups model. This suggests that net of the genetic similarities within families, having an obese parent leads to significant increases in BMI. Thus, the intergenerational association of obesity is not completely genetic. We need to fully consider the social processes contributing to both the parents and the child's weight. Finally, the estimated difference in BMI between blacks and whites is also larger in the multiple groups model, but this parameter change also reflects the fact that the omitted category for race and ethnicity changed in this model. In the multiple groups model, the omitted category not only includes whites, but also those classified as "other" and those in mixed race pairs.

Based on our best-fitting multiple group model, we can examine the explained variation within and between families for the monozygotic twins and the other three pair types. This is shown in Table 8. First, we can see the larger between-family variance for the monozygotic pairs (0.061) relative to the other pair types (0.020). One potential explanation for the between-family difference is that the sample of monozygotic pairs in the data are not a random sample of families with monozygotic pairs. This could result from the following two sources: (1) the listwise deletion created a systematic bias in the data for these pairs, or (2) the data collection efforts of Add Health for monozygotic pairs were not random. We need to carefully examine the potential biases introduced from the listwise deletion to help identify whether this is the source of the problem. In addition, alternative methods for handling the missing data could be utilized. The second option is not unreasonable, though, because Add Health included any and all twin pairs found at the in-home survey. Given the extraordinarily high level of between family variation for the monozygotic twins, the model explains very little of this variation. For the other three pair types, however, the model explains one quarter of the variation. The differences between the percent explained for these three pair types in Table 8 and the percentage explained across all pairs in Table 5 could be due to the omission of some latent constructs or, more likely, the absence of a measurement model in the multiple groups case.

The within-family variance components also differ between monozygotic twins and the other pair types, but in an unanticipated way. One would expect the within family variation to be smaller for the monozygotic twins, but here we find the opposite. It is not clear why this would be the case, but the sampling strategies of Add Health for the genetic sample or the bias from listwise deletion could play a role here as well. Despite this unusually high variation within families of monozygotic twins, we explain a significantly higher proportion of the within-family

variation for these pairs than for the other three types by merely controlling for age. On the basis of Table 8, we conclude that our sample of monozygotic twins from Add Health is different than the samples of the other three pair types. In addition, the relatively small size of the monozygotic pairs makes all estimates for this group less stable than that which we have for the other pair types. Similarly, the relatively small number of half siblings limits our ability to detect within-family variance differences between these pairs and the full siblings. We need to further investigate these difference to better understand our findings.

Despite the concerns raised by the variance estimates for the monozygotic twins, the multiple groups model provide an important contribution over the analysis across all sibling pair types. Most of the exogenous constructs that are significant in Model 5, the final model, of all sibling pairs are significant in the multiple groups model. The lone exception was for playing sports with a parent. In addition, the parameter estimates for these family-level factors were attenuated in the multiple groups model because the model accounted for the relatedness of the pairs. While the models of all sibling pairs provides better estimates of family-level factors for adolescent weight relative to an individual-level model, the multiple groups model by pair type appropriately bounds these estimates. Most importantly, the multiple groups models confirmed that social factors, in addition to genetic factors, contribute to an intergenerational association of weight and possibly obesity. Although social scientists might not find this result surprising, it is worth demonstrating to help communicate the multiple pathways through which families influence children's outcomes.

DISCUSSION

We employ structural equation sibling resemblance models in an attempt to assess the effects of family environmental and behavioral factors on adolescent weight after controlling for

both the unobserved characteristics of families and the genetic influences on adolescent weight. With the rapid increase in the prevalence of adolescent obesity, we need models that better illuminate the social determinants of adolescent obesity, while recognizing and appropriately modeling the influence of genetics. The proposed research seeks to arrive at such a model for a particular American cohort. Although we cannot explain the rapid increase in the prevalence of adolescent obesity, our research does provide better estimates of the effects of family background characteristics on adolescent weight. Relative to the behavioral genetics models, this research provides an alternative and, we would argue, improved method for analyzing genetic and environmental factors influencing adolescent weight. Finally, by recognizing the social relationships within the Add Health data, we can better model and understand how families and social contexts facilitate or hinder healthy characteristics in children.

We have several proposals for furthering this research effort. First, we will investigate the biases introduced with listwise deletion and consider alternative methods for handling missing data. Second, we will investigate various ways to have better measures of physical activity, inactivity, and food-related behavior. Third, we will test whether physical activity, inactivity, and food-related behaviors are better modeled as within-family processes rather than as family-level processes. Although we expect physical activity, inactivity, and food-related behaviors to vary both within and between families, we need to test what the appropriate specification of these processes is for these data. Fourth, we hope to investigate similar sibling models predicting the qualitative categories of "overweight" or "obese" using Mplus software. Modeling these categories of weight would not only better speak to policy concerns, but also allow us to determine the meaning of the estimated effects for eating meals at home and playing sports with parents since these constructs lead to increases in the continuous measures of BMI.

These family processes, however, may not lead to an increased probability of being overweight or obese. With these additional steps, we hope to continue improving our model of the family-level processes important for this genetically-determined outcome.

REFERENCES

Ainsworth, Barabara, Haskell, William, Leon, Arthur, Jacobs, David R Jr., Montoye, Henry, Sallis, James F. and Paffenbarger, Ralph S. Jr. 1993. "Compendium of Physical Activities: classification of energy costs of human physical activities*." Medicine and Science in Sports and Exercise* 25: 71-80.

Andersen, Ross E., Carlos J. Crespo, Susan J. Bartlett, Lawrence J. Cheskin, and Michael Pratt. 1998. "Relationship of Physical Activity and Television Watching With Body Weight and Level of Fatness Among Children: Results from the Third national Health and Nutrition Examination Survey." *Journal of the American Medical Association*, 279: 12: 938-942.

Barlow, Sarah E. and William H. Dietz. 1998. "Obesity Evaluation and Treatment:

Expert Committee Recommendations." Pediatrics 102: p. e29.

Calle, E.E., M.J. Thun, J.M Petrelli, C. Rodriguez and C.W. Heath, Jr. 1999. "Body-Mass Index and Mortality in a Prospective Cohort of U.S. Adults." *The New England Journal of Medicine* 341: 1097-1105.

Centers for Disease Control and Prevention (CDC). October 3, 2000. "Body Mass Index-for-Age." <http://www.cdc.gov/nccdphp/dnpa/bmi/bmi-for-age.htm>

Cole, T. J. 1991. "Weight-stature indices to measure underweight, overweight and obesity." In *Anthropometic Assessment of Nutritional Status* (Himes, J.H., ed.) pp. 83-112. Wiley and Sons, New York, NY.

Comuzzie, Anthony G. and David B. Allison. 1998. "The Search for Human Obesity Genes." Science, 280: 5368: 1374-1377.

Dietz, William H. 1998. "Health Consequences of Obesity in Youth: Childhood Predictors of Adult Disease." *Pediatrics*, 101: 3 (Supplement): 518-525.

_____. 1990. "You are What you Eat—What You Eat is What You Are," *Journal of Adolescent Health Care* 11: 76-81.

Goodman, Elizabeth, Beth R. Hinden, and Seema Khandelwal. 2000. "Accuracy of Teen and Parental Reports of Obesity and Body Mass Index." *Pediatrics*, 106: 52-58.

Gordon-Larsen, Penny, Robert G. McMurray, and Barry M. Popkin. 1999. "Adolescent physical activity and inactivity vary by ethnicity: The National Longitudinal Study of Adolescent Health." *The Journal of Pediatrics*, 135: 5: 310-306.

Gortmaker, Steven L., Aviva Must, James M. Perrin, Arthur M. Sobol, and William H. Dietz. 1993. "Social and Economic Consequences of Overweight in Adolescence and Young Adulthood." *The New England Journal of Medicine*, 329: 1008-1012.

Guo, S.S., A.F. Roche, W.C. Chumlea, JU.D. Gardner, and R.M. Siervogel. 1994. "The predictive value of childhood body mass index values for overweight at age 35." *American Journal of Clinical Nutrition,* 59: 810-819.

Hauser, Robert M. 1991. "Simple Models of Sibling Resemblance: Reply to Gillespie." *American Journal of Sociology*, 97: 1: 206-209.

_____. 1988. "A Note on Two Models of Sibling Resemblance." *American Journal of Sociology*, 93: 6: 1401-1423.

Hauser, Robert M., Jennifer T. Sheridan, and John Robert Warren. 1999. "Socioeconomic Achievements of Siblings in the Life Course: New Findings from the Wisconsin Longitudinal Study." *Research on Aging*, 21: 2: 338-378.

Hauser, Robert M. and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." Pp. 177-298 in Adrian E. Raftery (ed.), *Sociological Methodology*, 27. Cambridge: Basil Blackwell.

Himes, J.H. and Dietz, W. H. 1994. "Guidelines for overweight in adolescent preventative services: recommendations from an expert committee." *American Journal of Clinical Nutrition* 59: 307-316.

Jacobson, Kristen C. and David C. Rowe. 1998. "Genetic and Shared Environmental Influences on Adolescent BMI: Interactions with Race and Sex." *Behavior Genetics*, 28: 265-278.

Kilpatrick, M., C. Ohanessian, and J. B. Bartholomew. 1999. "Adolescent Weight Management and Perceptions: An Analysis of the National Longitudinal Study of Adolescent Health." *Journal of School Health* 69(4): 148-152.

Levinson, Richard, Brian Powell, and Lala Carr Steelman. 1986. "Social Location, Significant Others and Body Image Among Adolescents." *Social Psychological Quarterly*, 49: 4: 330-337.

Martin, Molly A., Ann M. Meier, and Gary D. Sandefur. 2002. "Family Background and Racial and Ethnic Differences in Adolescent Weight." Center for Demography and Ecology Working Paper No. 2002-11. Madison, WI: University of Wisconsin-Madison.

Mossberg, H.O. 1989. "40-year follow-up of overweight children." *Lancet.* 2: 491-493.

Popkin, Barry M. and J.Richard Udry. 1998. "Adolescent Obesity Increases Significantly in Second and Third Generation U.S. Immigrants: The National Longitudinal Study of Adolescent Health." *Journal of Nutrition*, 128: 701-706.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." Pp. 111-73 in Peter V Marsden (ed.), *Sociological Methodology*. Cambridge: Basil Blackwell.

Sallis, James F., Buono, Michael J., Roby, Julia J., Micale, Frank G. and Nelson, Julie A., 1993. "Seven-day recall and other physical activity self-reports in children and adolescents." *Medicine and Science in Sports and Exercise* 25: 99-108.

Shapiro, S.S. and M.B. Wilk. 1965. An analysis of variance test for normality. *Biometrika* 52: 591-611.

Sörbom, Dag and Karl G. Jöreskog. 1981 "The Use of LISREL in Sociological Model Building." Pp. 179-199 in D.J. Jackson and E.F. Borgatta (eds.), Factor Analysis and Measurement in Sociological Research: A Multidimensional Perspective. Beverly Hills: Sage Publications.

Troiano, Richard P. and Katherine M. Flegal. 1998. "Overweight Children and Adolescents: Description, Epidemiology, and Demographics." *Pediatrics*, 101: 3 (Supplement): 497-504.

Whitaker, Robert C., Jeffrey A. Wright, Margaret S. Pepe, Kristy D. Seidel, and William H. Dietz. 1997. "Predicting Obesity in Young Adulthood fro Childhood and Parental Obesity." *The New England Journal of Medicine*, 337: 13: 869-873.

Winkleby, Marilyn A., Thomas N. Robinson, Jan Sundquist, ad Helena Kraemer. 1999. "Ethnic Variation in Cardiovascular Disease Risk Factors Among Children and Young Adults: Findings from the Third National Health and Nutrition Examination Survey, 1988-1994." *Journal of the American Medical Association*, 281: 11: 1006-1013.

World Health Organization Expert Committee. 1995. "Physical Status: The Use and Interpretation of Anthropometry." *WHO Technical Report Series* no. 854, WHO, Geneva, Switzerland.

Figure 1. Individual Model of Adolescent BMI, All Sibling 1s

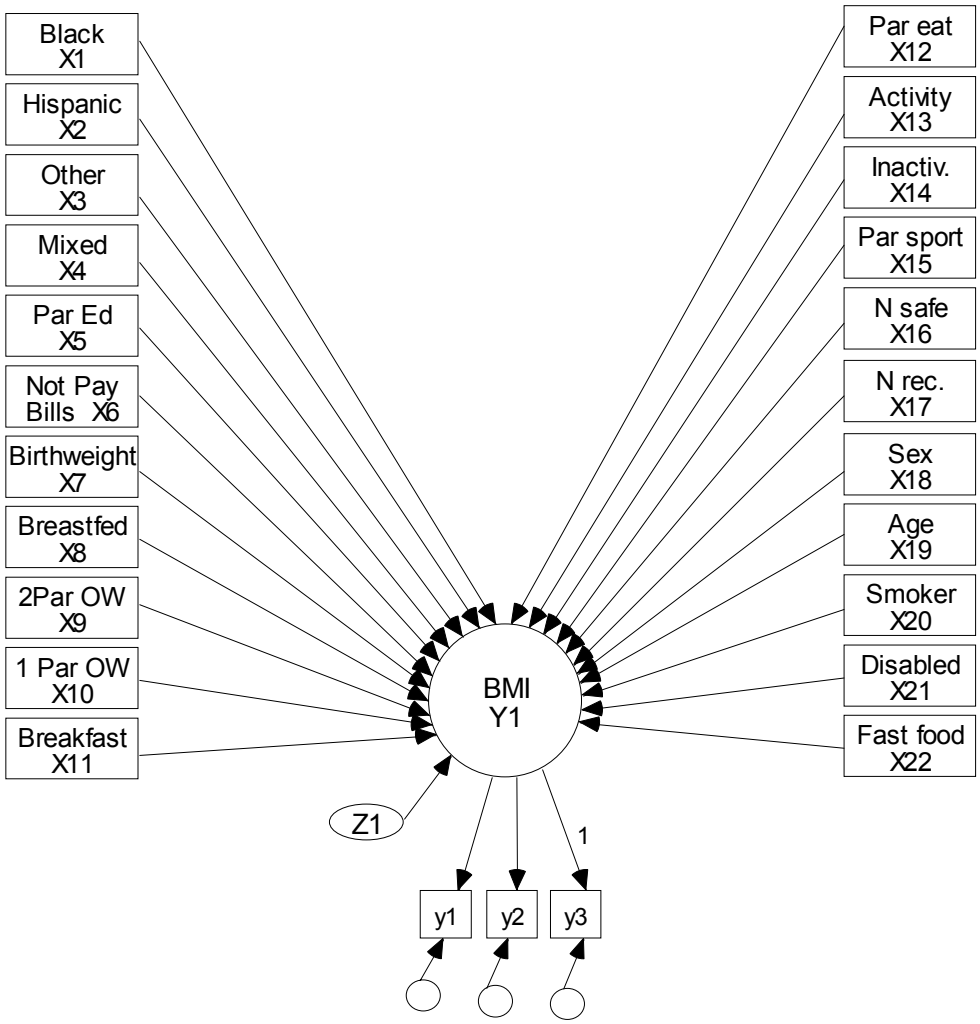
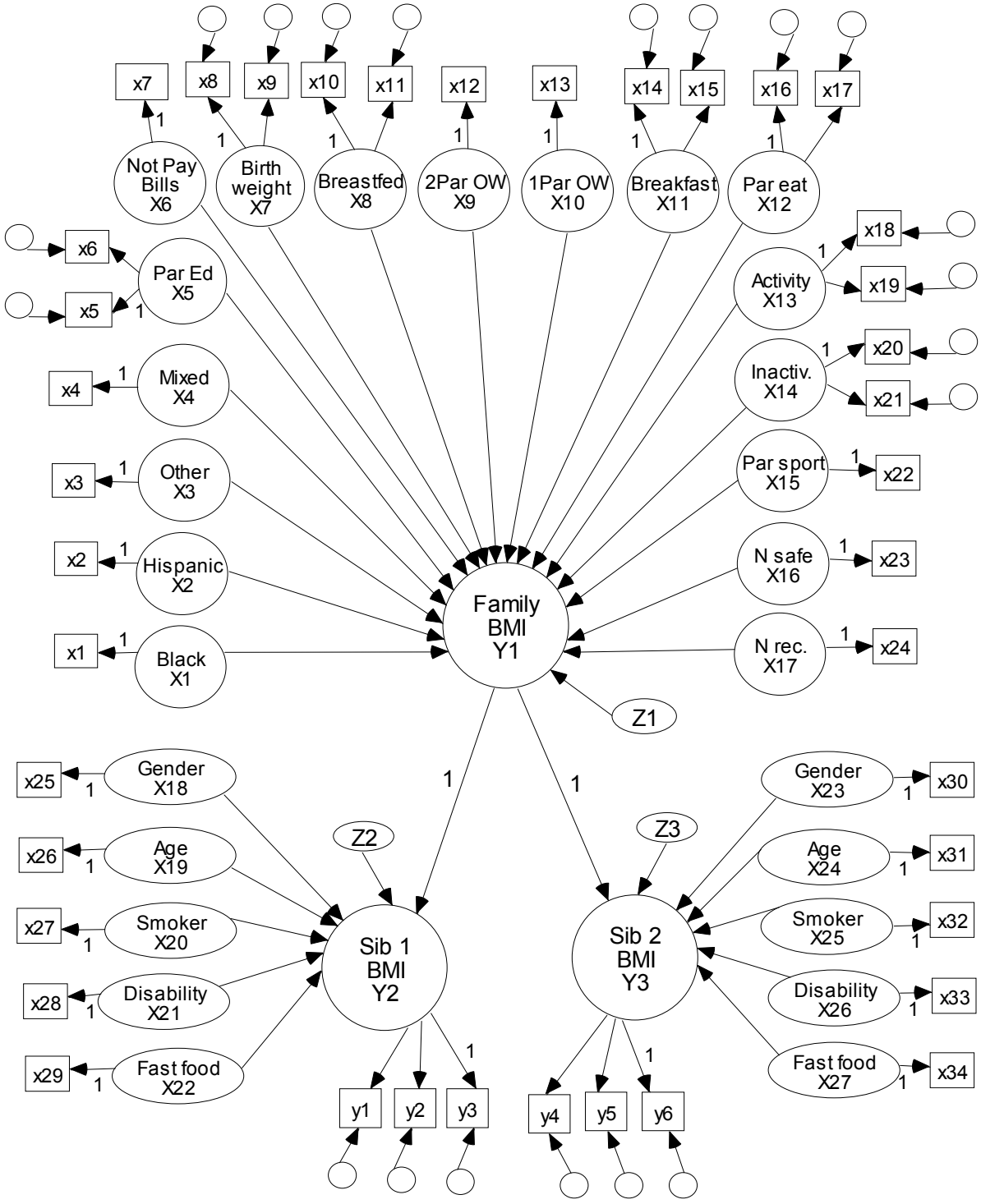
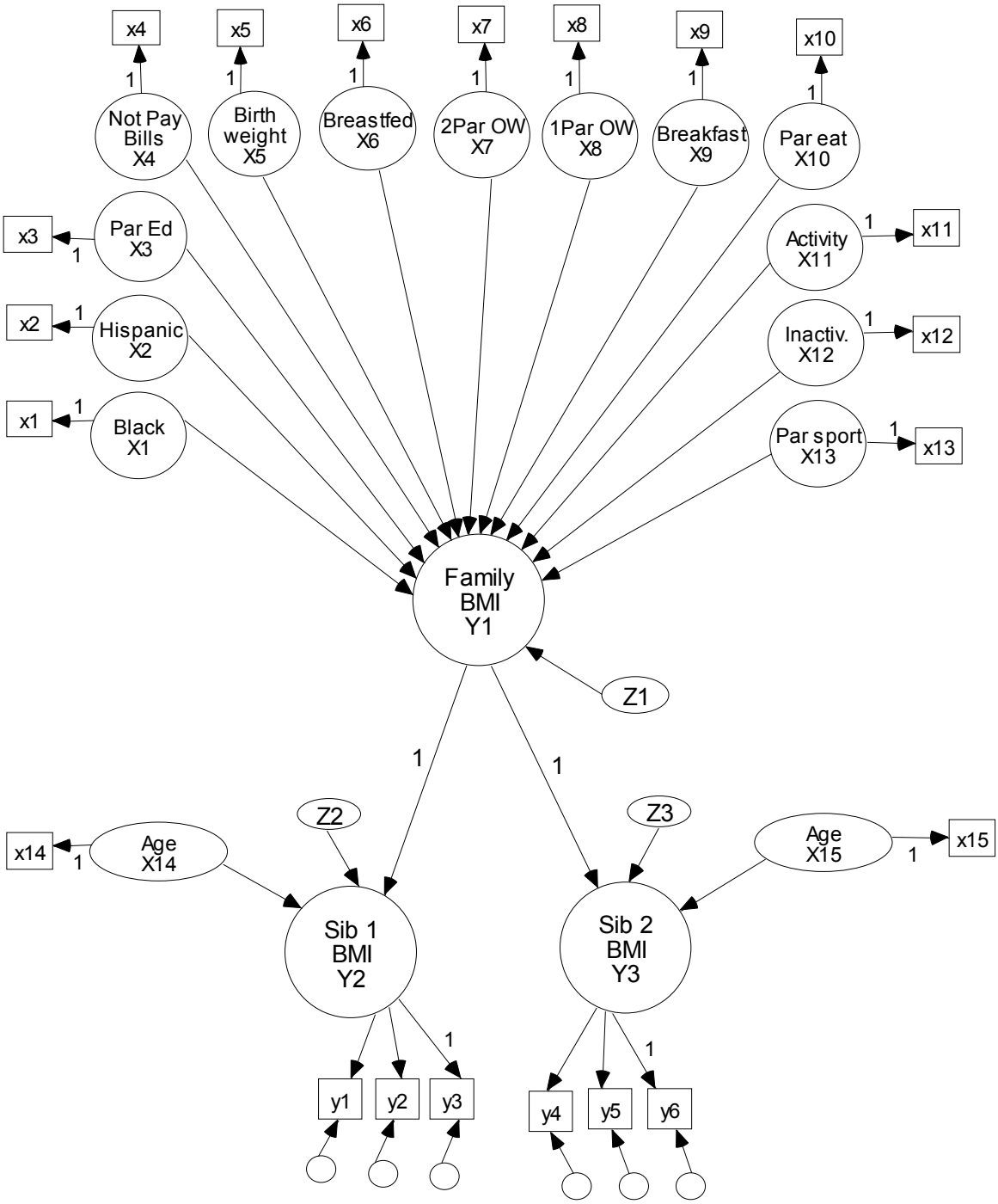


Figure 2. Sibling Structural Equation Model for Adolescent BMI, All Sibling Pairs



Note: X1-X27 are intercorrelated

Figure 3. Sibling Structural Equation Model for Adolescent BMI,Multiple Groups Model



Note: X1-X15 are intercorrelated

Table 1. Variable Descriptions, Transformations, Means, and Standard Deviations

Latent Label	Measured Variable Description	Variable Transformation	Orig. metric Weighted Mean	Orig. metric Weighted St.Dev.	Model metric Unweighted Mean	Model metric Unweighted .Dev.
Dependent Variables for BMI Factors						
S1BMI, S2BMI	BMI from Wave 1 self report, residualized	ln(BMI - 12.81)	22.2	135.0	0.00	0.4
	BMI from Wave 2 self report, residualized	ln(BMI - 13.28)	22.7	138.4	0.00	0.4
	BMI from Wave 2 physical measurements, residualized	ln(BMI - 13.31)	22.7	147.8	0.00	0.4
Independent Variables for Family-level Factors						
BLACK	Pair is non-Hispanic Black (= 1)	---	0.11	-	0.17	-
HISPANIC	Pair is Hispanic (= 1)	---	0.07	-	0.12	-
OTHER	Pair is non-Hispanic, Asian or "other" race (= 1)	---	0.04	-	0.05	-
MIXEDPR	Pair is mixed race/ethnicity (= 1)	---	0.02	-	0.03	-
PARED	Average parents' education in years, Parent report	$(ed^{1.03}-1)/1.03$	13.51	72.6	13.08	2.4
NPAYBILL	Not enough money to pay bills, Parent report		0.20	-	0.20	-
BTHWT	Birthweight in pounds, Parent report	$(birthwt^{1.37}-1)/1.37$	7.25	48.4	9.34	3.2
BRSTFD	Number of months breastfed, Parent report	---	3.62	197.1	2.84	5.0
BOBS	Both biological parents with obesity problems (= 1), Parent report	---	0.06	-	0.06	-
OOBS	One biological parent with obesity problems (= 1), Parent report	---	0.16	-	0.18	-
BK_SNA	Weekday breakfast - snack foods (=1), Wave 1	---	0.07	-	0.07	-
BK_NONE	Weekday breakfast - nothing (=1), Wave 1	---	0.18	-	0.18	-
BKDAYS	Days ate breakfasts last week, Wave 2	---	4.28	87.4	4.28	2.5
PARDIN	Number of dinners with a parent last week, Wave 1	---	4.76	80.5	4.59	2.4
ACTCT	Count of moderate to vigorous physical activities last week, Wave 1	---	4.23	31.6	4.18	0.9
INACCT	Hours spent doing inactive hobbies last week, Wave 1	$(inacct^{.0987}-1)/.0987$	21.90	698.6	3.11	1.7
PARSPT	Played a sport with a parent during last four months (=1), Wave 1	---	0.31	-	0.27	-
NESAFE	Does not usually feel safe in neighborhood (=1), Wave 1	---	0.08	-	0.09	-
NEREC	Uses a neighborhood recreation center, Wave 1	---	0.22	-	0.21	-

(Continued)

Table 1. Variable Descriptions, Transformations, Means, and Standard Deviations (Cont.)

Latent Label	Measured Variable Description	Variable Transformation	Orig. metric Weighted Mean	St.Dev.	Model metric Unweighted Mean	.Dev.
Independent Variables for Individual-level Factors						
SEX	Female (= 1)	---	0.48	-	0.49	-
AGE	Age, Wave 1	age/10	15.52	57.8	1.55	0.2
SMOKER	Regular smoker, self report (= 1), Wave 1	---	0.22	-	0.19	-
DISABL	Physical disability (= 1), Wave 1	---	0.02	-	0.02	-
FASTFD	Days ate fast food last week, Wave 2	---	2.25	62.2	2.28	1.8

Source: Individuals in the final sample of pairs from the National Suvery of Adolescent Health, Waves 1 and 2

a: Weighted statistics are calculated using Wave 2 sample weights.

Table 2: Model Fit Statistics, *N* = 2,324

	Model	df	Minimum Fit L^2	BIC
Individual Model				
0	All Variables	50	110.7	-276.8
Sibling Model				
1	Race & ethnicity family factor + all within family factors	78	226.4	-378.2
2	All variables	405	1253.5	-1885.7
3	2 + θ^{ε} correlated in all Wave 2 BMIs	399	1122.0	-1970.6
4	3 + all Wave 2 λ_ys = 1	403	1140.6	-1983.1
5	4 + λ_x =1 for S2 indicators of PARED, BIRTHWT, BRSTFD, INACCT	407	1140.9	-2013.8

Table 3. Selected Parameters of Final Sibling Model: Model 5

	Parameter	γ	e^{γ} [a]	SE	Standardized Estimate[b]
Effects of Family-level Variables on Family's Latent BMI					
Race/Ethnicity (Omitted: White Pair)					
Black Pair	$\gamma(1,1)$	0.043	1.044	0.021 **	0.069
Hispanic Pair	$\gamma(1,2)$	0.107	1.113	0.020 ***	0.151
Other, Same-group Pair	$\gamma(1,3)$	0.037	1.038	0.031	0.033
Mixed Pair	$\gamma(1,4)$	-0.035	0.966	0.038	-0.024
Parental Education	$\gamma(1,5)$	-0.008	0.992	0.003 ***	-0.079
Cannot pay bills	$\gamma(1,6)$	0.051	1.052	0.016 ***	0.087
Birthweight	$\gamma(1,7)$	0.012	1.012	0.003 ***	0.133
Months breastfed	$\gamma(1,8)$	0.000	1.000	0.002	-0.008
Parents Obese (Omitted: Neither)					
Both parents	$\gamma(1,9)$	0.208	1.231	0.026 ***	0.207
One parent	$\gamma(1,10)$	0.156	1.169	0.016 ***	0.257
Breakfast					
Weekday = snacks, W1	$\gamma(1,11)$	-0.010	0.990	0.024	-0.011
Weekday = nothing, W1	$\gamma(1,12)$	0.020	1.020	0.021	0.033
Days ate breakfast, W2	$\gamma(1,13)$	-0.033	0.968	0.008 ***	-0.239
Meals with parents	$\gamma(1,14)$	0.014	1.014	0.006 **	0.098
Mod.-Vigorous Phys. Activity	$\gamma(1,15)$	0.021	1.021	0.031	0.045
Inactivity	$\gamma(1,16)$	0.050	1.051	0.010 ***	0.220
Played sport with parent	$\gamma(1,17)$	-0.033	0.968	0.018 *	-0.063
Neighborhood safe	$\gamma(1,18)$	-0.009	0.991	0.022	-0.011
Use neighborhood rec. center	$\gamma(1,19)$	0.025	1.025	0.017	0.044
Effects of Individual-level Variables on each Sibling's Latent BMI[c]					
Gender (1=Female)	$\gamma(1,20)$, $\gamma(1,25)$	0.007	1.007	0.014	0.010
Age	$\gamma(1,21)$, $\gamma(1,26)$	0.405	1.499	0.045 ***	0.187
Regular smoker	$\gamma(1,22)$, $\gamma(1,27)$	-0.029	0.971	0.018	-0.032
Disabled	$\gamma(1,23)$, $\gamma(1,28)$	-0.083	0.920	0.045 *	-0.035
Days at fast food, W2	$\gamma(1,24)$, $\gamma(1,29)$	-0.021	0.979	0.004 ***	-0.103

Source: Symmetric sibling pair data from the National Survey of Adolescent Health, Waves 1 & 2

P-value significance: ***: $p < .01$, **: $p < .05$, *: $p < .10$

a: The exponentiated parameter is in the original BMI metric because the model uses a log transformation of BMI.

b: From LISREL's completely standardized solution so that both the latent and observed variables are standardized. Estimates can be interpretted like path coefficients or correlations.

c: Parameter estimates vary very slightly across Sibling 1 and Sibling 2 so the average of the two estimates are presented here. The standard errors are the same for Sibling 1 and Sibling 2.

Table 4. Reliabilities of Measured Variables: Model 5

Latent Label	Measure	
Dependent Variables for BMI Factors		
S1BMI, S2BMI	Transformed BMI reported, wave 1	0.90
	Transformed BMI reported, wave 2	0.82
	Transformed BMI measured, wave 2	0.76
Independent Variables for Family-level Factors [a]		
BLACK	Black Pair (= 1)	1.00
HISPANIC	Hispanic Pair (= 1)	1.00
OTHER	Other, Same-group Pair (= 1)	1.00
MIXEDPR	Mixed Pair (= 1)	1.00
PARED	Parental Education	0.99
NPAYBILL	Cannot pay bills (=1)	1.00
BTHWT	Birthweight	0.67
BRSTFD	Months breastfed	0.77
BOBS	Both parents obese (= 1)	1.00
OOBS	One parent obese (= 1)	1.00
BK_SNA	Breakfast weekday = snacks, W1	1.00
BK_NONE	Breakfast weekday = nothing, W1	1.00
BKDAYS	Days ate Breakfast, W2	
	Sibling 1	0.43
	Sibling 2	0.25
PARDIN	Days meals with parents	
	Sibling 1	0.44
	Sibling 2	0.35
ACTCT	Mod.-Vigorous Phys. Activity	
	Sibling 1	0.28
	Sibling 2	0.08
INACCT	Inactivity	0.36
PARSPT	Played sport with parent (= 1)	1.00
NESAFE	Neighborhood safe (= 1)	1.00
NEREC	Use neighborhood rec. center (= 1)	1.00
Independent Variables for Individual-level Factors [a]		
SEX	Female (= 1)	1.00
AGE	Age	1.00
SMOKER	Regular smoker (= 1)	1.00
DISABL	Disabled (= 1)	1.00
FASTFD	Days at fast food	1.00

a: Reliabilities equal to 1.0 are so by model design.

Table 5. Total and Unexplained Variance across Siblings: Model 5

Between Family Variance: Family BMI	
Total Variation	0.024
Unexplained Variation, $\Psi(1)$	0.017
	(0.001)
% Explained	29.2%
Within Family Variance: S1BMI or S2BMI[a]	
Total Variation	0.057
Unexplained Variation, $\Psi(2)$ or $\Psi(3)$	0.031
	(0.001)
% Explained	45.1%

a: Either S1BMI or S2BMI parameters can be used for estimates of within family variance; here we use an average of the two siblings' parameters.

Table 6: Model Fit Statistics for Multiple Groups Sibling Model, *N* = 2,280

	Model	df	Minimum Fit L^2	BIC
0	Baseline model	332	536.20	-2030.8
1	0 + only $\lambda_y(1,1)$ & $\lambda_y(4,2)$ free	340	559.87	-2069.0
2	1 + λ_ys invariant across groups	346	567.77	-2107.5
3	2 + θ^{ε}s invariant across groups	364	607.94	-2206.5
4	3 + Φs invariant across groups	724	5896.06	298.1
5	3 + γs invariant across groups	409	683.65	-2478.7
6	5 + $\Psi(1,1)$, $\Psi(2,2)$ & $\Psi(3,3)$ invariant across groups	418	880.82	-2351.1
7	5 + $\Psi(1,1)$ invariant across groups	412	811.06	-2374.5
8	5 + $\Psi(DZ,1,1) = \Psi(FS,1,1)$	410	683.65	-2486.4
9	8 + $\Psi(DZ,1,1) = \Psi(FS,1,1) = \Psi(HS,1,1)$	411	685.64	-2492.2
10	9 + $\Psi(DZ,2,2)=\Psi(FS,2,2)$ & $\Psi(DZ,3,3)=\Psi(FS,3,3)$	412	683.69	-2501.9
11	10 + $\Psi(HS) = \Psi(FS) = \Psi(DZ)$	415	687.69	-2521.1

Table 7. Selected Parameters of Final Multiple Groups Sibling Model: Model 11

	Parameter	γ or λ_y	e^{γ} or e^{λ} [a]	SE
Effects of Family-level Variables on Family's Latent BMI				
Race/Ethnicity (Omitted: White Pair)				
Black Pair	$\gamma(1,1)$	0.069	1.072	0.018 ***
Hispanic Pair	$\gamma(1,2)$	0.101	1.106	0.021 ***
Parental Education	$\gamma(1,3)$	-0.008	0.992	0.003 ***
Cannot pay bills	$\gamma(1,4)$	0.042	1.043	0.016 ***
Birthweight	$\gamma(1,5)$	0.005	1.005	0.003 **
Months breastfed	$\gamma(1,6)$	-0.001	0.999	0.001
Parents Overweight (Omitted: Neither)				
Both parents	$\gamma(1,7)$	0.241	1.273	0.027 ***
One parent	$\gamma(1,8)$	0.170	1.185	0.017 ***
Days ate breakfast, W2	$\gamma(1,9)$	-0.011	0.989	0.003 ***
Meals with parents	$\gamma(1,10)$	0.006	1.006	0.003 **
Mod.-Vigorous Phys. Activity	$\gamma(1,11)$	0.004	1.004	0.007
Inactivity	$\gamma(1,12)$	0.016	1.016	0.004 ***
Played sport with parent	$\gamma(1,13)$	-0.020	0.980	0.015
Effects of Individual-level Variables on each Sibling's Latent BMI				
Age	$\gamma(2,14)$	0.334	1.396	0.044 ***
	$\gamma(3,15)$	0.325	1.384	0.043 ***
Loading of Siblings' Latent BMI on Measured BMI				
BMI, Wave 1 reported: S1	$\lambda_y(1,2)$	0.878	2.407	0.011 ***
BMI, Wave 1 reported: S2	$\lambda_y(4,3)$	0.878	2.405	0.011 ***
BMI, Wave 2 reported	$\lambda_y(2,2)$, $\lambda_y(5,3)$	1.000		-
BMI, Wave 2 measured	$\lambda_y(3,2)$, $\lambda_y(6,3)$	1.000		-

Source: Symmetric sibling pair data from the National Survey of Adolescent Health, Waves 1 & 2

P-value significance: ***: $p < .01$, **: $p < .05$, *: $p < .10$

a: The exponentiated parameter is in the original BMI metric because the model uses a log transformation of BMI.

Table 8. Total and Unexplained Variance across Pair Types: Model 11

	MZ Twins	DZ Twins, Full Sibs, Half Sibs[a]
Between Family Variance: Family BMI		
Total Variation	0.061	0.020
Unexplained Variation, Ψ(1)	0.056	0.015
	(0.005)	(0.001)
% Explained	8.8%	25.5%
Within Family Variance: S1BMI or S2BMI[b]		
Total Variation	0.072	0.063
Unexplained Variation, Ψ(2) or Ψ(3)	0.010	0.042
	(0.003)	(0.002)
% Explained	86.5%	33.8%

a: Variances across these groups estimated to be equal.

b: Either S1BMI or S2BMI parameters can be used for estimates of within family variance; here we use an average of the two siblings' parameters.