2015

The Impact of Obesity in Children and Adults on Medical Care Expenditures: A Regression Calibration Instrumental Variables Approach

Adam Ira Biener
Lehigh University

The Impact of Obesity in Children and Adults on Medical Care Expenditures: A Regression Calibration Instrumental Variables Approach

by
Adam Biener

Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy
in
Business and Economics

Lehigh University
September 2015

This thesis is accepted and approved in partial fulfillment of the requirements for the Ph.D in Economics

---

Date: June 15, 2015

Dissertation Advisor: Chad Meyerhoefer

Committee Members:

Muzhe Yang

Irina Panovska

John Cawley

Acknowledgements

I would like to first sincerely acknowledge my advisor, Dr. Chad Meyerhoefer, for all of his patience and guidance in helping me conduct my research and grow in my career. I appreciate all of the time he spent with me discussing this work and for believing in my research. Thank you for always being kind and honest with me, and for pushing me to learn and work hard.

I am also deeply grateful to my committee members: Dr. John Cawley for his kindness and his generosity, as well as for all the opportunities he has afforded me to discuss and develop this research with himself and other esteemed colleagues; Dr. Muhze Yang for his invaluable training in econometrics, and for his very insightful guidence in how to organize and conduct empirical research; and Dr. Irina Panovska for her help in developing the framework and presentation of this dissertation.

I want to independently thank Drs. Meyerhoefer, Cawley and Dearden for their letters of recommendation, and guidance throughout my job search.

I am very thankful to Dr. Nathan Teft, Dr. Jeffery Sobal, and others who discussed this work in conferences or with me directly. I also need to thank Ray Kuntz, for all his help executing my countless programs at AHRQ.

Lastly I want to thank my friends and family for all of their support. Krishna, Rhea, Lizhong, Dante, Cheng, and the other graduate students, thank you for your being a strong support network. Mom, Dad, and Jen, thank you for your constant and ever-present love and support, without which I would never have found the passion to pursue this work. Marlene, You are the love of my life. Thank you for your unending encouragement and for always believing in me.

For my Dad.

Contents

List of Tables

List of Figures

Abstract

These papers are concerned with estimating the marginal effect of weight and obesity on medical expenditures for children and adults when heights and weights are misreported. The first chapter investigates the implications of reporting error and endogenous instruments in GLM and IV-GLM models. We evaluate standard IV estimation along with regression calibrated IV estimators (RCIV). We find that the standard regression calibrated IV model is fundamentally inconsistent, but propose an adjusted RCIV estimator (ARCIV) that corrects these faults, as well as incorporating correction for misreporting. Chapter 2 contains an application of the ARCIV from the first chapter to estimate the marginal effect of changes in BMI of children on their annual medical care costs. After correcting for reporting error and instrumenting, we find that the impact of youth obesity is significantly larger than found in previous research. Chapter 3 updates previous estimates the impact of adult obesity on medical expenditures. Although prior research has used instrumental variables to estimate the cost of adult obesity, controlling for measurement error using our ARCIV estimator shows that the true effect is still larger than found in non-IV models, but smaller than IV models with mismeasured BMIs.

Chapter 1

The Conditional Expectations Instrumental Variables Method for Models with Additive Measurement Error

1.1 Introduction

Estimates of the impact of BMI and obesity on medical expenditures generally rely on survey data. Data drawn from surveys are susceptible to measurement error. The error generating process may be due to lack of precision in recording true information. In the Medical Expenditure Panel Survey, used in subsequent chapters, BMIs are generated using self- and proxy-reported height and weight. Incorrect reporting can introduce measurement error that results in biased estimates of the impact of BMI and obesity on medical expenditures (Bound et al., 2001). A confounding concern in medical expenditure modelling is the endogeneity of BMI. Some earlier studies of the impact of BMI on medical expenditures admit that endogeneity of BMI is a potential limitation, but do not explicitly control for it in their analysis (Finkelstein et al., 2009; Monheit et al., 2009).

Cawley and Meyerhoefer (2012) addressed concerns about endogeneity by using instrumental variables to estimate the impact of BMI on medical expenditures. They used the BMI of biological children to instrument for the BMI of their parents in the MEPS in order to estimate the impact of obesity on adult medical expenditures. They find that after correcting for the endogeneity of BMI, coefficient estimates significantly increase in magnitude, suggesting that endogeneity is responsible for a considerable attenuation bias.

Cawley and Meyerhoefer (2012), and other studies using similar instruments, utilize BMI that is derived from self- and proxy-reported height and weight. IV estimation will be consistent if any reporting error is classical (has mean zero). However, there have been a number of empirical studies of adult self- and proxy-reporting behavior that suggests a general tendency to overstate height and understate weight, which results in under-reports of BMI (Stommel and Schoenborn, 2009; Gorber et al., 2007; Lundahl et al., 2014). The key concern is that under-reporting will generate non-classical error in the endogenous regressor. In one study, O'Neill and Sweetman (2013) describe the bias due to non-classical measurement error analytically for the log-linear model. They demonstrate that the bias due to non-classical error in BMI is determined by the joint distribution of reporting error and true BMIs, and in particular, the covariance between the instrument and the measurement error. They find evidence that under-reporting in BMI can result in upward bias when using instrumental variables to estimate the impact of obesity on income. Cawley and Meyerhoefer (2012) acknowledge that in the MEPS, the endogenous regressor and the instrument are under-reported. Thus IV estimation using the MEPS may be biased due to reporting error in BMI

In subsequent chapters, we use validation data to correct for measurement error in both the endogenous regressor and the instrument. Validation data can be used to generate correction equations whose arguments are present in the main sample. The correction equations are then used to predict the true value for the mismeasured variables in the main sample (Lee and Sepanski, 1995; Bound et al., 2001). Previous studies have used the NHANES as a source of validation data for mismeasured BMI. Cawley (2000)

used the NHANES III to correct for self-reporting error in height and weight in the National Longitudinal Survey of Youth (NLSY) in order to estimate the impact of body weight on employment disability. He uses an IV model, and corrects for reporting error in both the endogenous regressor as well as the instrument. This estimator can be thought of as a regression calibrated instrumental variables (RCIV) estimator, since fitted values conditional on validation data are substituted into the IV model. The same corrections are employed by Cawley (2004) in order to estimate the impact of obesity on wages. These previous studies have used corrected BMI in IV models. However, the subsequent chapters are the first studies to implement corrected BMI for IV estimates of the impact of obesity on medical expenditures.

This chapter updates the descriptions of bias due to reporting error by Bound et al. (2001) and O'Neill and Sweetman (2013) to accommodate the non-linear medical expenditure models used in Cawley and Meyerhoefer (2012), as well as in subsequent chapters. This chapter also evaluates the validation method of Cawley (2000) used to correct for measurement error in both the endogenous regressor and the instrument. We find that this estimator is only approximately consistent, and prone to bias even when the instrument is not measureed with error. We propose an alternative version of the RCIV estimator that are more robust to certain measurement error regimes. We then use Monte Carlo simulations to compare the RCIV estimators to uncorrected and IV estimators in both linear and non-linear contexts. We find the the standard RCIV estimator performs worse than uncorrected IV estimation, although it can be improved to match the performance of the IV.

The contributions of this chapter satisfy the concerns about validation literature described in Bound et al. (2001). In their review of validation methods for measurement error in survey data, they conclude that validation studies need to be more explicit about their assumptions regarding the model errors and their relationship to the true variables in the model. This chapter provides a clear theoretical description of bias under measurement error regimes that are suggested by empirical studies of self- and proxy-reporting behavior with regard to height and weight. We update previous research by describ-

ing possible biases in the situation where the endogeneity of BMI necessitates the use of IV estimation, but both the endogenous regressor and the available instruments are measured with error. Bound et al. (2001) also express concern that validation studies are not of practical use to data analysts and researchers. Their review provides a clear description of bias when the endogenous regressor is mismeasured, and they evaluate the efficacy of different validation procedures in those situations. We frame our findings in terms of the empirical model of the impact of BMI on medical expenditures, and only use available data to implement any validation method. To accommodate the diversity of modeling approaches, we demonstrate the effect of measurement error when estimating both linear and non-linear models of healthcare demand.

1.2 Measurement Error in Linear Models of Medical Expenditures

What follows is a description of the impact of measurement error on various models of healthcare demand. In order to make the conclusions here more portable into other chapters, we will consider the outcome y to be medical expenditures, and the regressor of interest to to be BMI. Thus we can focus on instances that are grounded in the empirical reality of these models. We assume that all models are correctly specified, as our main interest is the role of measurement error. Additional resources regarding bias due to measurement error in linear models can be found in Fuller (1987). Measurement error in non-linear models are discussed in Carroll et al. (2006). Our discussion overlaps somewhat with that in Bound et al. (2001) since we are also discussing models and measurement error that commonly incorporate survey data.

1.2.1 Least-Squares Estimator

Consider a simple linear model in which y is the outcome that depends on X, an Nxk column matrix of data such that $X = [x|Z]$, where x is the regressor of interest, and Z is a matrix of other covariates. To allow for measurement error, X is not directly

observable, and instead only measurements of X can be observed. The true model is,

$$y = X\beta + \varepsilon \tag{1.1}$$

We assume that $E(X\varepsilon) = 0$ so that the true model is correctly specified. We do not observe X, but instead observe measurement W such that,

$$W = X + u. \tag{1.2}$$

We begin by assuming that $Cov(Xu) = 0$ and $Var(u) \neq 0$ in equation 1.2. If $E(u) = 0$, then the measurement error is classical in nature. In this case, measurement error is also said to be nondifferential, as the observed BMI W does not contain any information about the expenditures y not already contained in X.[1]

Give the true model is linear, the best estimator for β is the least squares estimator. The asymptotic bias due to measurement error in equation 1.2 for $k = 1$ can be shown using the probability limit of the least-squares estimator,[2]

$$plim\beta_{OLS} = \frac{Cov(w, y)}{Var(w)} = \frac{\sigma_{wy}}{\sigma_w^2} = \frac{\beta\sigma_x^2}{\sigma_x^2 + \sigma_u^2} \tag{1.3}$$

As the measurement error variance increases, $\hat{\beta}_{OLS}$ attenuates towards zero.[3]

In practice, data drawn from surveys depends on self- or proxy-reports. It may be the case that X is correlated with u. For example, individuals may be more likely to under-report their weight as their own weight increase. In this case we would have nonclassical measurement error such that, $E(Xu) \neq 0$. Nonclassical measurement error would alter the probability limit to be,

$$plim\beta_{OLS} = \frac{\beta(\sigma_x^2 + \sigma_{ux})}{\sigma_x^2 + \sigma_u^2 + 2\sigma_{ux}}. \tag{1.4}$$

[1]Measurement error u in equation 1.2 would be differential if u was correlated with the error term ε. This is analogous to reporting error being an omitted variable in the true model, or W being a replicate measure, or proxy for X. Since we assume the model is correctly specified, we assume that measurement error is nondifferential.

[2]Bound et al. (2001) show this bias in multivariate regression.

[3]If $\sigma_u^2 = 0$, then x is not mismeasured and it is clear that β_{OLS} is a consistent estimator.

The direction of bias depends on the relationship between the measurement error variance σ_u^2 and the covariance between x and the additive error term u (σ_{ux}). If $\sigma_{ux} > 0$, then OLS will underestimate the true effect. If $\sigma_{ux} < 0$ OLS will still attenuate the estimated coefficients unless $\sigma_u^2 + \sigma_{ux} < 0$.

Another characterization of the bias due to u is in terms of the least-squares estimator from regressing the additive error u on w, which we will call β_{uw} (Bound et al., 2001). In the bivariate regression, the bias can be expressed as,

$$\beta_{OLS} = \beta(1 - \beta_{uw}). \tag{1.5}$$

The sign of β_{uw} depends on the relationships between σ_u^2 and σ_{ux} stated above.

In practice, it is possible that $\sigma_{u\varepsilon} \neq 0$, suggesting that u is endogenous to the error term. This may be due to x being correlated with unobservable factors that affect the outcome. In the context of the impact of obesity, reporting of height and weight may be correlated with factors such as culture or environment that also effect expenditures, but are unobservable and thus omitted from the estimated model. In this case, measurement error is said to be differential. Differential measurement error will cause further upward or downward bias in the estimated coefficients of all covariates in X.

If validation data are available, then it may be possible to correct for measurement error in x. This approach requires that there is at least one shared covariate that can be used to predict x in both datasets (Guo and Little, 2011). In previous research of the impact of obesity, the NHANES has served as validation data. The NHANES contains measured heights and weights as well as self-reported height and weight. Suppose that equation 1.2 describes self-reporting behavior for BMI, and that there exist validation data such that,

$$w_N = x_N + u_N, \tag{1.6}$$

Where the superscript denotes that the data are from the validation sample in the NHANES. Unlike the main sample, we observe x_N. This allows us to estimate the rela-

tionship between measured and reported BMI in the NHANES,

$$\hat{\alpha}_{xw} = (w'_N w_N)^{-1}(w'_N x_N) \tag{1.7}$$

We then estimate the least-squares estimator using the fitted values of $\hat{x} = w\hat{\alpha}_{xw}$. In general, replacing x in the analysis with the regression of X on (W, Z) is called regression calibration (Carroll et al., 2006). To disambiguate the method here from other applications of regression calibration, we refer to fitting the true predictor x in the main model with it's expected value conditional on the observed value w as the conditional expectation method (O'Neill and Sweetman, 2013; Lyles and Kupper, 1997). The validity of this approach depends on some assumptions about w. First, w must be a surrogate for y in both datasets, such that w does not contain any additional information about the distribution of y than x in both datasets. (this is identical in practice to assuming that the instrument is not endogenous to the error term. Second, the distribution of y given x and y given w must be the same in both datasets, which is known as transportability (Carroll et al., 2006; Lee and Sepanski, 1995). We assume here that these conditions are met.

The probability limit of $\hat{\beta}_{CE}$ from conditional expectations approach is,[1]

$$plim\hat{\beta}_{CE} = \beta + \frac{\sigma_{u\varepsilon}}{\sigma_x^2 + \sigma_{ux}}. \tag{1.8}$$

O'Neill and Sweetman (2013) demonstrate that the least-squares estimator using the fitted values $\hat{x}$ is consistent only when measurement error is classical. The conditional expectation estimator will exhibit the same bias as β_{OLS} when $Cov(X, u) = var(u)$, and the sign of β is the same as the sign of $Cov(u, \varepsilon)$. If $Cov(X, u) > Var(u)$, then non-classical measurement error will cause larger bias in β_{OLS} than in β_{CE}. Differential measurement error will cause larger bias in β_{OLS} than in $\hat{\beta}_{CE}$ as well. Thus using corrected BMI can tighten the bounds around the true coefficient relative to least-squares.

[1]The conditional expectation approach is equivalent to two-sample two-stage least squares (Inoue and Solon, 2010)

1.2.2 Instrumental Variables

Another possibility is that the true BMI x is itself endogenous, which is to say also correlated with ε. If this is true then observing the true value x will still result in biased estimates of β_{OLS}. For example, True BMI is endogenous in medical expenditure models because individuals with low SES tend to have higher rates of obesity, but poorer access to care, resulting in a confounding effect of elevated BMI on medical expenditures. We are unable to adequately control for SES, and associated cultural or environmental factors. The endogeneity of BMI has lead impact of obesity studies to use instrumental variables as a source of identification.

Previous research has established that the BMI of a biological relative is a valid instrument for BMI (Cawley, 2000, 2004). We assume that the BMI of a biological child s exists for our sample. For child BMI s to be a valid instrument in the bivariate linear model,

$$Cov(s, x) \neq 0, \tag{1.9a}$$

$$Cov(s\varepsilon) = 0. \tag{1.9b}$$

$$Cov(s, u) = 0, \tag{1.9c}$$

Equation 1.9a states that s must be correlated with x. Equation 1.9b states that the IV is uncorrelated with the model error term ε.[1] This is often referred to as the 'exclusion restriction' by which the IV must not be a part of the true model, and can only influence the outcome y through it's correlation with x.

Equation 1.9c states that s must be uncorrelated with measurement error $u = w - x$. Though typically not one of the canon assumptions of IV validity, this condition is a necessary assumption for the IV estimator to be consistent. If measurement error is classical random additive measurement error, then this assumption will always be satisfied.

[1] Non-linear IV methods requite a more stringent assumption that s is independent of the error term (Carroll et al., 2006).

Under the IV assumptions,

$$plim\hat{\beta}_{IV} = \frac{Cov(y,s)}{Cov(w,s)} = \frac{\sigma_{xs}\beta + \sigma_{s\varepsilon}}{\sigma_{xs} + \sigma_{su}} = \beta. \tag{1.10}$$

If any of the above assumptions are violated, $\hat{\beta}_{IV}$ will be biased. Following the empirical literature, we assume that condition 1.9a is satisfied such that we do not need to worry about bias due to weak instruments.[1] It is fundamentally unobservable whether the true instrument s is endogenous to the error term, and may result in attenuating bias or upward bias depending on the sign of $Cov(s,\varepsilon)$.

Non-classical measurement error in x, where $Cov(x,u) \neq 0$, may result in the true instrument being correlated with the additive error u and violating the condition in equation 1.9c. In survey data such as the MEPS, a single respondent will self-report their own height and weight and proxy-report the height and weight of their child. Although covariance is not transitive, it is feasible that equation 1.9a and $Cov(x,u) \neq 0$ will result in $Cov(s,u) \neq 0$. If adults tend to under-report their BMI, then we may find that $Cov(s,u) < 0$. This will cause the IV estimator to overestimate the true β. Further, the size of the bias does not depend on the size of the measurement error variance. This implies that even if few observations are measured with error, it can cause significant bias.

We now allow for the instrument to be measured with error such that,

$$t = s + v \tag{1.11}$$

where t is the observed instrument. We update the probability limit of $\hat{\beta}_{IV}$ to incorporate measurement error in the instrument,

$$plim\hat{\beta}_{IV} = \frac{Cov(y,t)}{Cov(w,t)} = \frac{\beta(\sigma_{xs} + \sigma_{xv}) + \sigma_{s\varepsilon} + \sigma_{v\varepsilon}}{\sigma_{xs} + \sigma_{xv} + \sigma_{su} + \sigma_{vu}}. \tag{1.12}$$

The IV estimator is consistent if the conditions in 1.9 hold for t and if $Cov(v,\varepsilon) =$

[1] See section 3.2.1 for a discussion of the empirical validity of this instrument, and Stock et al. (2002); Bound et al. (1995) for a discussion of the bias due to weak instruments.

$Cov(v, u) = 0$. In the context of reported weight, we may observe $Cov(v, u) \neq 0$ if a parent uses the same process to report their weight as they do their child's weight. This may be incidentally true if both adult and child BMI share the same relationship with their reporting error. An example is if both adult and child BMI are reported with non-classical measurement error such that $Cov(x, u) < 0$ and $Cov(s, v) < 0$. This is not a strict implication, and even if adult and child BMI are both increasingly under-reported as BMI increases, the correlation in reporting errors may be effectively zero.

IV estimation will be consistent if x is endogenous due to differential measurement error. If the IV suffers from differential measurement error, then it will be endogenous to the error term and will introduce bias. The direction of this bias is determined by the sign of $Cov(v, \varepsilon)$. Interestingly, it is possible for true child BMI s to be a valid instrument, yet for the mismeasured IV t to be invalid due to differential measurement error.

1.2.3 Regression Calibration in Linear IV Estimation

Validation data may exist for both x and s. The NHANES contains measured height and weight for both adults and children as well self-reported height and weight for adults, and can be used to correct for measurement error. For example, Cawley (2000) uses the NHANES III to correct for reporting error in both the endogenous regressor and the IV in the National Longitudinal Survey of Youth (NSLY).

Let us suppose that in addition to the presence of measured and reported height and weight as in equation 1.6, there exists a measured and reported child BMI in the NHANES,[1]

$$t_N = s_N + v_N. \tag{1.13}$$

Measured child BMI is regressed on reported BMI as in equation 1.7 to generate fitted

[1]With the exception of children above the age of 16, there are no reported measures for height and weight for children in the NHANES. We assume they exist to simplify the discussion here, which generalizes to validation using any valid matching surrogate. Courtemanche et al. (2014) demonstrate that the percentile rank of BMI can be used as as matching surrogate in lieu of self-reported BMI. Subsequent chapters implement validation using this method and demonstrate that percentile rank satisfies the conditions for the conditional expectation approach.

values $\hat{s} = t\hat{\lambda}$, where $\hat{\lambda}$ is the estimated coefficient for the instrument from the regression calibration equation in the validation sample. The conditional expectation IV estimator is

$$\beta_{RCIV} = (\hat{s}'\hat{x})^{-1}(\hat{s}'y) \tag{1.14}$$

The probability limit of β_{RCIV} is,

$$plim\hat{\beta}_{RCIV} = \frac{Cov(y, \hat{s})}{Cov(\hat{x}, \hat{s})} = \frac{\beta\sigma_{x\hat{s}} + \sigma_{\hat{s}\varepsilon}}{\sigma_{\hat{x}\hat{s}}}. \tag{1.15}$$

The RCIV estimator β_{RCIV} is approximately consistent if $Cov(x, \hat{s}) \approx Cov(\hat{x}, \hat{s})$. It is still a necessary assumption that $\hat{s}$ is not endogenous to the error term, which is by extension of the IV exclusion restriction in (1.9b). This closeness of $Cov(\hat{x}, \hat{s})$ to $Cov(x, \hat{s})$ depends on the the validation assumptions in Lee and Sepanski (1995) being satisfied.

Even if transportability holds, β_{RCIV} may still be inconsistent. This can be seen using the error due to imputation. We define the imputation error as,

$$\hat{u} = \hat{x} - x, \tag{1.16a}$$

$$\hat{v} = \hat{s} - s. \tag{1.16b}$$

The resulting condition for consistency is that $Cov(\hat{S}, \hat{u}) = 0$. We can expand (1.15) as,

$$plim\hat{\beta}_{RCIV} = \frac{\beta\sigma_{x\hat{s}}}{\sigma_{x\hat{s}} + \sigma_{\hat{s}\hat{u}}}.$$

Thus the estimator will be consistent only when $Cov(\hat{S}, \hat{u}) = 0$.

Another interpretation of the consistency conditions for RCIV is whether there exists fitted values that satisfy them. Consider estimating a RCIV model where first $\hat{\alpha}$ and $\hat{\lambda}$ are estimated using (1.7). The fitted conditional expectations are $\hat{x} = w\hat{\alpha}$ and $\hat{s} = t\hat{\lambda}$. Under transportability, we can assume that the pattern of reporting error is identical in the validation data as it is in the main sample. We can then solve for α^* that causes

β_{RCIV} to consistently estimate β,

$$\alpha^* = \frac{Cov(x,t)}{Cov(w,t)}. \tag{1.17}$$

The optimal alpha could be obtained from an IV regression of x on w in the main sample, but would require knowledge of true x. The validation sample can only produce α^* if both x_N and s_N are in the same dataset, and $E(x_N s_N) = E(xs)$. Replicate measures of x and s would be sufficient, however x and s may have separate validation. For example, the NHANES contains adult and child BMI, but it is not possible to match parents and children in the public-use NHANES.

If α^* cannot be obtained, regression calibration may generate inconsistent IV estimates. Under perfect transportability, consistency would require that, $Cov(x,w)/Var(w) = Cov(x,t)/Cov(w,t)$. It is easy to deduce that this will not hold whenever IV regression of x on w is warranted (x is endogenous, or measured with error).

To change the conditions under which RCIV estimation will be consistent, we propose as an alternative estimators adjusted regression calibrated IV (ARCIV) which is estimated as,

$$\beta_{ARCIV} = (\hat{s}'w)^{-1}(\hat{s}'y), \tag{1.18}$$

which is simply the IV regression of Y on w, using $\hat{S}$ as the instrument. The only difference between ARCIV and IV is that the covariance between error in the instrument and the model error term is replaced with $Cov(\hat{v}\varepsilon)$. If there is no differential measurement error in the IV, then the ARCIV estimator reduces to the IV estimator.

The degree of bias in the RCIV estimators is a function of the method of imputation. A poor regression calibration fit introduces more error into the fitted values, which can exacerbate the RCIV bias. Under more sophisticated validation methods that improve the fitting of $\hat{x}$ and $\hat{s}$, the magnitude of any bias in the RCIV will be reduced. This implies that in addition to transportability and exogenous true instruments, RCIV requires that the conditional expectations must be very close to the true values.

1.3 Regression Calibration in IV Log-Gamma Model of Medical Expenditures

1.3.1 Log-Gamma Model

Early literature in health expenditure modeling focuses on the non-linear relationship between medical expenditures and BMI. This non-linear relationship is partly due to the distribution of medical expenditures, which only has support over non-negative values, typically has a mass point at zero, and can have severe positive skewness (Jones, 2000). In some circumstances, it is sufficient to get consistency by estimating a linear model using a transformed outcome variable, such as the log of expenditures. However, a log-linear model may not fully correct for heteroscedasticity. A commonly used method to model expenditures is to employ a Generalized Linear Model to relate BMI to medical expenditures (Manning and Mullahy, 2001). We are particularly interested in a Gamma GLM with log-link, as this is the GLM specification that empirical studies have identified as the correct fit for the data (Manning and Mullahy, 2001; Finkelstein et al., 2009; Cawley and Meyerhoefer, 2012).[1]

In this section a very simple framework for GLM will be established. We follow Carroll et al. (2006) and describe GLM as a mean and variance model, called a quasilikelihood and variance function (QVF). The advantage is it is easy to show how measurement error introduces bias, and to demonstrate how regression calibration and instrumental variables can be used to correct for error. We narrow the focus to Gamma GLM with log-link, and describe the conditional expectations instrumental variables estimator for this specification. We will assume that the GLM is correctly specified, as the concern here is consistency when covariates are measured with error, and not the choice of model.[2]

Mean and variance models specify the mean and variance conditional on model covariates (thus variance can be said to be conditionally heteroscedastic). In general, we

[1]For a general description of measurement error in non-linear models, see Lee and Sepanski (1995).

[2]A set of heuristics for selecting the appropriate model for non-linear outcomes can be found in Manning and Mullahy (2001).

express this as,

$$E(y|X) = \mu(X,\theta), \tag{1.19a}$$

$$Var(y|X) = \phi V(\mu(X,\theta)), \tag{1.19b}$$

Where $V(\mu(Z,X,\theta))$ is a non-negative function, ϕ is a scale parameter, and θ is a vector containing parameters that specify both functions. The Log-gamma model is then specified as,

$$\mu = exp(\eta), \tag{1.20a}$$

$$V(\mu) = \mu^2, \tag{1.20b}$$

Where η is the linear predictor $X\beta$. The inverse link function 1.20a relates the linear predictor to the conditional mean of outcome y. The gamma conditional variance is the square of the inverse link. The GLM is fit by setting up and solving a set of estimating equations (Hardin and Hilbe, 2012). For the Log-Gamma Model, the equations to solve for β are,

$$\sum_{i=1}^{n} \frac{y_i - \mu_i}{V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_k} = \sum_{i=1}^{n} \frac{y_i - exp(X_i\beta_k)}{exp(X_i\beta_k)} X_i = 0, \forall\ k. \tag{1.21}$$

In practice, the GLM model is fit using iteratively reweighted least squares (IRLS). This is the method used in the STATA **QVF** command employed in subsequent chapters (Hardin et al., 2003; Hardin and Carroll, 2003a). For a complete description of IRLS for the Log-Gamma Model, see Appendix A.1. If the IRLS algorithm converges to a fixed coefficient vector β_{IRLS}, then it is a MLE of β, as long as x is measured without error. The β_{IRLS} for the Log-Gamma model specifically can be thought of as the result of iterated least squares. The iterations do not need to be reweighted as the Log-Gamma model is a special case where the weighting matrix that relates the linear predictor to the variance devolves to the identity matrix. (This is the same as GLS where $Var(\varepsilon|X) = I$.)

Measurement error in X will cause β_{IRLS} to fit the GLM by minimizing the naive estimating equations (combining (1.6) and (1.21)),

$$\sum_{i=1}^{n} \frac{y_i - exp(W_i\beta_k)}{exp(W_i\beta_k)} W_i = 0, \forall\ k. \tag{1.22}$$

As long as $Var(u) \neq 0$, β_{IRLS} will be inconsistent. Interestingly, measurement error in X will introduce the same magnitude asymptotic bias as in OLS. This is because the measurement error variance cannot influence the weighting matrix in the IRLS algorithm ($\Xi = I$). This implies that the linear model can serve as a conceptual proxy for the Log-Gamma model when it comes to quantifying the asymptotic bias due to measurement error in X when it is both classical or non-classical in nature.

1.3.2 Conditional Expectations Instrumental Variables in Log-Gamma Models

In the context of medical expenditure models, If X is BMI, then it will be endogenous to the error term even in the non-linear model whether it is measured with error or not. This endogeneity can be corrected by using instrumental variables. Hardin et al. (2003) describe the IV estimation procedure for mean and variance models. Their IV-GLM estimator is robust to differential measurement error in X, and is the only correction method so far attempted to mitigate the bias due to endogeneity of BMI in non-linear medical expenditure models (Cawley and Meyerhoefer, 2012; Cawley et al., 2014). The estimator is effectively a two-stage non-linear least squares. The first stage is linear, in which observed BMI W is regressed on matrix R, which is an augmented matrix of exogenous variables $[Z,\ S]$ consisting of instruments S and other covariates Z, in order to obtain the coefficient vector $\hat{\gamma}_j$ for $j = 1, ..., k$. The second stage fits the specified GLM using IRLS with general estimating equations.

$$\sum_{i=1}^{n} \frac{y_i - \mu_i}{V(\mu_i)} \frac{\partial \mu_i}{\partial \eta_i} (Z,\ R\hat{\gamma}) = 0, \quad \forall\ k. \tag{1.23}$$

As in the linear case, the IV-GLM model will be inconsistent if $Cov(S, u) \neq 0$.

If validation data are available, the IV-GLM estimator can be corrected by using the conditional expectation approach. The IV-GLM model estimated using IRLS with both W and T (allowing the IV to be mismeasured as well) replaced by their conditional expectations $\hat{X}$ and $\hat{S}$ respectively from equation 1.7. For the Log-Gamma Model with all covariates measured with error, the estimating equation of the GLM is,

$$\sum_{i=1}^{n} \frac{y_i - exp(\hat{S}\hat{\gamma}_i\beta_k)}{exp(\hat{S}\hat{\gamma}_i\beta_k)}\hat{S}\hat{\gamma}_i = 0, \quad \forall\ k, \tag{1.24}$$

where the linear first stage regression is,

$$\hat{\gamma} = (\hat{X}'\hat{X})^{-1}(\hat{X}'\hat{S}). \tag{1.25}$$

As in the linear case, the RCIV-GLM model will be approximately consistent if $Cov(\hat{S}, \hat{u}) = 0$, as long as the true instrument is not endogenous, and the validation assumptions are satisfied.

1.4 Monte Carlo Simulations

1.4.1 Simulation Framework

We follow the general simulation framework from Manning and Mullahy (2001) to demonstrate the bias due to measurement error in the context of Log-OLS and Log-Gamma models. Unlike their simulations, we regulate the correlations between variables by drawing from a joint-normal distribution. The main data $X = [x|z_1, z_2, z_3]$, instrument s with its own associated covariates z_4, z_5, and z_6, additive error u and v and model error ε are drawn from a single joint normal distribution. Mismeasured covariates are generated using (1.2) and (1.11).

The corresponding variables from the validation data are drawn from a separate distribution. The measurement error in x and s are identical in both datasets, which would imply transportability. The covariance between x and s in the validation draws are set to zero and $Cov(S_N, u_N) = 0$. We set $Var(x) = Var(s) = 1$. The variance of all covariates

are set equal to one as well. All exogenous variables have mean 0.5. Select variables, z_1, z_2, z_4, z_5, are transformed into binary variables that are 1 when $z \geq 0.5$ and zero otherwise.

We construct linear predictor $\eta = \beta_0 + 1x + 1d_1 + 1d_2 + 1z_3$. The log outcome is then $ln(y) = \eta + \varepsilon$. Error in the log-scale generates skewness on raw scale. The raw scale relationship is $E(y|X) = exp(\beta_0 + \beta X + 0.5\sigma_\varepsilon^2)$, where β_0 is chosen with σ_ε to set $E(y|X) = 1$. The gamma distributed outcome is drawn from a gamma distribution with $\mu = exp(\eta)$ and a scale parameter of 0.5.[1] We set $E(\varepsilon) = 0$, so that all models are correctly specified.

We compare OLS, IV, and RCIV under two validation procedures. First we reproduce the standard validation method in Cawley (2000) by regressing measured outcomes in the validation data on higher order functions of their mismeasured values and additional covariates. We also reproduce the validation method proposed in Courtemanche et al. (2014) that uses the percentile rank of of w and w_N as the matching surrogate in addition to exogenous covariates. In our simulations, the standard method has $R^2 = .92$ on average; The percentile-rank method has $R^2 = .99$.

For clarity of presentation, we initially do not allow x to be endogenous. When $E(x\varepsilon) \neq 0$, Even though the OLS estimator may appear consistent under certain measurement error conditions, it is understood that the motivation for IV estimation is that OLS is biased due to endogeneity. The relative magnitudes of measurement error variances presented here are calibrated using the NHANES validation data for adults used in Chapter 3. We use 500 simulations, and set the number of observations to 1,500. We alternatively used 10,000 observations and found that such a large sample size obscures variation in the finite sample performance of the estimators.

1.4.2 Simulated Measurement Error

Table 1.1 contains the simulation results where regression calibration is performed using percentile rank. The first panel contains the baseline results where x or s are not

[1] We set all zero draws to .000001, otherwise the IV-GLM models will not converge.

measured with error. The sample covariances are computed using the fitted values and predicted $\tilde{u}$ and $\tilde{v}$ where, $\tilde{u} = w - \hat{x}$ and $\tilde{v} = t - \hat{s}$.

The second panel introduces classical measurement error in x. Uncorrected estimates are predictably biased downward as in equation 1.3. Interestingly, the RCIV estimator is slightly biased upwards. Since the instrument is not endogenous and the true coefficient $\beta = 1$, equation 1.15 tells us $\beta_{RCIV} = Cov(x, \hat{s})/Cov(\hat{x}, \hat{s})$. When $\sigma_u^2 = 0.10$, $\beta_{RCIV} = 0.248/0.236 = 1.051$. The difference between $Cov(x, \hat{s})$ and $Cov(\hat{x})$ is equal to $Cov_(\hat{s}\hat{u}) = 0.013$. This term can differ from zero even if $Cov(s, u) = Cov(s, v) = Cov(x, v) = 0$ if $\alpha \neq \alpha^*$. The ARCIV estimator does not impute x and will not exhibit this bias. Panel 3 introduces under-reporting which shifts the measurement error distribution. There is little effect of this shift on the average coeffecient estimates, although the variance of the GLM estimates increases slightly.

Panels four and five contain the results for non-classical measurement error in x where $Cov(x, u < 0$. These clearly illustrate that the RCIV estimators bias stems from the regression calibration stage. When $\sigma_u^2 = 0.1$ and $\sigma_{xu} = -.1$, equation 1.4 shows that $\hat{\alpha} = 1$. In this case, none of the RCIV estimators are biased. However, when $\hat{\alpha} \neq 1$, The RCIV estimator is again biased. ARCIV is as robust to non-classical measurement error in x as the regular IV model will be. Correlation between x and its own error does not explicitly factor into any asymptotic bias in β_{IV}, but is a kind of omitted variables bias in β_{OLS} for which IV, and in turn ARCIV can correct.

Table 1.2 contain the simulation results in which measurement error in the instrument is also introduced. In panel one, both x and s are measured with classical measurement error. IV estimates are robust to classical error, so long as it does not lead to a weak IV problem. ARCIV is also robust to this problem. The standard RCIV is upwardly biased, since $Cov(x, \hat{s})/Cov(\hat{x}, \hat{s}) = 0.241/0.228 = 1.057$. This difference is again driven by $Cov(\hat{s}, \hat{u}) \neq 0$. Imputation of the instrument is responsible for $Cov(x\hat{s}) \neq Cov(x, s)$. Thus even classical measurement error in the IV can cause bias in the RCIV estimator. The RCIV estimator is similarly biased when the measurement error in s is non classical.

Panels three and four contain estimated coefficients when both x and s exhibit non-

classical error. OLS estimates are only incidentally close to β, and will be attenuated when x is endogenous. IV and ARCIV are both robust to non-classical error in both x and s, even with $E(x) < 0$. The standard RCIV estimator understates the true effect. Notably in panel 4, ARCIV is less biased than IV. However, IV still performs significantly better than RCIV, and is slightly more efficient than ARCIV. When we increase the number of observations, ARCIV performs no better than IV. ARCIV may have slightly better finite sample performance over IV, at least in this data scenario.

Even though $Cov(s, u) = 0$, the estimated covariance is not zero. However, the estimated covariance $Cov(\hat{s}, \tilde{u})$, does provide an estimate of $Cov(\hat{s}, \hat{u})$ even when both x and s exhibit non-classical measurement error. We can estimate $Cov(x, \hat{s}) = Cov(\hat{x}, \hat{s}) + Cov(\hat{s}, \tilde{u})$ and then explain the difference between the RCIV and the ARCIV estimator. We find this relationship in the subsequent empirical chapters, suggesting that the ARCIV estimates may be consistent, with RCIV generating slight underestimates, as they are in panels three and four of Table 1.2. We characterize this relationship as $\beta_{RCIV} = \beta_{ARCIV} \times \delta$ where,

$$\delta = \frac{\sigma_{\hat{s}\hat{x}} + \sigma_{\hat{x}\hat{v}} + \sigma_{\hat{s}\hat{u}} + \sigma_{\hat{u}\hat{v}}}{\sigma_{\hat{s}\hat{x}} + \sigma_{\hat{x}\hat{v}}}. \tag{1.26}$$

In panel 1 for instance, $\delta = (0.222 + 0.013)/(0.222) \approx 1.059$, and $\beta_{RCIV} = 1.050 = 0.991 \times 1.059$. In panel 4, $\delta = (0.278 - 0.013)/(0.278) \approx 0.953$, and $\beta_{RCIV} = 0.950 = 1.004 \times 0.953$.

Panel five in Table 1.2 introduces correlation between the instrument and measurement error in x. All IV and RCIV estimators will dramatically overstate the true effect. In the presence of this correlation, it may only be possible to bound the true effect. β_{OLS} will likely under-state the true effect due to endogeneity or measurement error in x. The RCIV estimator is potentially less upwardly biased than IV estimation, and can potentially tighten the bounds around the true effect, even when regression calibration is unable to produce a consistent estimate of β.

Table 1.3 contains the simulation results from panel three of Table 1.2 with additional endogeneity. Panel one incorporates correlation between x and the model error term ε.

We focus only on the Log-Linear Estimator since correlation with error in the log-scale causes the GLM estimator to be mis-specified. OLS is biased downward, and IV is still consistent since the instrument is exogenous. The RCIV estimator is downwardly biased, but not by as much as OLS. The ARCIV estimator still mirrors the IV estimator and consistently estimates β. The robustness of IV and RCIV estimators to endogeneity is reassuring as the endogeneity of x is the primary motivation for using instrumental variables to begin with.

Panel two in Table 1.3 adds correlation between measurement error in the IV and the model error term. Such a correlation may arise even if the true instrument is exogenous. For instance, reporting error may be associated with unobservable SES or cultural characteristics that are also correlated with the outcome. All estimators are biased downward when $Cov(v, \varepsilon) < 0$. However, the IV and ARCIV estimators are less biased than OLS and RCIV. Conceptually, if ARCIV perfectly or near-perfectly predicts, s, then ARCIV may perform better than RCIV by causing $Cov(\hat{v}, \varepsilon) < Cov(v, \varepsilon)$.

1.5 Discussion

This chapter describes bias due to measurement error in IV estimation used in common models of health health expenditures and describe how regression calibration may improve the performance of IV estimation of the marginal effect of BMI on spending. Analytical results demonstrate that regression calibration of both the endogenous regressors and the instrument may introduce bias into IV estimates, as the fitted values contain information about the mismeasured variables used to predict them. We show that in the standard method of regression calibration, this bias is unavoidable in the RCIV model. We propose two alternative estimators to RCIV. ARCIV simply performs IV estimation using fitted values for the instrument to IV regress the outcome on observed, possibly mismeasured covariates. In simulations we find that this estimator performs equally well to IV estimation, both of which are consistent when RCIV is biased.

Using the simulations only, it is difficult to select between IV and ARCIV solely based

on consistency. However, these simulations do have limitations. The draws are all from joint-normal draws, and thus all data have the same distribution shapes. The regression calibrated estimators may have better performance than IV under different data conditions. Another reason ARCIV may be preferable to IV is that, under sufficiently strong validation, the imputed error may be equal to zero in expectation, and thus diminish the covariance between imputed error and variables used in the estimation. The drawback of ARCIV is that standard errors are more difficult to compute as they must incorporate the variance due to imputation. Despite these concerns, ARCIV is no more biased than IV, and may potentially mitigate certain biases. For this reason we estimate the main results in subsequent chapters using the ARCIV estimator.

Regression calibration is not robust to differential measurement error, or correlation between the IV and measurement error in x. Substituting regression calibration with other stochastic imputation methods may improve the performance of the RCIV and ARCIV estimators in these severe situations (O'Neill and Sweetman, 2013). If knowledge of the outcome variable was available in the validation data, stochastic multiple imputation could produce RCIV and ARCIV estimators more robust to differential measurement error (Freedman et al., 2009). It is possible to use improved regression calibration or stochastic imputation in lieu of regression calibration. In the absence of additional data, stochastic imputation can only improve on regression calibration if the regression model is not flexible enough to fit the relationship between true values and their mismeasured counterparts. Otherwise, stochastic imputation will simply reproduce fitted values with the same properties as those from regression calibration. No validation data exists with both measured BMI and medical spending. If the NHANES were to survey a sub-sample of respondents about medical expenditures, or the MEPS to measure the heights and weights of a weighted sub-sample of household respondents, then more refined correction methods could be employed that are more robust to differential measurement error, and possibly other severe measurement error problems, than the conditional expectation estimators.

Table 1.1: Simulation Results - Measurement error in x

			Log-Linear Estimator			Log-Gamma Estimator			Sample Cov.	
No Measurement Error			Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.250
$\sigma_{xs} =$ 0.25	$\sigma_{xt} =$	0.250	OLS	1.000	0.028	GLM	1.013	0.069	$\sigma_{x\hat{s}} =$	0.248
$E(u) =$ 0	$E(\varepsilon) =$	1	IV	0.994	0.101	IV-GLM	0.997	0.177	$\sigma_{\hat{x}\hat{s}} =$	0.246
$\sigma_u^2 =$ 0	$\sigma_{xu} =$	0	RCIV	0.999	0.106	RCIV-GLM	1.002	0.181	$\sigma_{\hat{x}\tilde{u}} =$	-0.007
$\sigma_v^2 =$ 0	$\sigma_{sv} =$	0	ARCIV	0.994	0.102	ARCIV-GLM	0.996	0.177	$\sigma_{\hat{s}\tilde{v}} =$	-0.006
									$\sigma_{\hat{s}\hat{u}} =$	0.002
									$\sigma_{\hat{s}\tilde{u}} =$	0.002
Classical M.E. in x			Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.251
$\sigma_{xs} =$ 0.25	$\sigma_{xt} =$	0.251	OLS	0.895	0.029	GLM	0.911	0.058	$\sigma_{x\hat{s}} =$	0.249
$E(u) =$ 0	$E(\varepsilon) =$	1	IV	1.001	0.113	IV-GLM	1.001	0.161	$\sigma_{\hat{x}\hat{s}} =$	0.235
$\sigma_u^2 =$ 0.1	$\sigma_{xu} =$	0	RCIV	1.063	0.124	RCIV-GLM	1.063	0.173	$\sigma_{\hat{x}\tilde{u}} =$	0.043
$\sigma_v^2 =$ 0	$\sigma_{sv} =$	0	ARCIV	1.002	0.113	ARCIV-GLM	1.002	0.161	$\sigma_{\hat{s}\tilde{v}} =$	-0.005
									$\sigma_{\hat{s}\hat{u}} =$	0.014
									$\sigma_{\hat{s}\tilde{u}} =$	0.015
Classical M.E. in x										
$E(u) < 0$			Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.250
$\sigma_{xs} =$ 0.25	$\sigma_{xt} =$	0.251	OLS	0.890	0.026	GLM	0.901	0.055	$\sigma_{x\hat{s}} =$	0.237
$E(u) =$ -0.1	$E(\varepsilon) =$	1	IV	0.995	0.120	IV-GLM	1.019	0.171	$\sigma_{\hat{x}\hat{s}} =$	0.213
$\sigma_u^2 =$ 0.1	$\sigma_{xu} =$	0	RCIV	1.107	0.138	RCIV-GLM	1.135	0.194	$\sigma_{\hat{x}\tilde{u}} =$	0.089
$\sigma_v^2 =$ 0	$\sigma_{sv} =$	0	ARCIV	0.995	0.122	ARCIV-GLM	1.020	0.172	$\sigma_{\hat{s}\tilde{v}} =$	0.043
									$\sigma_{\hat{s}\hat{u}} =$	0.024
									$\sigma_{\hat{s}\tilde{u}} =$	0.024
Non-Classical M.E. in x										
$Cov(x,u) < 0$			Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.248
$\sigma_{xs} =$ 0.25	$\sigma_{xt} =$	0.251	OLS	0.943	0.028	GLM	0.951	0.065	$\sigma_{x\hat{s}} =$	0.248
$E(u) =$ 0	$E(\varepsilon) =$	1	IV	1.004	0.115	IV-GLM	1.002	0.163	$\sigma_{\hat{x}\hat{s}} =$	0.247
$\sigma_u^2 =$ 0.1	$\sigma_{xu} =$	-0.05	RCIV	1.006	0.120	RCIV-GLM	1.008	0.169	$\sigma_{\hat{x}\tilde{u}} =$	0.047
$\sigma_v^2 =$ 0	$\sigma_{sv} =$	0	ARCIV	1.003	0.116	ARCIV-GLM	1.004	0.164	$\sigma_{\hat{s}\tilde{v}} =$	-0.001
									$\sigma_{\hat{s}\hat{u}} =$	0.001
									$\sigma_{\hat{s}\tilde{u}} =$	0.001
Non-Classical M.E. in x										
$Cov(x,u) < 0$			Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.249
$\sigma_{xs} =$ 0.25	$\sigma_{xt} =$	0.251	OLS	0.882	0.030	GLM	0.901	0.056	$\sigma_{x\hat{s}} =$	0.235
$E(u) =$ 0	$E(\varepsilon) =$	1	IV	1.004	0.123	IV-GLM	1.021	0.177	$\sigma_{\hat{x}\hat{s}} =$	0.234
$\sigma_u^2 =$ 0.1	$\sigma_{xu} =$	-0.1	RCIV	1.010	0.126	RCIV-GLM	1.026	0.182	$\sigma_{\hat{x}\tilde{u}} =$	-0.006
$\sigma_v^2 =$ 0	$\sigma_{sv} =$	0	ARCIV	1.004	0.123	ARCIV-GLM	1.020	0.178	$\sigma_{\hat{s}\tilde{v}} =$	0.041
									$\sigma_{\hat{s}\hat{u}} =$	0.001
									$\sigma_{\hat{s}\tilde{u}} =$	0.002

Notes: $N = 1,500$. Results averaged over 500 simulations. True coefficient $\beta = 1$. Predicted values $\hat{x}$ and $\hat{s}$ are fitted by regressing x_N and s_N on percentile rank of observed values interacted with exogenous covariates. Some sample covariances use predicted $\tilde{u}$ and $\tilde{v}$ where, $\tilde{u} = w - \hat{x}$ and $\tilde{v} = t - \hat{s}$. Unknown error introduced by imputation $\hat{u}$ and $\hat{v}$ are defined in (1.16).

Table 1.2: Simulation Results - Measurement Error in x and s

				Log-Linear Estimator			Log-Gamma Estimator			Sample Cov.	
Classical M.E. in x and s				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.250
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.251	OLS	0.894	0.027	GLM	0.908	0.061	$\sigma_{x\hat{s}} =$	0.236
$E(u) =$	0	$E(\varepsilon) =$	1	IV	0.990	0.116	IV-GLM	1.005	0.175	$\sigma_{\hat{x}\hat{s}} =$	0.222
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	0	RCIV	1.050	0.125	RCIV-GLM	1.065	0.190	$\sigma_{\hat{x}\tilde{u}} =$	0.042
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	0	ARCIV	0.991	0.118	ARCIV-GLM	1.005	0.178	$\sigma_{\hat{s}\tilde{v}} =$	0.044
										$\sigma_{\hat{s}\hat{u}} =$	0.013
										$\sigma_{\hat{s}\tilde{u}} =$	0.014
Classical M.E. in x											
Non-Classical M.E in s				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.251
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.251	OLS	0.898	0.027	GLM	0.911	0.061	$\sigma_{x\hat{s}} =$	0.264
$E(u) =$	0	$E(\varepsilon) =$	1	IV	1.006	0.102	IV-GLM	1.016	0.175	$\sigma_{\hat{x}\hat{s}} =$	0.249
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	0	RCIV	1.067	0.112	RCIV-GLM	1.078	0.187	$\sigma_{\hat{x}\tilde{u}} =$	0.045
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	1.006	0.102	ARCIV-GLM	1.016	0.178	$\sigma_{\hat{s}\tilde{v}} =$	-0.058
										$\sigma_{\hat{s}\hat{u}} =$	0.016
										$\sigma_{\hat{s}\tilde{u}} =$	0.015
Non-Classical M.E. in x and s				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.248
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.248	OLS	1.000	0.031	GLM	1.016	0.066	$\sigma_{x\hat{s}} =$	0.261
$E(u) =$	0	$E(\varepsilon) =$	1	IV	1.004	0.100	IV-GLM	1.018	0.169	$\sigma_{\hat{x}\hat{s}} =$	0.274
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	-0.1	RCIV	0.954	0.100	RCIV-GLM	0.967	0.164	$\sigma_{\hat{x}\tilde{u}} =$	-0.057
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	1.004	0.101	ARCIV-GLM	1.018	0.164	$\sigma_{\hat{s}\tilde{v}} =$	-0.056
										$\sigma_{\hat{s}\hat{u}} =$	-0.013
										$\sigma_{\hat{s}\tilde{u}} =$	-0.013
Non-Classical M.E. in x and s											
$E(u) < 0$				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.251
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.251	OLS	1.001	0.030	GLM	1.018	0.062	$\sigma_{x\hat{s}} =$	0.264
$E(u) =$	-0.1	$E(\varepsilon) =$	1	IV	0.999	0.099	IV-GLM	1.010	0.155	$\sigma_{\hat{x}\hat{s}} =$	0.278
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	-0.1	RCIV	0.950	0.045	RCIV-GLM	0.957	0.152	$\sigma_{\hat{x}\tilde{u}} =$	-0.057
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	1.000	0.099	ARCIV-GLM	1.007	0.156	$\sigma_{\hat{s}\tilde{v}} =$	-0.058
										$\sigma_{\hat{s}\hat{u}} =$	-0.013
										$\sigma_{\hat{s}\tilde{u}} =$	-0.013
Non-Classical M.E. in x and s											
$\sigma_{su} < 0$				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.150
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.250	OLS	0.999	0.031	GLM	1.015	0.068	$\sigma_{x\hat{s}} =$	0.262
$E(u) =$	0	$E(\varepsilon) =$	1	IV	1.696	0.210	IV-GLM	1.691	0.288	$\sigma_{\hat{x}\hat{s}} =$	0.165
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	-0.1	RCIV	1.615	0.200	RCIV-GLM	1.607	0.277	$\sigma_{\hat{x}\tilde{u}} =$	-0.055
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	1.697	0.210	ARCIV-GLM	1.688	0.288	$\sigma_{\hat{s}\tilde{v}} =$	-0.055
$\sigma_{su} =$	-0.05									$\sigma_{\hat{s}\hat{u}} =$	-0.098
										$\sigma_{\hat{s}\tilde{u}} =$	-0.008

Notes: $N = 1,500$. Results averaged over 500 simulations. True coefficient $\beta = 1$. Predicted values $\hat{x}$ and $\hat{s}$ are fitted by regressing x_N and s_N on percentile rank of observed values interacted with exogenous covariates. Some sample covariances use predicted $\tilde{u}$ and $\tilde{v}$ where, $\tilde{u} = w - \hat{x}$ and $\tilde{v} = t - \hat{s}$. Unknown error introduced by imputation $\hat{u}$ and $\hat{v}$ are defined in (1.16).

Table 1.3: Simulation Results - Endogeneity and Differential Measurement Error

				Log-Linear Estimator			Log-Gamma Estimator			Sample Cov.	
Non-Classical M.E. in x and s											
$Cov(x,\varepsilon) < 0$				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.251
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.251	OLS	0.592	0.028	GLM	–	–	$\sigma_{x\hat{s}} =$	0.249
$E(u) =$	0	$E(\varepsilon) =$	1	IV	1.000	0.112	IV-GLM	–	–	$\sigma_{\hat{x}\hat{s}} =$	0.248
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	-0.1	RCIV	1.005	0.037	RCIV-GLM	–	–	$\sigma_{\hat{x}\tilde{u}} =$	-0.007
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	1.000	0.112	ARCIV-GLM	–	–	$\sigma_{\hat{s}\tilde{v}} =$	-0.007
$\sigma_{x\varepsilon} =$	-0.25									$\sigma_{\hat{s}\hat{u}} =$	0.001
										$\sigma_{\hat{s}\tilde{u}} =$	0.002
Non-Classical M.E. in x and s											
$Cov(v,\varepsilon) < 0$				Estimator	Mean	S.D.	Estimator	Mean	S.D.	$\sigma_{wt} =$	0.250
$\sigma_{xs} =$	0.25	$\sigma_{xt} =$	0.250	OLS	0.998	0.032	GLM	–	–	$\sigma_{x\hat{s}} =$	0.263
$E(u) =$	0	$E(\varepsilon) =$	1	IV	0.801	0.105	IV-GLM	–	–	$\sigma_{\hat{x}\hat{s}} =$	0.276
$\sigma_u^2 =$	0.1	$\sigma_{xu} =$	-0.1	RCIV	0.761	0.037	RCIV-GLM	–	–	$\sigma_{\hat{x}\tilde{u}} =$	-0.058
$\sigma_v^2 =$	0.1	$\sigma_{sv} =$	-0.1	ARCIV	0.801	0.104	ARCIV-GLM	–	–	$\sigma_{\hat{s}\tilde{v}} =$	-0.055
$\sigma_{x\varepsilon} =$	0	$\sigma_{v\varepsilon} =$	-0.05							$\sigma_{\hat{s}\hat{u}} =$	-0.014
										$\sigma_{\hat{s}\tilde{u}} =$	-0.014

Notes: $N = 1,500$. Results averaged over 500 simulations. True coefficient $\beta = 1$. Predicted values $\hat{x}$ and $\hat{s}$ are fitted by regressing x_N and s_N on percentile rank of observed values interacted with exogenous covariates. Some sample covariances use predicted $\tilde{u}$ and $\tilde{v}$ where, $\tilde{u} = w - \hat{x}$ and $\tilde{v} = t - \hat{s}$. Unknown error introduced by imputation $\hat{u}$ and $\hat{v}$ are defined in (1.16). GLM results are not interpretable as $\sigma_{x\varepsilon} \neq 0$ implies the GLM model is mis-specified.

Chapter 2

The Medical Care Costs of Youth Obesity

2.1 Introduction

Over the last 50 years, the prevalence of youth obesity has significantly increased. In 1965, under 5% of US children and adolescents aged 2-19 years were obese. As of 2012, the prevalence of youth obesity is an estimated 16.9%, with significantly higher rates of obesity among Hispanic and non-Hispanic black children and adolescents (NCHS 2010; Ogden et al., 2012). Further, about 33% of boys and 30.4% of girls are considered to be overweight or obese.[1]

This prevalence of overweight currently exceeds 40% among black and Hispanic adolescents (Ogden et al., 2012). Simply being overweight or obese increases their likelihood of type 2 diabetes, gallbladder disease, sleep apnea, joint problems, and cardiovascular risk factors during childhood and adolescence (Han et al., 2010; Guo and Chumlea, 1999; Dietz and Robinson, 2005; Ogden et al., 2002). In addition to physical consequences, being overweight or obese is also associated with negative self-image, low self-esteem, and behavioral and learning difficulties (Dietz, 1998). Overweight children are more likely to become overweight and obese adults (Biro and Wien, 2010), and children who are over-

[1]Weight status for children is determined by percentile rank in gender-age specific CDC growth charts. Table 2.1 contains youth BMI percentile cutoffs. Table 2.2 contains the prevalence rates of obesity from 1963 to 2010 calculated from the National Health and Nutrition Examination Survey.

weight are more likely to have coronary heart disease, type 2 diabetes, and other serious health problems (NIH, 1998).

Previous research has assessed the short-term costs of youth obesity[1]. Finkelstein and Trogden (2008) found that, on average, obese children and adolescents incur $220 more in medical spending than normal weight children and that overweight children incur $180 more than normal weight children. Monheit et al. (2009) estimated separate models by gender and found that adolescent girls who become obese cost $790 more per year than normal weight girls. They found no significant effect for boys. These studies suggest that the short term return-on-investment to youth obesity interventions is small, and the financial benefits of targeted interventions only exceed the costs by incorporating the expected future costs of adult obesity (Finkelstein and Trogden, 2008).

The costs associated with youth obesity can be seen in higher utilization of hospital care, as well as less acute care in outpatient settings. Trasande et al. (2009) estimated that childhood obesity was responsible for $237.6 million in hospitalizations in 2005, up from $125.9 million in 2001. Trasande and Chatterjee (2009) used the 2002 – 2005 MEPS to estimate the impact of youth obesity on health care utilization. They found that overweight and obese children aged 6 to 19 years old had higher utilization of inpatient and outpatient care, as well as higher prescription drug expenditures. They also found that the increases in costs and utilization were concentrated among adolescents. They aggregate these additional expenses to attribute $14.1 billion annually in additional medical care costs to elevated BMI in children.

It is of both academic and policy concern whether these additional costs are borne by the individuals and families whose health and diet behaviors can influence their weight status, or whether third parties such as public or private insurance plans bear the additional costs of obesity. Children with low socioeconomic tend to have higher BMI on average, and are more likely to be enrolled in public programs like Medicaid (Freedman et al., 2006), which may lead to public programs covering a larger share of medical care costs associated with youth obesity. For instance, the cost of additional hospitaliza-

[1]For an overview of studies estimating not only the short-term costs of childhood obesity, but the long-term costs and economic impact, see Trasande and Elbel (2012) and Pelone et al. (2012)

tions associated with obesity related conditions among children and adolescents was by Trasande et al. (2009) found to be paid mostly by Medicaid, even though private payers cover more obesity treatments.

If families do not face the financial burden of poor health behaviors, then health insurance may, in effect, subsidize obesity (Bhattacharya and Sood, 2005). Bhattacharya and Sood (2005) show that the external costs of obesity (costs not faced by the obese individual) can represent a negative externality if premiums are not risk-adjusted to reflect the weight of the enrollee. Families most prone to obesity face the smallest out-of-pocket share of medical costs for conditions associated with youth obesity due to public insurance programs such as Medicaid and SCHIP. If these public programs bear the larger share of the external costs of obesity, even though they cover a small share of the US population, then these costs may represent a loss in societal welfare.

A major limitation of these prior studies is that they only estimate the association between obesity and youth medical expenditures, and not the causal effect, as they do not explicitly account for possible endogeneity of weight or measurement error in body weight and height. As a result, these studies likely underestimate the true causal effect. Weight is endogenous because it is correlated with unobserved socio-economic status (SES) or access to care, both of which affect medical spending. Families with lower SES are more likely to be obese and have higher incidence of poor health; have unobservable health problems; or engage in other risky behaviors (Fontaine and Bartlett, 2000). However, due to poor access to care, these families may have lower expenditures on medical care (Burkhauser and Cawley, 2008). To reconcile the endogeneity of child weight, we use the BMI of a biological relative to instrument for child BMI (Cawley, 2004). Prior research has used the weight of biological relatives as a source of exogenous instruments. For example, Cawley and Meyerhoefer (2012) employ restricted-use biological linkage variables in the MEPS to match parents to their biological children in order to estimate the impact of obesity on adult medical expenditures. We use the BMI of each child's biological parents as an IV for the child's BMI.

In data sets drawn from surveys, in which height and weight are reported and not

measured, it is plausible to observe random additive or classical measurement error as well as measurement error that is non-classical. For example, a number of studies find that individuals tend to under-report their own weight, and that this under-reporting is correlated with their own BMI (Gillum and Sempos, 2005; Rowland, 1990). This type of misreporting can invalidate measurement error models only suitable for random additive error (Villanueva, 2001; O'Neill and Sweetman, 2013). In prior studies of the impact of youth obesity, child BMI is usually derived from parental proxy-reports. Parental self-reports as well as their proxy-reports of child height and weight are likely mis-measured. If this mis-reporting is not random, then measurement error can introduce asymptotic bias, even when using instrumental variables (O'Neill and Sweetman, 2013). In the absence of repeat measures of BMI and correctly measured instruments, we can use validation data to correctly identify the true impact of youth obesity, even in the presence of non-classical measurement error (Bound et al., 2001). We correct for measurement error in child and parent BMI by using the National Health and Nutrition Examination Survey (NHANES) as a source of measured BMI data. The NHANES contains both measured and self-reported heights and weights and can be used to correct reported BMI in the MEPS (Cawley, 2004).

After correcting for reporting error and endogeneity of weight, we find that medical expenditures are significantly higher for children, and in particular, girls aged 11-17 who are overweight and obese relative to those who are healthy weight. Our estimates are larger in magnitude than those from previous studies that do not use instrumental variables. This suggests that endogeneity of weight and measurement error have a large role in biasing non-IV estimates of the impact of obesity on child medical expenditures. Further, previous studies have likely underestimated the cost of youth obesity and in turn, the cost effectiveness of interventions that target youth obesity before they reach adulthood. We also find evidence that the costs of youth obesity are borne mostly by third party-payers.

2.2 Identification Strategy

2.2.1 Instrumental Variables

We estimate the effect of BMI on medical expenditures using the IV-GLM procedure described in Carroll et al. (2006). In order for IV estimation to produce consistent estimates, the instrument must be sufficiently correlated with the endogenous or mismeasured variable. Our first stage partial F-statistics range from 697.44 - 2,064.07, well in excess of the rule-of-thumb F-stat = 10 for all of our specifications (Stock et al., 2002). More difficult to establish is that the instrument is independent of the error in the structural model.[1]

Independence would be violated if parents and their children's BMI are affected by common environmental factors that also influence the child's medical care costs. Parents are typically responsible for their children's healthcare decisions. Parenting decisions may stem from parent attitudes towards health, and may be reflected in the parent's BMI. If these attitudes are correlated with unobserved cultural or environmental factors that also effect the child's medical expenditures, then the instrument will be correlated with the error term in the model. This cannot be directly tested in our data.

There is however, a substantial behavioral genetic literature validating the genetic relationship between the weight of biological relatives. (Haberstick et al., 2010; Smith et al., 2009; Wardle et al., 2008; Grilo and Pogue-Geile, 1991) For example, Grilo and Pogue-Geile (1991) find that 40–70% of the variation in obesity-related phenotypes, such as body mass index, skinfold thickness, fat mass and leptin levels, is inheritable. Adoption studies have found that the strong genetic correlation between children and their biological parents is not weaker for children raised by adoptive parents (Stunkard et al., 1986; Srensen and Stunkard, 1993). Twin studies also find that the correlation in the weight of twins does not depend on whether they were raised together or separately. (Price and Gottesman, 1991; Maes et al., 1997)

This same behavioral genetics literature finds little to no evidence in favor of shared

[1]Independence is a stronger assumption than for linear IV, but is required for IV-GLM (Carroll et al., 2006).

household or environmental effects on BMI (Haberstick et al., 2010; Wardle et al., 2008; Maes et al., 1997; Grilo and Pogue-Geile, 1991). Taken at face value, no shared household effect of BMI implies that a parent's BMI can only effect their child's medical expenditures through it's effect on child BMI. We interpret this lack of evidence for environmental effects not as a statement that location, shared culture, or shared diet do not effect child BMI, but that all of the correlation between parent BMI and these factors is captured by the strong correlation between parent and child BMI due to the strong influence of shared genetics.

Prior economic research has employed genetic variation in weight to generate instrumental variables. Cawley (2000) use the National Longitudinal Survey of Youth (NLSY) in order to measure the effect of body weight on employment disability. He uses the BMI of a biological child to instrument for adult BMI. Kline and Tobias (2008) use the BMI of parents to instrument for the BMI of their children and estimate the impact of BMI on wages in Britain. Lindeboom et al. (2010) estimate the effect of obesity on employment, using rich data from the British National Child Development Study (NCDS). The results show a significant negative association between obesity and employment even after controlling for a rich set of demographic, socioeconomic, environmental and behavioral variables. Cawley (2004) uses the BMI of a sibling to instrument for adult BMI in estimating the impact of obesity on wages. More recently Cawley and Meyerhoefer (2012) and Cawley et al. (2014) use the BMI of children to instrument for the BMI of their biological parents in the MEPS in order to estimate the impact of obesity on adult medical expenditures.

2.2.2 Proxy-Reporting Error in BMI

Parents in the MEPS are responsible for not only reporting their height and weight, but proxy-reporting their childrens' height and weight. If parental mis-reporting behavior is correlated with their own weight, then the strong correlation between the BMI of parents and children may results in correlation between parents BMI and the measurement error in their child's BMI. This correlation will result in biased IV estimates of the

impact of BMI on medical expenditures.

Prior research has documented that adults under-reported their own weight (Shiely et al., 2013; Gorber et al., 2007; Villanueva, 2001). A number of small studies have tried to determine the manner in which parents report their children's weight and height. Studies of parental reporting of adolescent BMI show slight over-reporting of height and under-reporting of weight, resulting in under-reports of BMI (Brettschneider et al., 2012; O'Connor and Gugenheim, 2011; Goodman, Hinden, and Khandelwal, Goodman et al.; Reed and Price, 1998). For example, O'Connor and Gugenheim (2011) surveyed parents of children 2 to 17 years old at an outpatient orthopedic clinic, asking them to report their child's height and weight prior to having them measured. They found that mean weight error in parental reports increased with child age and with age-specific child BMI z-score. Parents tended to under-report their child's weight, leading to 21 percent of children measured as obese being misclassified as not-obese in parental reports.

In their meta-analysis of parental proxy reporting studies, Lundahl et al. (2014) find that 14.3% of parents underestimate the weight of their normal weight children, and that over 50% of parents underestimate the weight of their overweight and obese children. They find that parents are more likely to underestimate child weight in the overweight/obese range, and that parents become more accurate estimating the weight of children with much larger BMI. This indicates that children with a BMI-for age percentile just over the cutoff for overweight were at greater risk for being misclassified by parents than those with a much higher BMI. The higher incidence of mis-classification around the thresholds may be due to higher density of observations compared with the tails of the BMI distribution. They speculate that the improvement in the accuracy of parental reporting for very high BMI children may be due to increased measurement since BMI is a more salient factor in healthcare decisions. This trend of improved reporting for children with high BMIs does not confound the negative correlation between child BMI and measurement error across the entire BMI distribution. They also found that parents were more likely to under-report normal weight boys, than normal weight girls. Importantly for our identification, they found that misreporting did not vary across year

or the country/state of the study.

Parental under-reporting of their children's weight may be intentional if parents know the true weight of their child, but choose to misreport it, possibly due to stigma associated with obesity (Puhl and Heuer, 2010; Latner et al., 2005). Parents who intend to be accurate may still incorrectly estimate their children's weight due to their own perception of what an obese child looks like, or what an obese child's lifestyle is compared to their child. Another possibility is that parents are unaware of their child's true height and weight, and respond with the most recent measurements they recall or inaccurate measurements. This may cause random measurement error, but may result in under-reports as children's height and weight tend to rise over time regardless of obesity status. In either case, the consequence of under-reporting is a negative correlation between the child's BMI and the proxy-reporting error in child BMI. This negative correlation, combined with positive correlation between parent and child BMI does not imply a negative correlation between parent BMI and the reporting error in child BMI, regardless of the reason parents misreport their children's weight.[1] Were parent BMI to be correlated with the measurement error in child BMI, then using parent BMI as an IV for child BMI would not result in consistent estimates of the impact of child obesity on medical expenditures.

2.2.3 Bias Due to Measurement Error

Systematic misreporting of a covariate is a kind of omitted variables problem that is most easily analyzed using simple linear regression. Like random-additive error, under-reporting will lead to biased OLS coefficients (Bound et al., 2001). This can be demonstrated using the probability limit of the beta coefficient in a bivariate linear IV model[2].

Consider a true linear model:

$$Y = \beta x + \varepsilon, \tag{2.1}$$

[1]Covariance (and correlation, which is normalized covariance) is not transitive. (Langford et al., 2001)

[2]See Bound et al. (2001) for a more detailed description.

such that $E(x\varepsilon) = 0$. We cannot observe X directly in the data, and instead observe w such that,

$$w = x + u. \tag{2.2}$$

Thus, the true data X is observed with some additive error term u. If $E(u) = 0$ and $Cov(x, u) = 0$, any variance in u causes an attenuating bias in OLS estimates of β. If $Cov(x, u) \neq 0$, then OLS is still biased towards zero (more significantly if $Cov(x, u) > 0$).

We introduce instrument S such that,

$$x = \gamma s + \zeta \tag{2.3}$$

Where $\gamma \neq 0$ and ζ has mean zero.

IV estimation will consistently estimate β as long as S is uncorrelated with u and ε, and is sufficiently correlated with X. We assume $Cov(s, \varepsilon) = 0$ and $Cov(s, x) \neq 0$ and focus on the relationship between the IV and the measurement error in x. Typically instruments are also measured with error such that,

$$t = s + v. \tag{2.4}$$

The probability limit of β_{IV} can be expressed in terms of covariances, as

$$plim\beta_{IV} = \frac{Cov(t, Y)}{Cov(t, w)} = \frac{\beta Cov(t, x)}{Cov(t, w)}. \tag{2.5}$$

Changing notation and substituting equations (2.2) and (2.4) into equation (2.5) we can derive,

$$plim\hat{\beta}_{IV} == \frac{\beta(\sigma_{xs} + \sigma_{xv}) + \sigma_{s\varepsilon} + \sigma_{v\varepsilon}}{\sigma_{xs} + \sigma_{xv} + \sigma_{su} + \sigma_{vu}}. \tag{2.6}$$

Although all the covariance terms in the probability limit depend on unobservable relationships, IV estimation will be robust to non-classical error in both x and s. The non-linear instrumental variables models used in this analysis incorporate linear first stages. The second stage performs a GLM fit of the outcome y using the fitted values for X from the first stage.

2.2.4 Corrected Body Mass Index

OLS and IV estimates are not robust to all forms of measurement error. Differential measurement error or correlation between the true instrument and the measurement error in the endogenous variable may still bias IV estimates. It is possible to reduce the magnitude of these biases by using validation techniques to replace mismeasured variables with error-corrected fitted values (Bound et al., 2001). Using mismeasured heights and weights will also lead to incorrectly estimating the share of the population of children who are overweight or obese. Prior studies have used measured validation data to correct for measurement error in BMI when estimating the impact of BMI on health and labor outcomes (Cawley, 2000, 2004; Dutton and McLaren, 2014). Validation data are typically used to generate correction equations that are used to predict true BMI in the principle sample, which only contains mismeasured BMI. (Courtemanche et al., 2014). We follow the imputation procedure described in Courtemanche et al. (2014) using the National Health and Nutrition Examination Survey (NHANES) as a source of validation data.[1] A summary of the imputation procedure in Courtemanche et al. (2014) in the context of our model follows here.

Imputation requires there is a surrogate, or matching agent, present in the principal and validation datasets, as well as the transportability of the surrogate across both datasets. Each observation has true BMI X and observed BMI w_j in the principle sample $j = M$ and in the validation sample $j = N$.[2] the first condition is that for true BMI x, there exists a surrogate (typically the observed BMI) w_j^{sur} such that the distribution of the outcome Y given (x, w_j^{sur}) is the same as Y given x. Essentially the observed BMI cannot contain information about the outcome that is not already reflected in true BMI. Another interpretation is that measurement error in BMI cannot be correlated with unobserved variables that influence the outcome. The second condition is the transportability of the surrogate; that the underlying distributions of true BMI in both datasets are equal

[1] We follow the arguments of Lee and Sepanski (1995) and impute BMI directly instead of predicting height and weight and using the predictions to calculate BMI. The advantage is we are directly predicting the BMI distribution, and in turn the non-linear relationship between BMI and expenditures.

[2] In our analysis, the MEPS is the principle sample and the NHANES serves as the validation sample.

conditional on the surrogate. Transportability implies that,

$$E(x|w_M, Z_M) = E(x|w_N, Z_N),^{1} \tag{2.7}$$

where covariates (Z_j) are shared in both datasets (gender, age, race, etc...).

In prior research using the NHANES as validation data for self-reported adult BMI, the self-reported BMI in the NHANES is used as the surrogate for imputation (Cawley, 2000, 2004). We can only follow this method to impute adult BMI for adults who self-report their own height and weight, which will be used to generate their imputed BMI. In practice, mothers and fathers may have different self-reporting behavior, which may vary by age or race. In order to maintain transportability, we focus on children whose mothers are the primary survey respondent, and self-report their own height and weight, and then control for age and race in the imputation steps.

Even after restricting the sample, we are unable to use proxy-reported BMI as the surrogate for children because the NHANES does not include proxy-reported height and weight for children, as it only contains measured heights and weights. To impute child BMI, as well as adult BMI, we follow the method in Courtemanche et al. (2014), and use the percentile rank of BMI as the matching surrogate. In this way, misreporting manifests not only as a stochastic process on additive error terms u and v, but as a shifting in the distribution of BMIs. Using the percentile rank as the matching surrogate only requires that the expected value of true BMI conditional on reported BMI is monotonically increasing in reported BMI (Courtemanche et al., 2014). This monotonicity implies that individuals who report higher BMI are expected to have higher a BMI than those who report report lower BMI. Courtemanche et al. (2014) test this monotonicity assumption in the NHANES, and do not find evidence to reject it.

The percentile rank is a transportable surrogate if,

$$E(x|BMI\ Rank_M, Z_M) = E(x|BMI\ Rank_N, Z_N). \tag{2.8}$$

[1]This is known as weak transportability (Lee and Sepanski, 1995).

For both adults and children, we predict measured BMI in the NHANES as a function of gender, race, and their interactions with linear splines generated from the percentile rank of BMI and higher orders of age (age in months for children). We then use the estimated equations to predict true BMI for adults and children in the principal sample (MEPS).

2.3 Data and Empirical Model

2.3.1 Data

The Medical Expenditure Panel Survey (MEPS) is a comprehensive, nationally representative survey of the U.S. civilian non-institutionalized population. In the MEPS, families are surveyed five times during a two year period about their medical care utilization and expenditures. For each family (the responding unit in the MEPS), a single individual is the primary respondent. For most families in the MEPS, the mother is the primary respondent. In two parent households, the second most common primary respondent is the father. We can identify the primary respondent in each family (usually a parent) and we use restricted-use biological linkage variables to match parents to their biological children. Heights and weights are not measured in the MEPS. The primary respondent typically reports the heights and weights of everyone in the reporting unit. This means that the primary respondent self-reports her height and weight, and heights and weights for her spouse and children are generated from proxy-reports.

We use data from the 2000-2010 household component of the MEPS and inflate all expenditures in each year to 2010 dollars. We limit the sample to households with biological children younger than 18 and older than 11, and mother's between 20 and 64; both with non-missing BMI. We do not incorporate children younger than 11 due to high rates non-response for height and weight. We eliminate children whose parents had BMI in excess of 80 and below 10 (2,152 observations). We finally eliminate 1,590 underweight children from the sample. The resulting estimation dataset has 27,002 children aged 11-17 with mothers who were the primary-respondent and proxy-reported the child's BMI.

In the MEPS, medical expenditures and the source of payment are collected directly from households as well as from the households medical care providers for every medical event. In addition to total medical expenditures, we estimate the impact of obesity on expenditures by all third party payers (typically, public and private insurers), and also expenditures by all payers on specific categories of care: inpatient, outpatient, prescription drugs, and other (which includes dental, vision, home health care services, and medical equipment but excludes spending on over-the-counter medications).

MEPS data are collected through a stratified multi-stage probability design, which we account for in the calculation of the standard errors for our marginal effects. In particular, we use the method of balanced repeated replications to estimate standard errors in our non-IV and IV two-part GLM models. This method accounts for clustering at the PSU-level, stratification, and weighting.

We correct for reporting error in BMI by using the National Health and Nutrition Examination Survey (NHANES) as a validation dataset. The NHANES is a nationally representative survey of adults and children that combines interviews and physical examinations (CDC / National Center for Health Statistics, 2014). In the NHANES, adults and children aged 16 or older self-report their weight, and all survey respondents have their heights and weights measured[1]. The NHANES does not contain medical expenditures, but does share with the MEPS a rich set of covariates.

The continuous version of the survey data are released in two-year waves (survey-cycles), beginning in 1999. We construct the validation dataset by appending six survey cycles of the data, from the 1999-2000 survey cycle to the 2009-2010 survey cycle. Like the MEPS, the NHANES provides weights to account for the complex survey design (including oversampling), survey non-response, and post-stratification (CDC / National Center for Health Statistics, 2013). We construct the appropriate survey weights for a 12 year span of the data (1999-2010).[2]

[1]Survey participants are not aware that they will be weighed until after they self-report their height and weight.

[2]We extrapolate the method to construct weights when combining survey cycles for 12 years of data using the 10 year method in task 2 in CDC / National Center for Health Statistics (2013). Our code is displayed in Appendix A.

2.3.2 Model Specification

As with adult medical expenditures, medical spending on children is highly positively skewed with a substantial number with zero expenditures in any survey year (Monheit et al., 2009; Finkelstein and Trogden, 2008). To account for the shape of distribution of expenditures we employ a two-part model of medical expenditures (Jones, 2000). The first part is a Logit model that estimates the probability of having positive expenditures. The second part estimates the level of medical expenditures conditional on having positive spending and is specified as a a GLM with Gamma variance structure and log link. We conduct the specification tests suggested by Manning and Mullahy (2001) to identify the proper link function and distribution for our data.[1] We also perform a modified Hosmer-Lemeshow test by regressing prediction error from each model on deciles of the distribution of predicted expenditures. We fail to reject the null-hypothesis that the decile coefficients are jointly equal to zero, indicating the choice of distribution and link function is appropriate.[2]

Both parts of the model include child characteristics: gender, race/ethnicity (white, black, Hispanic, other race), child age in months, education levels for both parents (no high school diploma, high school graduate, some college, bachelors degree or higher) and household characteristics: census region (northeast, midwest, south, or west), whether the respondent lives in an MSA, household composition (number of household members age 0-5 years, 6-17 years). We acknowledge that income can partly determine child medical expenditures, but we do not include income in the model as it is likely endogenous. We also exclude insurance status or coverage type due to similar endogeneity concerns. Measures of both parent's education serve as proxys for parental SES. We do not control for parent race as it is highly co-linear with child race. We also control for whether the child self-reported their height and weight, although fewer than one percent of children self-reported their BMI when their mother was the primary respondent.

We implement both parts of the model as Generalized Linear Models (the first stage is

[1] We conduct Park tests to confirm our choice of conditional variance, in particular that the variance is proportional to the square of the conditional mean. We find $\lambda = 1.87 - 1.99$ across our samples.

[2] Monheit et al. (2009) tested alternate specifications on the 2001-2003 MEPS and selected a two-part model with a Probit first stage and a Log-Gamma second stage.

a GLM using the binomial distribution and a Logit link.). In this way we can incorporate instrumental variables into both parts of the two-part model using the IV-GLM estimator proposed by described in Hardin and Carroll (2003a).[1] Their method is essentially two-stage nonlinear least squares. The first stage is a linear regression of the mismeasured covariates on the set of included and excluded instruments. The second stage is a GLM fit of the outcome on the known covariates and the fitted values of the mismeasured covariate from the first stage[2] (Hardin and Carroll, 2003b).

It is possible to estimate models using the z-score of child BMI as the explanatory variable. Child weight status is determined using the z-score of BMI, however, it is difficult to interpret the weight change associated with a unit change in the z-score. We use child BMI as the measure of youth obesity and include controls for child gender and age in months. BMI is an imperfect measure of body fat since it does not distinguish fat from fat-free mass such as muscle and bone (Burkhauser and Cawley, 2008; ONeill, 2015), but BMI is the only measure of fatness in the MEPS.

We do not estimate marginal effects using binary indicators of obesity. Instrumenting for binary measures of obesity can cause biased estimates in the the first stage regression. We can only estimate bounds for the true effect when instrumenting for a binary indicator derived from reported child BMI that is both endogenous and non-classically mismeasured (Frazis and Loewenstein, 2003). To estimate the average marginal effect of moving from normal weight to overweight or obese, we alternatively predict expenditures as the mean BMI within each category, and take the difference in predicted expenditures.

2.4 Results

2.4.1 Descriptive Statistics

Descriptive statistics for the model covariates are presented in tables 2.3 and 2.4. Mean expenditures are reported only for children with non-zero expenditures (around

[1]For a thorough treatment on non-linear measurement error models, see Chen et al. (2011).

[2]We estimate the IV-GLM in STATA using the **qvf** command in (Hardin et al., 2003). The command fits the GLM using iteratively re-weighted least squares (IRLS). Hardin and Hilbe (2012) (2012) detail the steps of the IRLS algorithm.

88% of children). Boys (girls) aged 11-17 have on average $2,098.93 ($2,264.74) in medical expenditures a year from all payers. These magnitudes are smaller than the mean expenditures for adults (From chapter 3). Although average expenditure is slightly higher for girls than for boys, other covariates are very similar across gender.

Table 2.5 contains descriptive statistics for boys and their parents' BMI and Table 2.6 shows the same descriptive statistics for girls. Comparing the mean reported BMI in the MEPS to the imputed measured BMI from the NHANES suggests that BMIs in the MEPS are under-reported. For example, the mean reported BMI of girls in the sample is 22.18. The mean imputed BMI from the NHANES is higher at 22.49. The same relationship exists for mothers, whose self-reports are lower, on average, than their imputed measured BMI levels. The magnitude of under-reporting is larger for obese children than for overweight and healthy weight children. This indicates that there is some correlation between child BMI, and the degree to which child BMI is under-reported. Tables 2.5 and 2.6 also show the mean mothers' BMI, and indicate similar under-reporting on average.

2.4.2 Proxy-Reporting Error

Figure 2.1 compares the distribution of self-reported BMI in the NHANES to the measured BMI of adult females. The self-reported distribution is shifted to the left of the measured BMIs, suggesting that mothers under-report their own BMI. A necessary assumption to impute measured BMI from the NHANES into the MEPS is that the misreporting behavior is the same in both samples. Only primary respondents in the MEPS self-report their data, the large majority of which are women. Figure 2.2 shows the distribution of self-reported BMI in the MEPS and measured BMIs imputed from the NHANES. The self-reports in the MEPS are similarly distributed to those in the NHANES. When measured BMI is imputed, we observe similar reporting as in the NHANES. There is less mass in the distribution of measured BMI relative to self-reported BMI for BMI > 30. There is greater mass in the distribution of self-reports over the normal weight and overweight range of BMI.

Figure 2.3 shows the distribution of child BMIs proxy-reported in the MEPS and the distribution of measured child BMI in the NHANES. Mothers' misreporting of their children's height and weight results in under-reporting of BMI similar to that caused by mothers' self-reports. There is greater mass in the distribution of imputed measured BMI above BMI = 25 up to approximately 40 and less mass below BMI of around 23. The standard BMI cutoffs for weight status do not strictly apply for children, whose BMIs tend to be lower than adults. Thus misreporting that over-represents the number of children with BMI under 20 may be indicative of the same over-representation of normal weight individuals that we observe in mothers' self-reports.

We estimate the sample covariances between measurement error in reported BMI and imputed measured parent and child BMI by calculating the error terms in equations (2.2) and (2.4). This allows us to indirectly observe how mothers' misreporting can generate measurement error in child BMI that is correlated with both the child's and parent's BMI. Figure 2.4 plots the additive error in boys' proxy-reported BMI against their imputed measured BMI. There is a moderate relationship ($\rho = -0.40$) between boys' BMIs and the degree to which parents misreport their height and weight. At lower imputed BMIs the proxy reporting error is mean zero, turning negative as imputed BMI increases. There is a greater portion of boys' BMIs that are under-reported than over-reported. Figure 2.5 separates the plot by age. Although the 11 and 13 year old boys have positive correlation coefficients, The gradient of under-reporting also becomes steeper as boys become older.

Figure 2.6 plots the additive error in girls' BMI values against their imputed measured BMIs. The negative correlation is slightly stronger ($\rho = -0.48$) than for boys, although far more observations are under-reported. When we split up observations by age in figure 2.7 we observe some variation in parental reporting by child age, but a consistent downward gradient due to under-reporting. Interestingly, the degree of misreporting approaches zero at very high BMI levels for girls in some age groups. This finding is consistent with Lundahl et al. (2014), who observe in their meta-study that parents reporting accuracy improves as child BMI increases. One possible explanation is that parents are more acutely aware of their child's weight because it is significantly higher

than average, or because their children have weight-related comorbidities that increase their contact with the healthcare system. It may be that this salience of higher BMI is associated with more honest reporting; either by resolving lack of information or by mitigating the effect of stigma associated with relatively high BMI.

2.4.3 Impact of Youth Obesity on Medical Expenditures

We begin by estimating a non-IV two-part model with the same specification as our main model. The results are presented in the left column of Table 2.7. Each estimated marginal effect represents a one unit increase in BMI on annual medical expenditures. From our non-IV model we find that, on average, an additional BMI unit raises medical care costs by $17.45 per year, however this effect is not precisely estimated.

To estimate the impact of a normal weight child becoming obese (overweight), we use the model to compute the average marginal effect of moving from the mean BMI among normal weight children to the mean BMI among the obese children (overweight)[1]. (Cawley et al., 2014) Table 2.8 presents the effects of moving from normal weight to overweight and moving from normal weight to obese for the models estimated in columns 1 - 4 in table 2.7. When not instrumenting the magnitudes are small and insignificant, suggesting that the cost curves associated with BMI are flat and close to zero.

Column 2 presents the coefficient estimates using the mother's self-reported BMI as the IV for the proxy-reported BMI of their child. Compared to the non-IV estimates, the effect sizes are much larger, and statistically significant for the total sample, and for the subsample of just girls. These estimates use the same model as Cawley and Meyerhoefer (2012), and suggest that the marginal effect of a BMI increase for children ($92.66) is significantly lower than for adults ($149), which is reasonable given children have lower incidence of hospitalizations ad chronic illness.

Columns 3 and 4 contain estimates from the RCIV and ARCIV estimators described in Chapter 1. The RCIV estimator uses predicted child and adult BMI in place of the mismeasured values from the MEPS. The RCIV estimates are smaller than the IV esti-

[1]The dataset from which the child BMI is observed is used to determine child weight status.

mates. This is consistent with non-classical measurement error in child BMI, which can cause downward bias in the RCIV estimator. Column 4 Contains the ARCIV estimator in which only the instrument is replaced with it's fitted value. The ARCIV estimates are in between the IV and RCIV estimates. The simulation results in Chapter 1 suggest that the IV and RCIV estimates should be very similar.

In the simulation results from chapter 1, we can use the sample covariances to explain the difference between the RCIV and the ARCIV estimates when ARCIV is consistent. Column 5 in table 2.7 contains the ARCIV estimates multiplied by δ from (1.26). We find that this relationship holds as the estimates in column 5 are very close to the RCIV estimates in Column 3. When both BMI and child BMI are systematically under-reported, the RCIV estimator will be biased downward. We do find evidence of under-reporting in both adult and child BMI in the MEPS, and conclude that the RCIV estimator is biased downward.

In simulations, the ARCIV estimator should be very similar to IV. We observe that they are similar, although the ARCIV estimates are slightly smaller. It is not possible to know if this difference is due to improved consistency by using ARCIV, or if the difference is attributable to finite sample size. If imputation improves the estimators performance under some unknown measurement error, then ARCIV is the better estimator. If IV is consistent, then the estimated coefficients using ARCIV are slightly conservative estimates of the true effect. For these reasons, and because we can used sample covariances to explain the bias in the RCIV estimator relative to the ARCIV estimator, we use the ARCIV estimator as our preferred estimator.

We find that, on average, an increase of one BMI unit is associated with a \$87.39 increase in medical expenditures. For boys (girls), this effect is \$98.60 (\$92.18) for each BMI unit. We use the estimated models to calculate average marginal effects of moving from normal weight to overweight and obese on medical expenditures. We find that, on average, overweight children incur \$409.74 in additional expenditures relative to normal weight children, and boys (girls) who are overweight incur \$419.43 (\$483.78) in additional expenditures relative to normal weight girls. Moving from normal weight to obese

increases boys (girls) annual expenditures by $1,163.93 ($1,273.28).

2.4.4 Economic Burden of Youth Obesity

To better understand the impact of youth obesity on medical expenditures we estimate separate models by payment type. Child medical expenses are often covered by their families private health insurance plan (such as an employer sponsored plan), or a public insurance plan such as Medicaid or CHIP. Table 2.9 shows the results from the preferred model estimated separately for expenditures paid out-of-pocket, and for expenditure paid by any third party. We find that a one BMI unit increase is associated with a $102.56 increase in third-party expenditures. This is larger than effect of a one BMI unit increase on total expenditures, despite the fact that third-party payments are a subset of total expenditures. We find a negative effect of a BMI unit increase on out-of-pocket payments. These complimentary effects sum very close to the effect BMI on total expenditures. The effect of a one unit increase in BMI on boys' third-party expenditures is $123.96. However, the effect of BMI on boys' expenditures paid out-of-pocket is $-31.38. For girls, the effect of BMI on third-party payments is $102.03, with a less precisely estimated negative effect of BMI on out-of-pocket payments of $-17.61. Table 2.10 contains the estimated marginal effects of overweight and obesity estimated from the models in Table 2.9.

One possible interpretation for the negative effect of elevated BMI on out-of-pocket expenditures is that becoming overweight or obese shifts costs onto third parties. Obesity related comorbidities are associated with higher levels of care, such as inpatient stays or heavy utilization of outpatient facilities. High utilization consumers of healthcare are likely spending beyond their deductibles, where levels of cost-sharing are lower relative to those spending their first dollars, even though their total dollar spending is higher. Obese children may be receiving a greater share of their care in higher cost environments. For example, a healthy weight child with joint problems may receive regular out-patient care, whereas an obese child with joint problems is more likely to have a severe event such as a sprain or muscle tear and end up in receiving more expensive treatment in an emergency department or hospital floor. Insurance design typically discourages individuals from

substituting hospital care for other types of care in order to lower their out-of-pocket costs by taking advantage of more generous cost-sharing for inpatient care. Yet, there is still considerable heterogeneity among plan enrollees, and obese individuals in any plan are likely receiving higher levels of care, potentially at more generous rates of coverage.

Another possible explanation is that the rates of overweight and obesity are higher among Medicaid enrollees compared to private insurance. We do not control for insurance status as it is endogenous in the main model. Thus the negative effect may be capturing the difference in average cost-sharing across insurance types, which is confounded with the higher rates of obesity among Medicaid enrollees. Omitting insurance status from the model may be directly contributing to a spurious correlation between higher BMI and lower out-of-pocket costs. We test this possibility by including indicator variables for coverage by private insurance and by Medicaid. We find that after including controls for insurance status, the negative coefficients for out-of-pocket expenditures are smaller and only strongly significant for boys.[1] There is a corresponding drop in the size of the effect of BMI on third-party expenditures, reinforcing that the total impact of BMI can be expressed as the sum of the the out-of-pocket and third-party expenditures. There is still considerable heterogeneity in underlying health across enrollees with the same type of insurance coverage, and there may still be some negative effect of obesity on out-of-pocket costs due to higher utilization of more intense care among overweight and obese children.

Given that third parties seem to bear most of the medical care costs associated with youth obesity, it is of interest to determine what share is paid by public payers as compared to private insurance. Table 2.11 contains estimates of the impact of increased BMI on expenditures paid by Medicaid among Medicaid enrollees and expenditures paid by private insurance plans among those enrolled in private insurance. We find that a one BMI unit increase causes the level of expenditures paid for by Medicaid to increase by $145.51. When we restrict the sample to only those enrolled in Medicaid, we find that a one BMI unit increase raises boys' medicaid expenditures by $135.64, and girls' expen-

[1] We also estimate these models including the Log of family income, and find no discernible difference whether income is included in the model or not.

ditures by $266.48. We find no effect on the level of expenditures paid for by private insurance among those enrolled in private insurance plans. Table 2.12 contains estimates of the impact of overweight and obesity on expenditures paid by Medicaid among Medicaid enrollees and expenditures paid by private insurance plans among those enrolled in private insurance. Moving from healthy weight to overweight or obese significantly increases both boys' and girls' annual medical expenditures. The estimated effects for girls are again larger than those for boys. The estimated effects of overweight and obesity on the privately insureds' third-party payments are small and imprecisely estimated. Taken with the results in Tables 2.9 and 2.10, it seems that the costs of youth obesity accrue mostly to third parties, of which public insurance pays the larger share.

We use the IV estimates to estimate the aggregate cost of youth obesity. Table 2.13 contains estimates of annual medical care expenditures (in 2010 USD) associated with youth overweight and obesity, and Table 2.14 contains estimates of the costs just associated with obesity among children aged 11-17 with biological mothers who were the primary respondents.[1] Among this restricted IV sample, we find that from 2000 to 2010, the explicit costs of youth overweight and obesity was $4.75 billion on average. This cost estimate is composed of an average annual $5.64 billion increase in third-party payments, which is offset by an average annual reduction in out-of-pocket payments of $1.16 million. We aggregate the individual effects of overweight and obesity among children enrolled in Medicaid from Table 2.11 for the estimated 2 million children aged 11-17 who are enrolled. We find that the aggregate cost of medical care associated with youth overweight and obesity among Medicaid enrolled children is $3.46 billion, of which nearly all can be attributed to costs associated with only youth obesity.

Under the (admittedly, strong) assumption that the effect of obesity in our subpopulation generalizes to the full non-institutionalized population of adults aged 18 and older, we can scale up the aggregate costs in the IV sample used to estimate our model up to the entire population of children aged 11-17 by multiplying the subpopulation aggregate costs by the ratio of the US population of adults to the US population of adults with biological

[1]Using the MEPS sampling weights, we determined the population of children aged 11-17 was on average 19.97 million children from 2000 to 2010.

children. For example, the total effect of overweight and obesity is $2.68 billion *(30.4 million / 19.97 million) = $7.24 billion.[1] Using the generalized costs, we find that youth overweight and obesity is associated with $5.40 billion a year is medicaid expenditures, of which 96% is attributable solely to obesity.

2.5 Discussion

We estimate the effect of youth obesity on medical expenditures using data from the 2000-2010 Medical Expenditure Panel Survey (MEPS) using instrumental variables. IV estimation will correct for bias due to measurement error in child BMI and endogeneity of weight. IV estimation may still be vulnerable to bias due to non-classical measurement error in the IV, as parents also tend to under-report their own BMI. We correct for reporting error in the instrument by imputing measured BMI values into the MEPS from the National Health and Nutrition Examination Survey. Imputation allows us to indirectly identify the proxy-reporting behavior of parents with regard to their children's BMI. Our indirect observations correspond to the findings from empirical studies in two key ways. First, we find evidence of systematic under-reporting by mothers when reporting their children's weight, which is more pronounced for daughters. Second, we find that parental reporting improves as BMI increases beyond a certain point (as observed in Lundahl et al. (2014).) Maternal under-reporting is consistent with findings in previous research as well as in the NHANES where each adult both self-reports and has their weight measured. We these indirect observations of reporting error as theoretical justification for using instrumental variables estimation with error-corrected instruments.

We estimate a non-IV model and are unable to find significant effects of obesity on child medical expenditures. Finkelstein and Trogden (2008) estimated a two-part GLM model on 2001-03 MEPS with binary indicators for obesity and found that obese children and adolescents incur $220 more medical expenditures than those of normal weight, and overweight incur $180 more in medical expenditures. Monheit et al. (2009) estimated separate models by gender and found that adolescent girls who become obese

[1]The population counts are derived from the MEPS sample and sample weights.

cost $790 more per year than normal weight girls. Unlike these studies, we cannot use the full sample of children in the MEPS, and subset to children with reporting biological mothers due to availability of instruments. Children in single parent households may have different underlying health or access to care that will cause our estimates to differ from prior studies. In all non-IV studies of the impact of obesity, we cannot interpret estimated coefficients as causal effects, and expect that these effects are influenced by attenuating bias due to the endogeneity of weight, as well as measurement error.

Instrumenting using self-reported mothers' BMI in the MEPS dramatically increases the estimated effects of BMI on expenditures. Although IV estimation mitigates bias due to endogeneity and proxy-reporting error in child BMI it may still be biased due to reporting error in parent BMI. WE follow Chapter 1 and use regression calibrated child and parent BMI to estimate RCIV and ARCIV models of medical expenditures. We find that ARCIV estimates are slightly lower than IV estimates, but larger than RCIV estimates. We compute the sample covariance between the fitted instrument and predicted measurement error in child BMI, and show that this covariance is not equal to zero, and is responsible for the RCIV estimator under-estimating the true effect. The results in Chapter 1 show that correlation between the regression calibrated instrument and predicted measurement error in child BMI does not imply that the true instrument is correlated with the true reporting error in child BMI. Taken together, our findings are consistent with non-classical error and under-reporting in both child and parent BMI, but no correlation between the IV and measurement error in child BMI. IV estimation may still be biased if reporting error in the instrument is correlated with the model error term. The ARCIV estimator may reduce this bias, or simply generate the same coefficient estimates as IV. We find ARCIV estimates are slightly lower than those from IV, and interpret them as either the true effect, or more conservative under-estimates compared to the larger IV estimates.

After adjusting for self-reporting error in mothers' BMI, we find that, on average, an increase of one BMI unit is associated with a $87.39 increase in medical expenditures. For boys (girls), this effect is $98.60 ($92.18) for each BMI unit. When we compare medical

costs of obese children to those who are healthy weight, we find that obese children incur $1,091.23 more in medical care costs relative to healthy weight children. These effect sizes are significantly larger than those found in our non-IV model, both overall, and for boys and girls. The effects for boys are notable, as previous studies have not precisely estimated effects of increased BMI on boys' medical expenditures.

We are also the first to estimate the impact of youth obesity on expenditures by payer type. We find that virtually all the increase in medical expenditures caused by youth obesity is paid by third parties; primarily public programs like Medicare. We are the first to find a positive effect of obesity on medical expenditures for boys, and determine the reason previous studies find no effect of elevated BMI on boys' medical expenditures is that increases in expenditures covered by third parties are offset by reductions in out-of-pocket payments. These results have serious implications for policy-makers concerned with the US obesity epidemic. Not only are the costs larger than previously understood, but federal social insurance covers nearly all the costs of care associated with youth obesity.

There are some limitations to our modeling approach. Like previous research using instrumental variables, the validity of the IV depends on an untestable exclusion restriction. There is a large behavioral genetics literature that supports the genetic linkage between the weight of biological relatives, and little support for shared environmental effects that are correlated with weight. But genes which influence weight may be inherited alongside genes that also affect demand for medical care. We cannot observe genetic information in the MEPS and acknowledge this possible limitation. Further, BMI is an imperfect measure of fatness. BMI is strongly correlated with more accurate measures of obesity such as body fat or waist circumference (Burkhauser and Cawley, 2008; ONeill, 2015). However, BMI is the only measure of fatness available in the MEPS and the NHANES.

Chapter 1 shows IV models will be inconsistent under differential measurement error in the IV. Error corrected IV models using fitted values from regression calibration may mitigate this bias by replacing the instrument with its fitted expectation conditional on mismeasured mother's BMI and her observed covariates. To test the robustness of our

imputations we compare our our results to those using alternative methods of imputation. We impute using hotdeck imputation matching on reported BMI for children and adults. The advantage of hotdeck imputation is that it does not impose a functional form on the validation equation. This can allow for improved fitting of error corrected BMI when the relationship between true BMI and it's surrogate (typically observed BMI) is highly irregular. Estimates of the impact of BMI on annual youth medical expenditures are presented in Appendix Table B2. Using these alternate imputations, we find slightly larger effect sizes than using the regression based imputation. Interestingly, when using the ARCIV estimator, the effects significantly attenuate, and then increase when child BMI is imputed to perform RCIV. This suggests that this method of imputation, though perhaps correct in the aggregate, introduces idiosyncratic error into the imputations that may result in additional bias. One metric to evaluate alternative imputations is to see if sample covariances can be used to explain the difference between the RCIV and ARCIV estimates, suggesting that ARCIV is consistent. We find that the analytical relationship does not hold across the regression calibrated estimators. Additionally, our calibration regressions from the validation data have $R^2 = 0.999$, suggesting that the regression model is a strong fit for the relationship between true BMI and reported BMI. It is unlikely that hotdeck imputation, though more flexible, can improve the model fit by regression calibration.

Despite these limitations, we make an important contribution by being the first to use instrumental variables to estimate the the causal impact of youth obesity on medical expenditures. The discrepancy between our IV estimates and the non-IV estimates from previous studies suggest the costs of youth obesity are larger than previously believed. This has important implications for the cost effectiveness of weight-management interventions targeted at children. Previous estimates of the cost of childhood obesity have been used to motivate local and national level childhood obesity interventions (Trasande, 2010; Brown et al., 2007; Wang et al., 2003) For example, Trasande (2010) estimated the cost-effectiveness of government spending to reduce childhood obesity using estimates from Finkelstein et al. (2009). Our IV estimates suggest that Trasande (2010) signifi-

cantly underestimated the economic impact of these interventions. Whitlock et al. (2010) examined the effectiveness of both behavioral and pharmacologic weight-management interventions, and found that comprehensive behavioral interventions of medium-to-high intensity resulted in 1.9 to 3.3 BMI unit reduction.[1] Using our estimates and the most conservative policy effect size, a 1.9 BMI unit reduction would translate into a $166.04 cost savings per student, per year. Our estimates suggest that these and other treatments may be cost-effective, or even lead to cost savings, where using previous estimates may under-estimate the benefits relative to the costs of these interventions.

[1]They also found that behavioral interventions combined with prescription medicine can cause small to moderate reductions in BMI.

Table 2.1: Clinical Weight Classifications for Youth

Clinical Weight Classification	BMI Percentile Range for Youth
Underweight	BMI $\leq 5^{\text{th}}$
Healthy Weight	$5^{th} \leq$ BMI $< 85^{th}$
Overweight	$85^{th} \leq$ BMI$< 95^{th}$
Obesity	BMI $\leq 95^{th}$
Severe Obesity	BMI $\geq 99^{th}$

Source: Centers for Disease Control and Prevention (CDC) (2014); Skelton et al. (2009)
Notes: The percentiles correspond to a reference (historic), not the current, distribution of weight-for-height that is specific to gender and age. Youths defined as aged 2 to 19 years.

Table 2.2: Prevalence of Youth Obesity in the United States Defined Using BMI

Study	Years	Ages 2-5	Ages 6-11	Ages 12-19
NHES II	1963-1965		4.2	
NHES III	1966-1970			4.6*
NHANES I	1971-1974	4.0	6.1	
NHANES II	1976-1980		6.5	5.0
NHANES III	1988-1994	7.2	11.3	10.5
NHANES Continuous	1999-2002	10.3	15.9	16.0
	2003-2006	12.5	17.0	17.6
	2007-2010	11.2	18.8	18.2

Source: National Center for Health Statistics (NCHS) (2014)
Notes: Based on measured weight and height from the nationally representative samples in the National Health and Nutrition Examination Surveys. Obesity defined as a weight-for-height exceeding the 95th percentile in a historic reference population; see Table 2.1. Youth defined as individuals aged 2 to 19 years. NHANES stands for National Health and Nutrition Examination Survey. NHES stands for National Health Examination Survey. NHES I sampled adults aged 18-79, NHES II included children aged 6-11, NHES III included youths aged 12-17 years. *NHES III included youths aged 12-17, not 12-19.

Table 2.3: Descriptive Statistics for Boys in Two-Parent Households

Variables	**Mean**	**S.D.**	**Min**	**Max**
Has positive medical expenditures	0.86	(.35)	0	1
Annual Medical Expenditures*	$1,959.22	(5,374.2)	1	226,914.90
BMI (MEPS)	22.49	(4.91)	14.50	67.10
Hispanic	0.16	(.37)	0	1
Black	0.14	(.35)	0	1
Other race	0.04	(.19)	0	1
Age in months	170.91	(24.)	127	221
Self-reported	0.01	(.08)	0	1
Mother				
HS diploma	0.32	(.47)	0	1
Some college	0.26	(.44)	0	1
Bachelor's degree	0.17	(.38)	0	1
BA plus	0.09	(.29)	0	1
Father				
HS diploma	0.21	(.41)	0	1
Some college	0.14	(.35)	0	1
Bachelor's degree	0.11	(.31)	0	1
BA plus	0.08	(.27)	0	1
Household				
People in the household aged 0 - 5	0.19	(.50)	0	6
People in the household aged 6 - 17	2.12	(1.04)	1	9
Northeast	0.18	(.38)	0	1
Midwest	0.23	(.42)	0	1
South	0.37	(.48)	0	1
West	0.22	(.42)	0	1
Urban	0.82	(.39)	0	1
Year 2000	0.08	(.28)	0	1
Year 2001	0.09	(.29)	0	1
Year 2002	0.09	(.29)	0	1
Year 2003	0.09	(.29)	0	1
Year 2004	0.10	(.30)	0	1
Year 2005	0.10	(.30)	0	1
Year 2006	0.09	(.29)	0	1
Year 2007	0.09	(.28)	0	1
Year 2008	0.09	(.28)	0	1
Year 2009	0.09	(.29)	0	1
Year 2010	0.09	(.28)	0	1

Notes: Data: MEPS 2000–2010. N = 13,718 (11,178 have positive expenditures). All entries are in 2010 dollars.

Table 2.4: Descriptive Statistics for Girls in Two-Parent Households

Variables	**Mean**	**S.D.**	**Min**	**Max**
Has positive medical expenditures	0.88	(.33)	0	1
Annual Medical Expenditures*	$2,103.89	(4,806.27)	1.27	97,217.22
BMI (MEPS)	22.08	(4.82)	14.40	73.50
Hispanic	0.16	(.37)	0	1
Black	0.15	(.36)	0	1
Other race	0.04	(.19)	0	1
Age in months	171.66	(24.26)	127	221
Self-reported	0.01	(.08)	0	1
Mother				
HS diploma	0.31	(.46)	0	1
Some college	0.28	(.45)	0	1
Bachelor's degree	0.16	(.37)	0	1
BA plus	0.09	(.29)	0	1
Father				
HS diploma	0.21	(.40)	0	1
Some college	0.13	(.34)	0	1
Bachelor's degree	0.11	(.32)	0	1
BA plus	0.08	(.27)	0	1
Household				
People in the household aged 0 - 5	0.19	(.49)	0	4
People in the household aged 6 - 17	2.11	(1.03)	1	9
Northeast	0.17	(.38)	0	1
Midwest	0.24	(.43)	0	1
South	0.35	(.48)	0	1
West	0.23	(.42)	0	1
Urban	0.81	(.39)	0	1
Year 2000	0.08	(.27)	0	1
Year 2001	0.09	(.28)	0	1
Year 2002	0.09	(.29)	0	1
Year 2003	0.09	(.28)	0	1
Year 2004	0.09	(.29)	0	1
Year 2005	0.09	(.29)	0	1
Year 2006	0.10	(.30)	0	1
Year 2007	0.10	(.30)	0	1
Year 2008	0.10	(.29)	0	1
Year 2009	0.09	(.29)	0	1
Year 2010	0.09	(.28)	0	1

Notes: Data: MEPS 2000–2010. N = 13,284 (11,058 have positive expenditures). All entries are in 2010 dollars.

Table 2.5: BMI of Boys and Their Mothers

	Observations	**Mean**	**S.D.**	**Min**	**Max**
Reported Child BMI	13,718	22.49	4.91	14.50	67.10
Normal Weight	8,014	19.54	1.97	14.50	24.90
Overweight	2,341	23.82	1.42	20.00	27.32
Obese	3,363	29.71	4.50	22.60	67.10
Imputed Child BMI	13,718	22.80	5.34	13.74	49.00
Normal Weight	8,014	19.55	2.24	13.74	25.23
Overweight	2,341	24.34	1.91	20.11	28.83
Obese	3,363	30.73	4.39	22.91	49.00
Mothers					
Reported BMI	13,718	27.95	6.71	14.70	73.20
Imputed BMI	13,718	29.14	7.17	16.00	67.28

Notes: Data: MEPS 2000–2010. Overweight and obese labels are determined using imputed BMI from the NHANES. Measured BMI data are imputed from the NHANES by matching on BMI percentile rank, age, race, and gender.

Table 2.6: BMI of Girls and Their Mothers

	Observations	**Mean**	**S.D.**	**Min**	**Max**
Reported Child BMI	13,284	22.08	4.82	14.40	73.50
Normal Weight	7,914	19.38	1.97	14.40	24.20
Overweight	2,460	23.71	1.62	20.00	28.10
Obese	2,910	29.85	4.71	22.70	73.50
Imputed Child BMI	13,284	23.17	5.25	14.75	52.00
Normal Weight	7,914	20.10	2.13	14.75	25.40
Overweight	2,460	25.18	1.92	20.66	29.88
Obese	2,910	31.87	4.28	24.01	52.00
Mothers					
Reported BMI	13,284	28.22	6.69	14.40	71.60
Imputed BMI	13,284	29.45	7.19	16.00	67.25

Notes: Data: MEPS 2000–2010. Overweight and obese labels are determined using imputed BMI from the NHANES. Measured BMI data are imputed from the NHANES by matching on BMI percentile rank, age, race, and gender.

Figure 2.1: Distribution of Mothers' BMI in the NHANES

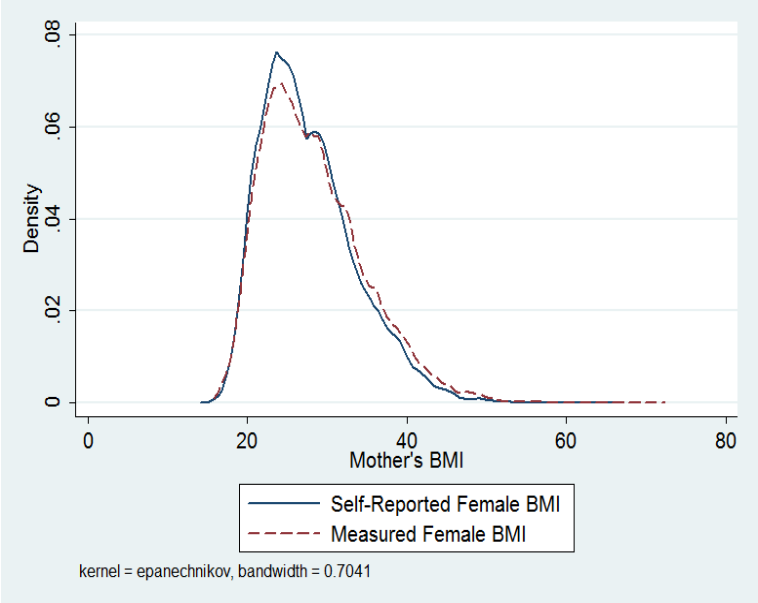

Figure 2.2: Distribution of Mothers' BMI in the MEPS

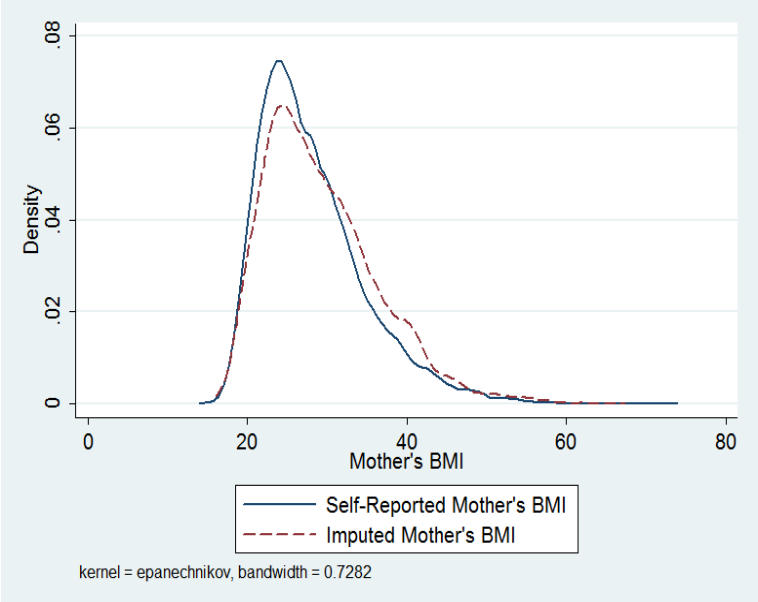

Figure 2.3: Distribution of Child BMI in MEPS and NHANES

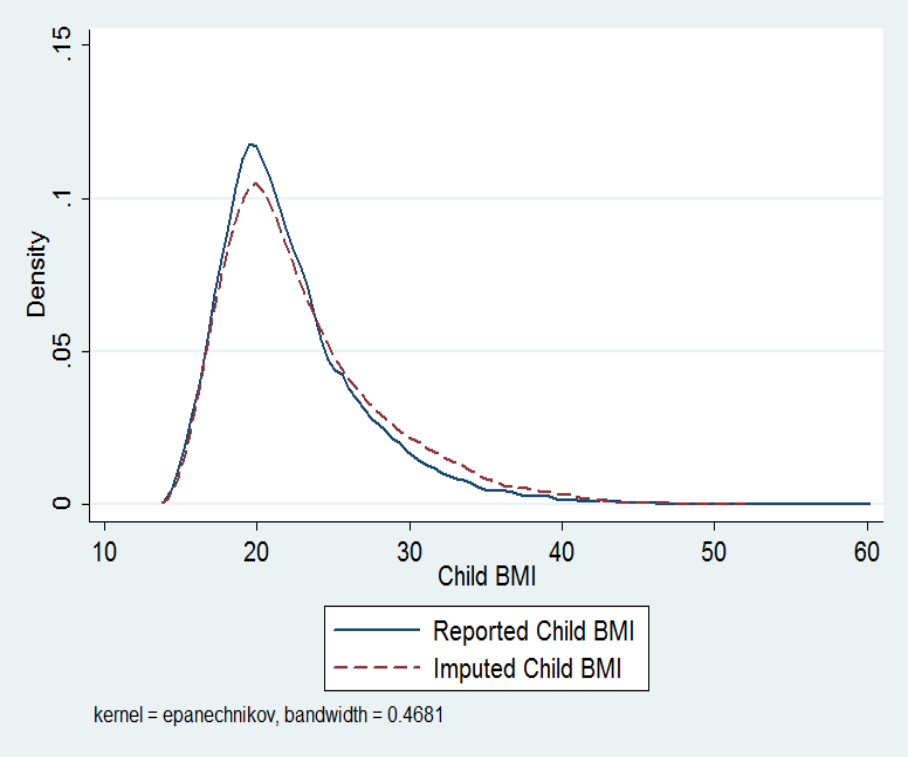

Figure 2.4: Proxy-Reporting Error in Boys' BMI

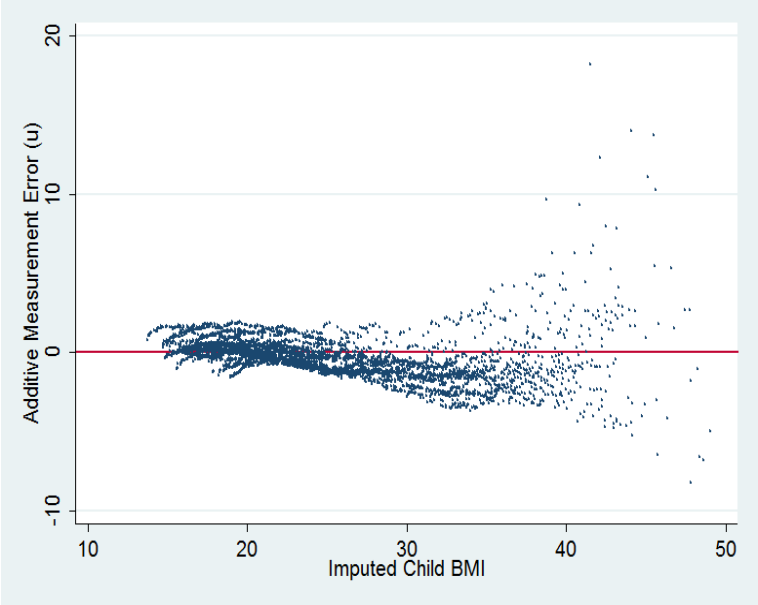

Figure 2.5: Proxy-Reporting Error in Boys' BMI by Age

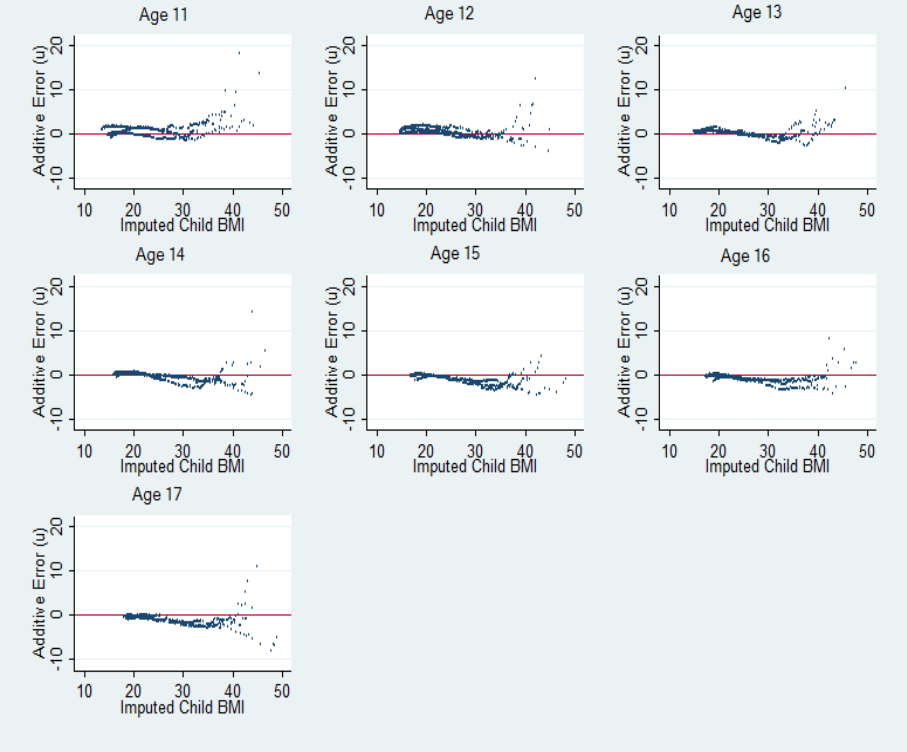

Figure 2.6: Proxy-Reporting Error in Girls' BMI

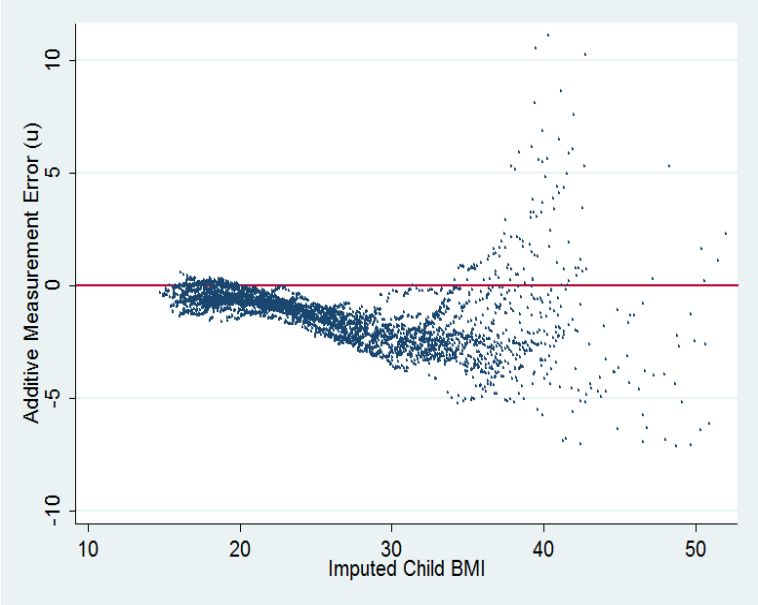

Figure 2.7: Proxy-Reporting Error in Girls' BMI by Age

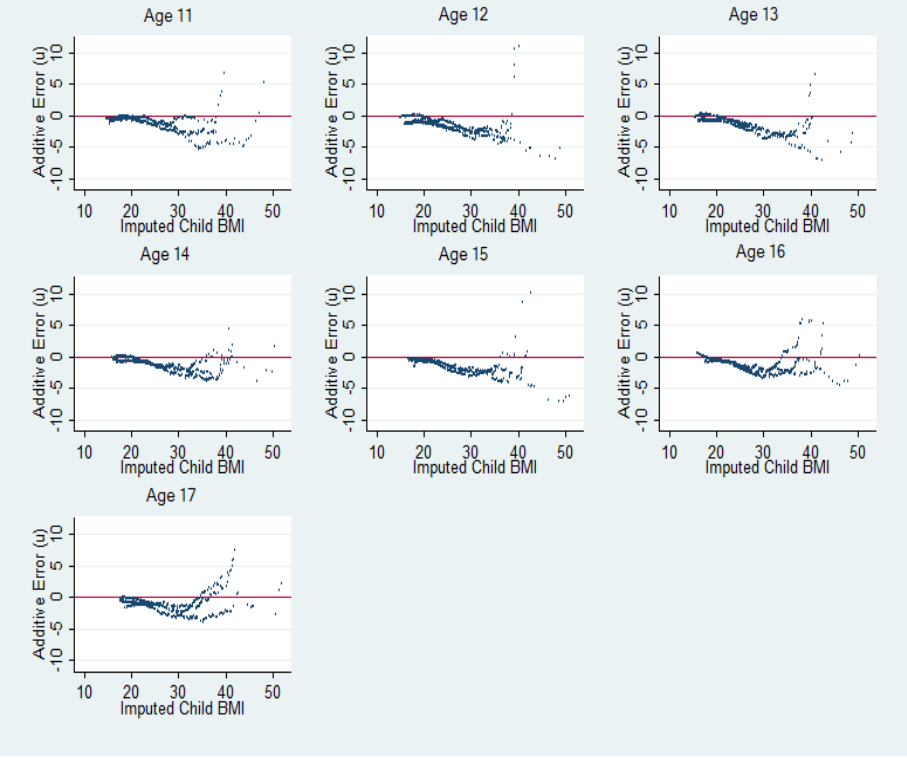

Table 2.7: Marginal Effects of Child BMI on Annual Medical Expenditures

Estimator:	Non-IV	IV	RCIV	ARCIV	ARCIV$\times\hat{\delta}$
Column	(1)	(2)	(3)	(4)	(5)
Boys & Girls	17.45	92.66***	81.52***	87.39***	81.20
N= 27,002	(10.89)	(32.35)	(29.74)	(32.08)	
		[1,657.48]	[1,829.79]	[2,064.07]	{.929}
Boys	19.02	101.02**	91.70**	98.60**	93.18
N=13,718	(15.77)	(51.18)	(46.76)	(50.54)	
		[905.86]	[991.80]	[1,002.41]	{.945}
Girls	16.80	100.04***	85.89**	92.18**	84.98
N= 13,284	(14.31)	(36.31)	(33.48)	(36.11)	
		[697.44]	[750.08]	[975.86]	{.922}

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. First-stage F-statistics in brackets. All entries are in 2010 dollars. For column 5, $\hat{\delta}$ displayed in curly brackets.

Table 2.8: Marginal Effects of Overweight and Obese on Annual Medical Expenditures

Estimator:	Non-IV	IV	RCIV	ARCIV
Column:	(1)	(2)	(3)	(4)
Marginal Effect of Overweight Relative to Normal Weight				
Boys & Girls	83.67	433.46***	386.36***	409.74***
N= 27,002	(52.38)	(143.29)	(131.43)	(143.23)
Boys	84.66	429.39**	416.31**	419.43**
N=13,718	(70.15)	(196.36)	(189.75)	(194.87)
Girls	88.68	524.11***	423.79***	483.78**
N= 13,284	(76.13)	(184.71)	(155.39)	(185.13)
Marginal Effect of Obese Relative to Normal Weight				
Boys & Girls	192.76	1,164.47**	1,032.54**	1,091.23**
N= 27,002	(124.23)	(460.38)	(416.2)	(451.9)
Boys	202.23	1,193.82*	1,129.77*	1,163.93*
N=13,718	(172.75)	(682.17)	(640.59)	(671.37)
Girls	197.45	1,399.35**	1,152.51**	1,273.28**
N= 13,284	(175.13)	(598.37)	(506.01)	(585.02)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level. BRR standard errors in parentheses. Columns 1 - 3 use differences in mean proxy-reported BMI. Column 4 uses differences in mean imputed BMI.

Table 2.9: Marginal Effects of Child BMI on Annual Medical Expenditures by Payment Type

Estimator: Population	ARCIV Total Expenditures	ARCIV Third Party Expenditures	ARCIV Out-of-Pocket Expenditures	ARCIV Third Party Expenditures	Out-of-Pocket Expenditures
Boys & Girls	87.39***	102.56***	-24.95***	86.38***	-15.62*
N= 27,002	(32.08)	(31.27)	(8.69)	(29.87)	(8.79)
Boys	98.60**	123.96**	-31.38***	104.71**	-23.35**
N=13,718	(50.54)	(52.99)	(10.58)	(47.1)	(10.31)
Girls	92.18**	102.03***	-17.61*	87.71***	-6.53
N= 13,284	(36.11)	(32.06)	(11.79)	(31.79)	(12.06)
Additional Controls for Insurance Status				X	X
Observations with positive medical expenditures	27,002	20,575	17,177	20,575	17,177

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars.

Table 2.10: Marginal Effects of Child Overweight and Obese on Annual Medical Expenditures by Payment Type

Population	Total Expenditures	Third Party Expenditures	Out-of-Pocket Expenditures	Third Party Expenditures	Out-of-Pocket Expenditures
Marginal Effect of Overweight Relative to Normal Weight					
Boys & Girls	409.74***	465.15***	-115.48***	395.49***	-72.99*
N= 27,002	(143.23)	(124.38)	(37.87)	(122.97)	(39.7)
Boys	419.43**	493.35***	-136.12***	473.44***	-101.88**
N=13,718	(194.87)	(165.47)	(43.47)	(180.47)	(43.42)
Girls	483.78**	408.97***	-90.06	451.94***	-33.91
N= 13,284	(185.13)	(109.8)	(56.87)	(152.79)	(61.15)
Marginal Effect of Obese Relative to Normal Weight					
Boys & Girls	1,091.23**	1,334.09***	-228.18***	1,099.59***	-151.69*
N= 27,002	(451.9)	(463.56)	(63.39)	(431.48)	(74.13)
Boys	1163.93*	1,533.36**	-265.75***	1,429.39**	-209.25***
N=13,718	(671.37)	(711.21)	(70.39)	(732.86)	(76.02)
Girls	1,273.28**	1,238.38***	-184.75*	1,241.43**	-73.66
N= 13,284	(585.02)	(430.26)	(101.27)	(533.87)	(125.38)
Additional Controls for Insurance Status				X	X
Observations with positive medical expenditures	12,765	11,660	10,735	11,660	10,735

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars.

Table 2.11: Marginal Effects of Child BMI on Third-Party Medical Expenditures by Payer Type

Population	Third Party Expenditures	Medicaid Expenditures	Private Insurance Expenditures
Marginal Effect of BMI			
Boys & Girls (Full Sample)	102.56***	145.51*	-12.58
N = 27,002	(31.27)	(84.37)	(19.69)
Boys & Girls		255.06	13.13
N = 10,389 / 14,190	–	(157.45)	(29.76)
Boys	–	135.64*	25.92
N = 5,215 / 7,288		(71.16)	(38.97)
Girls	–	266.48**	24.43
N = 5,174 / 6,902		(130.28)	(32.16)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars. Row 1 incorporates the full sample. Subsequent rows restrict the sample to individuals covered by Medicaid / Private insurance respectively.

Table 2.12: Marginal Effects of Overweight and Obese on Third-Party Medical Expenditures by Payer Type

Population	Third Party Expenditures	Medicaid Expenditures	Private Insurance Expenditures
Marginal Effect of Overweight Relative to Normal Weight			
Boys & Girls (Full Sample)	465.15***	316.67***	-62.11
N = 27,002	(124.38)	(49.89)	(95.77)
Boys & Girls	–	662.75***	63.43
N = 10,389 / 14,190		(134.73)	(143.88)
Boys	–	484.21***	127.40
N = 5,215 / 7,288		(150.16)	(188.5)
Girls	–	616.07***	120.12
N = 5,174 / 6,902		(106.08)	(156.67)
Marginal Effect of Obese Relative to Normal Weight			
Boys & Girls (Full Sample)	1,334.09***	1,704.55***	-153.06
N = 27,002	(463.56)	(536.96)	(202.27)
Boys & Girls	–	2,830.95***	141.36
N = 10,389 / 14,190		(981.99)	(339.46)
Boys	–	1,698.30***	293.86
N = 5,215 / 7,288		(780.12)	(472.76)
Girls	–	2,938.28***	283.15
N = 5,174 / 6,902		(1,006.97)	(396.12)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars. Row 1 incorporates the full sample. Subsequent rows restrict the sample to individuals covered by Medicaid / Private insurance respectively.

Table 2.13: Aggregate Costs of Youth Overweight and Obesity on Annual Medical Expenditures

Year	Total Expenditures	Third Party Expenditures	Out-of-Pocket Expenditures	Medicaid Expenditures
2000	3.07 (1.11, 5.04)	3.32 (1.61, 5.03)	-0.96 (-1.48, -0.44)	0.96 (0.39, 1.52)
2001	3.90 (1.42, 6.39)	4.60 (2.13, 7.08)	-1.02 (-1.54, -0.49)	1.42 (0.62, 2.23)
2002	4.26 (1.44, 7.08)	4.85 (2.15, 7.56)	-1.12 (-1.71, -0.53)	3.32 (1.15, 5.50)
2003	4.08 (1.52, 6.65)	4.63 (2.28, 6.98)	-1.10 (-1.69, -0.52)	2.77 (1.53, 4.02)
2004	4.64 (1.86, 7.43)	5.51 (2.87, 8.16)	-1.16 (-1.79, -0.52)	3.19 (1.67, 4.70)
2005	5.32 (1.86, 8.78)	6.31 (2.95, 9.67)	-1.29 (-1.96, -0.62)	4.16 (2.12, 6.21)
2006	5.53 (2.32, 8.73)	6.61 (3.57, 9.65)	-1.34 (-2.07, -0.61)	3.66 (1.99, 5.32)
2007	4.45 (1.55, 7.35)	5.25 (2.44, 8.06)	-1.10 (-1.68, -0.51)	2.95 (1.55, 4.35)
2008	5.45 (2.08, 8.83)	6.55 (3.30, 9.80)	-1.31 (-1.99, -0.63)	3.95 (2.07, 5.83)
2009	5.45 (2.04, 8.87)	6.81 (3.40, 10.22)	-1,15 (-1.72, -0.58)	5.29 (2.65, 7.88)
2010	6.12 (2.45, 9.79)	7.61 (3.93, 11.28)	-1.23 (-1.90, -0.56)	6.46 (3.23, 9.69)
Generalized to population 2010	9.70 (3.88, 15.53)	12.06 (6.24, 17.88)	-1.95 (-3.01, -0.89)	10.23 (5.12, 15.34)
2000 - 2010 average	4.75 (1.87, 7.63)	5.64 (2.95, 8.33)	-1.16 (-1.75, -0.57)	3.46 (1.97, 4.96)
Generalized to population 2000 - 2010 average	7.24 (2.85, 11.63)	8.60 (4.50, 12.69)	-1.77 (-2.66, -0.88)	5.40 (3.07, 7.73)

Notes: Data: MEPS 2000–2010. All expenditures are in billions of 2010 dollars. 90% BRR Confidence intervals in parentheses are adjusted for the complex design of the MEPS. Generalized effects for total, third-party, and out-of-pocket expenditures are computed by multiplying the average effect into the ratio of the general population of children aged 11-17 and children aged 11-17 in two parent households. The general effects for Medicaid expenditures are generalized to the population of children aged 11-17 enrolled in Medicaid. The population counts are derived from the MEPS sample and sample weights.

Table 2.14: Aggregate Costs of Youth Obesity on Annual Medical Expenditures

Year	Total Expenditures	Third Party Expenditures	Out-of-Pocket Expenditures	Medicaid Expenditures
2000	2.21 (0.67, 3.75)	2.48 (1.03, 3.93)	-0.60 (-0.92, -0.29)	0.88 (0.25, 1.51)
2001	3.02 (1.01, 5.02)	3.71 (1.57, 5.84)	-0.68 (-1.02, -0.34)	1.38 (0.47, 2.29)
2002	3.48 (1.04, 5.92)	4.09 (1.59, 6.59)	-0.81 (-1.22, -0.40)	3.07 (0.77, 5.37)
2003	3.05 (1.00, 5.10)	3.60 (1.57, 5.63)	-0.71 (-1.07, -0.35)	2.60 (1.20, 4.00)
2004	3.49 (1.26, 5.71)	4.34 (2.04, 6.63)	-0.74 (-1.12, -0.35)	3.35 (1.37, 5.33)
2005	3.94 (1.21, 6.68)	4.90 (2.04, 7.76)	-0.82 (-1.21, -0.42)	4.30 (1.74, 6.86)
2006	3.92 (1.42, 6.41)	4.96 (2.35, 7.56)	-0.78 (-1.16, -0.39)	3.73 (1.66, 5.79)
2007	3.42 (1.06, 5.78)	4.26 (1.76, 6.71)	-0.71 (-1.06, -0.36)	3.02 (1.31, 4.72)
2008	4.01 (1.35, 6.67)	5.05 (2.25, 7.85)	-0.81 (-1.21, -0.41)	3.75 (1.62, 5.87)
2009	4.20 (1.41, 6.99)	5.47 (2.46, 8.49)	-0.76 (-1.13, -0.39)	4.90 (2.09, 7.72)
2010	4.69 (1.68, 7.71)	6.05 (2.81, 9.30)	-0.82 (-1.24, -0.39)	5.81 (2.25, 9.37)
Generalized to population 2010	7.44 (2.66, 12.23)	9.60 (4.45, 14.75)	-1.29 (-1.96, -0.62)	9.20 (3.57, 14.83)
2000 - 2010 average	3.59 (1.26, 5.91)	4.44 (2.09, 6.80)	-0.75 (-1.10, -0.40)	3.34 (1.57, 5.12)
Generalized to population 2000 - 2010 average	5.46 (1.93, 9.00)	6.77 (3.18, 10.36)	-1.14 (-1.68, -0.60)	5.21 (2.45, 7.89)

Notes: Data: MEPS 2000–2010. All expenditures are in billions of 2010 dollars. 90% BRR Confidence intervals in parentheses are adjusted for the complex design of the MEPS. *,**,*** indicate significance at 10%, 5%, 1% level respectively. Generalized effects for total, third-party, and out-of-pocket expenditures are computed by multiplying the average effect into the ratio of the general population of children aged 11-17 and children aged 11-17 in two parent households. The general effects for Medicaid expenditures are generalized to the population of children aged 11-17 enrolled in Medicaid. The population counts are derived from the MEPS sample and sample weights.

Chapter 3

The Impact of Obesity on Adult Medical Care Costs

3.1 Introduction

The prevalence of obesity has more than doubled in the US over the last 30 years (Burkhauser et al., 2009). In 2009−2010, the prevalence of obesity (BMI $\geq$ 30) was 35.5% of adult men and 35.8% of adult women.[1] Additionally, over two-thirds of US adults are considered overweight, having 25 $\leq$ BMI $<$ 30 (Flegal et al., 2012). Obesity is associated with an increased risk of many health conditions, including cardiovascular disease, stroke, type-2 diabetes, certain cancers, osteoarthritis, asthma, and depression (Dixon, 2010; Hu, 2008). Incurring treatment for these conditions can represent a considerable increase in medical expenditures, particularly if they result in prolonged inpatient stays or the development of chronic illness.

A number of previous studies have estimated the impact of obesity on medical care costs by estimating cross-sectional models of the relationship between BMI or obesity on medical expenditures. (Finkelstein et al., 2009; Trasande et al., 2009; Trasande and Chatterjee, 2009; Thorpe et al., 2004; Finkelstein et al., 2003). For example, Finkelstein et al. (2009) use the 1998 and 2006 Medical Expenditure Panel Survey (MEPS) to estimate the impact of obesity on medical expenditures. They find that on average, being

[1]Body mass index is defined as weight in kilograms divided by height in meters squared.

obese was associated with $1,429 (2006 dollars) in additional spending on average. These studies have identified significant medical costs associated with obesity, and have raised the concern that much of these additional expenditures are paid for by state and Federal insurance programs such as Medicare and Medicaid.

A limitation of these previous studies is that estimating cross-sectional models can only estimate the correlation between BMI and medical expenditures, and cannot be interpreted as causal estimates. Another limitation is that BMI and weight status are endogenous in these models, which can lead to biased estimates of the association between BMI and medical expenditures. For example, a prolonged inpatient stay may result in increased expenditures as well as weight loss, resulting in underestimation of the true effect. The correlation will also likely underestimate the true effect as individuals with less access to care tend to have higher rates of obesity (Fontaine and Bartlett, 2000).

More recent research has addressed the endogeneity of weight while estimating the causal effect of obesity on medical expenditures. Cawley and Meyerhoefer (2012) use the BMI of the oldest biological child as an instrumental variable for the BMI of the respondent. Unlike Least-Squares regression, instrumental variables estimation is consistent in the presence of endogeneity as well as classical measurement error in BMI. They find that instrumenting significantly increases the magnitude of the impact of obesity on medial expenditures relative to non-IV estimates. They estimate that obesity raises annual medical costs by $2,741 (2005 dollars) per obese individual, compared to $656 when not using instrumental variables. Cawley et al. (2014) update this model using the 2000 - 2010 MEPS, and find that obesity raises annual medical care costs by $3,508 (2010 dollars) per obese individual on average.

As costs of obesity increase, given the structure of health insurance coverage, a large share of these costs may represent a negative externality if individuals do not face the financial consequences of their poor health behaviors (Bhattacharya and Sood, 2005). For example, Bhattacharya and Sood (2005) suggest that if premiums are not risk-adjusted to reflect the weight of the enrollee then health insurance may, in effect, subsidize obesity. If public programs bear the larger share of the external costs of obesity, then these costs

may represent a loss to societal welfare as families most prone to obesity face the smallest out-of-pocket share of medical costs for conditions associated with obesity. Earlier studies have suggested that there is a substantial externality of obesity. In their estimates using the MEPS, Finkelstein et al. (2009) found that obesity related costs to Medicare and Medicaid were 32% of total obesity related costs.

Cawley and Meyerhoefer (2012) point out that heights and weights in the MEPS are not measured, but instead reported by a single survey respondent, who self-reports their own height and weight, and proxy-reports the other members of the household. Studies of adult reporting behavior suggest that individuals tend to overstate their height and under-report their weight, resulting in smaller BMIs, and that the degree to which BMI is understated increases with BMI, generating a correlation between measurement error and the true value (Stommel and Schoenborn, 2009; Gorber et al., 2007). We expect this reporting error will lead to misclassification of individuals, understating the true rates of overweight and obesity. Previous IV estimates of the impact of obesity on medical expenditures may still result in biased coefficient estimates if measurement error is non-classical, as under-reporting will generate additive error that is not mean zero and potentially correlated with unobserved factors that also affect medical expenditures (Bound et al., 2001; O'Neill and Sweetman, 2013).

This paper uses data from the 2000 - 2010 Medical Expenditure Panel Survey to estimate the impact of BMI and obesity on adult medical expenditures. We acknowledge the limitations of previous studies as well as the empirical evidence of under-reporting, and use validation data to predict true BMI their children, whose BMI serve as an instrument. We use the National Health and Nutrition Examination Survey (NHANES) as a source of measured BMI data. The NHANES contains both measured and self-reported heights and weights for adults as well as measured child heights and weights, and can be used to correct reported BMI in the MEPS (Cawley, 2004). Due to the unavailability of proxy-reports for child height and weight in the NHANES to correspond to proxy-reports in the MEPS, we follow the imputation procedure in Courtemanche et al. (2014) and use the percentile rank of BMI as the matching agent to impute child BMI.

We are able to use the imputed adult and child BMIs to indirectly observe adult self- and proxy-reporting behavior, and provide evidence of under-reporting in the MEPS. We use the ARCIV estimator from chapter 1 to estimate the impact of obesity on annual medical expenditures. For most sub-populations we find estimates very similar to previous IV estimates. We also find evidence that third-parties bear a larger share of additional costs associated with obesity, suggesting that there are considerable negative externalises associated with obesity.

3.2 Identification Strategy

3.2.1 Instrumental Variables

We estimate the effect of BMI on medical expenditures using the IV-GLM procedure described in Carroll et al. (2006). In order for IV estimation to produce consistent estimates, the instrument must be sufficiently correlated with the endogenous or mismeasured variable. Our first stage partial F-statistics range from 80.75 - 773.98, well in excess of the rule-of-thumb F-stat = 10 for all of our specifications (Stock et al., 2002). More difficult to establish is that the instrument is independent of the error in the structural model.[1]

Independence would be violated if parents and their childrens' BMI are affected by common environmental factors that also influence the child's medical care costs. Parents are typically responsible for their children's healthcare decisions. Parenting decisions may stem from parent attitudes towards health, and may be reflected in the parent's BMI. If these attitudes are correlated with unobserved cultural or environmental factors that also effect the child's medical expenditures, then the instrument will be correlated with the error term in the model. This cannot be directly tested in our data.

There is however, a substantial behavioral genetic literature validating the genetic relationship between weight of biological relatives. (Haberstick et al., 2010; Smith et al., 2009; Wardle et al., 2008; Grilo and Pogue-Geile, 1991) For example, Grilo and Pogue-

[1]Independence is a stronger assumption than for linear IV, but is required for IV-GLM (Carroll et al., 2006).

Geile (1991) find that 40–70% of the variation in obesity-related phenotypes, such as body mass index, skinfold thickness, fat mass and leptin levels, is inheritable. Adoption studies have found that the strong genetic correlation between children and their biological parents is not weaker for children raised by adoptive parents (Stunkard et al., 1986; Srensen and Stunkard, 1993). Twin studies also find that the correlation in the weight of twins does not depend on whether they were raise together or separately. (Price and Gottesman, 1991; Maes et al., 1997)

This same behavioral genetic literature finds little to no evidence in favor of shared household or environmental effects on BMI (Haberstick et al., 2010; Wardle et al., 2008; Maes et al., 1997; Grilo and Pogue-Geile, 1991). Taken at face value, no shared household effect of BMI implies that a parent's BMI can only effect their child's medical expenditures through it's effect on child BMI. We interpret this lack of evidence for environmental effects not as a statement that location, shared culture, or shared diet do not effect child BMI, but that all of the correlation between parent BMI and these factors is captured by the strong correlation between parent and child BMI due to the strong influence of shared genetics.

Prior economic research has employed genetic variation in weight to generate instrumental variables. Cawley (2000) use the National Longitudinal Survey of Youth (NLSY) in order to measure the effect of body weight on employment disability. He uses the BMI of a biological child to instrument for adult BMI. Lindeboom et al. (2010) estimated the effect of obesity on employment, using rich data from the British National Child Development Study (NCDS). The results show a significant negative association between obesity and employment even after controlling for a rich set of demographic, socioeconomic, environmental and behavioral variables. Kline and Tobias (2008) used the BMI of parents to instrument for the BMI of their adult offspring and estimate the impact of BMI on wages in Britain. Cawley (2004) used the BMI of a sibling to instrument for adult BMI in estimating the impact of obesity on wages. More recently, Cawley and Meyerhoefer (2012) and Cawley et al. (2014) used the BMI of children to instrument for the BMI of their biological parents in the MEPS in order to estimate the impact of obesity on adult

medical expenditures.

3.2.2 Reporting Error in BMI

Prior research has documented the relationship between adult's height and weight and reporting error in their own weight. In their systematic review, Gorber et al. (2007) found an overall trend of under-reporting in adult self-reports. Villanueva (2001) studied the NHANES III, and found that measurement error in weight was negatively associated with measured BMI. More recently, Stommel and Schoenborn (2009) studied self-reporting in the 2001 - 2006 NHANES and found that, while adults tended to over-estimate BMI at low levels of BMI < 22, the tended to under-estimate BMI for BMIs > 28. In the MEPS, many adults have their BMI proxy-reported by another adult. (usually men being proxy-reported by their wives.) Few studies look at the accuracy of proxy-reporting. Reed and Price (1998) find that first degree relatives provide fairly accurate weight estimates, within 3–5% of measured weight, but that these proxy reports also tend to under-report weight and over-estimate height, leading to underestimation of BMI.

IV estimation is robust to classical measurement error, but may be biased in the presence of non-classical measurement error if this error is differential.[1] Chapter 1 contains analytical and simulation results demonstrating the properties of the IV estimator and regression calibrated IV estimators under non-classical measurement error.

3.2.3 Corrected Body Mass Index

OLS and IV estimates are not robust to all forms of measurement error. Differential measurement error or correlation between the true instrument and the measurement error in the endogenous variable may still bias IV estimates. It is possible to reduce the magnitude of these biases by using validation techniques to replace mismeasured variables

[1]Under-reporting that increases with BMI will lead to a negative correlation between BMI and the measurement error in BMI. This negative correlation, combined with positive correlation between parent and child BMI does not necessarily imply negative correlation between child BMI and the measurement error in adult BMI, as covariance (and correlation, which is normalized covariance) is not transitive (Langford et al., 2001). If child BMI is correlated with the measurement error in their parent's BMI, then using child BMI as an IV for their parent's BMI may not result in consistent estimates of the impact of obesity on medical expenditures.

with error-corrected fitted values (Bound et al., 2001). Prior studies have used measured validation data to correct for measurement error in BMI when estimating the impact of BMI on health and labor outcomes (Cawley, 2000, 2004; Dutton and McLaren, 2014). Validation data are typically used to generate correction equations that are used to predict true BMI in the principle sample, which only contains mismeasured BMI. (Courtemanche et al., 2014). We follow the imputation procedure described in Courtemanche et al. (2014) using the National Health and Nutrition Examination Survey (NHANES) as a source of validation data.[1] A summary of the imputation procedure in Courtemanche et al. (2014) in the context of our model follows here.

Imputation requires that there is a surrogate, or matching agent, present in the principal and validation datasets, as well as the transportability of the surrogate across both datasets. Each observation has true BMI X and their observed BMI w_j in the principle sample $j = M$ and in the validation sample $j = N$.[2] The first condition is that for true BMI x, there exists a surrogate (typically the observed BMI) w_j^{sur} such that the distribution of the outcome Y given (x, w_j^{sur}) is the same as Y given x. Essentially the observed BMI cannot contain information about the outcome that is not already reflected in true BMI. Another interpretation is that measurement error in BMI cannot be correlated with unobserved variables that influence the outcome. The second condition is the transportability of the surrogate; that the underlying distributions of true BMI in both datasets are equal conditional on the surrogate. Transportability implies that,

$$E(x|w_M, Z_M) = E(x|w_N, Z_N),^3 \tag{3.1}$$

where covariates (Z_j) are shared in both datasets (gender, age, race, etc...).

In prior research using the NHANES as validation data for self-reported adult BMI, the self-reported BMI in the NHANES is used as the surrogate for imputation (Cawley, 2000). We can only follow this method to impute adult BMI for adults who self-report

[1] We follow the arguments of Lee and Sepanski (1995) and impute BMI directly instead of predicting height and weight and using the predictions to calculate BMI. The advantage is we are directly predicting the BMI distribution, and in turn the non-linear relationship between BMI and expenditures.

[2] In our analysis, the MEPS is the principle sample and the NHANES serves as the validation sample.

[3] This is known as weak transportability (Lee and Sepanski, 1995).

their own height and weight, which will be used to generate their imputed BMI. In practice, mothers and fathers may have different self-reporting behavior, which may vary by age or race. Further, there are no proxy-reports of child BMI in the NHANES to match to proxy-reports in the MEPS.

In order to impute both adult and child BMI, we follow the method in Courtemanche et al. (2014), and use the percentile rank of BMI as the matching surrogate. In this way, misreporting manifests not only as a stochastic process on additive error terms u and v, but as a shifting in the distribution of BMIs. Using the percentile rank as the matching surrogate only requires that the expected value of true BMI conditional on reported BMI is monotonically increasing in reported BMI (Courtemanche et al., 2014). This monotonicity implies that individuals who report higher BMI are expected to have higher a BMI than those who report report lower BMI. Courtemanche et al. (2014) test this monotonicity assumption in the NHANES, and do not find evidence to reject it.

The percentile rank is a transportable surrogate if,

$$E(x|BMI\ Rank_M, Z_M) = E(x|BMI\ Rank_N, Z_N). \tag{3.2}$$

For both adults and children, we predict measured BMI in the NHANES as a function of gender, race, and their interactions with linear splines generated from the percentile rank of BMI and higher orders of age (age in months for children). We then use the estimated equations to predict true BMI for adults and children in the principal sample (MEPS).[1]

3.2.4 Empirical Model

To accommodate the non-trivial number of observations with no observed expenditures in any survey period (16% on average), we employ a two-part model of medical expenditures to estimate the impact of BMI on medical spending (Jones, 2000). The first part of the model is specified as a Logit model that estimates the probability of having

[1]To test the robustness of our imputations we alternatively impute using hotdeck imputation matching on percentile rank for children and adults. Using these alternate imputations. We find very similar results to those using the regression based imputation.

positive expenditures. The second part of the model estimates the level of medical expenditures conditional on having positive spending. We follow previous studies that have estimated the impact of BMI on medical expenditures and specify the second part as a GLM with Gamma variance structure and log link (Cawley et al., 2014; Cawley and Meyerhoefer, 2012; Finkelstein et al., 2009). This specification is more flexible than a linear model using transformed expenditures (such as log expenditures) and thus accommodates the non-linear relationship between BMI and medical expenditures, even without using higher orders of BMI in the model. We follow Manning and Mullahy (2001), and perform modified Park tests to test our choice of conditional variance and HosmerLemeshow tests to confirm that our choice of link function is consistent with the data generating process.[1]

We implement both parts of the model as Generalized Linear Models (the first stage is a GLM using the binomial distribution and a Logit link.). In this way we can incorporate instrumental variables into both parts using the IV-GLM estimator proposed by described in Hardin and Carroll (2003a). Their method is essentially two-stage non-linear least squares. The first stage is a linear regression of the mismeasured covariates on the set of included and excluded instruments. The second stage is a GLM fit of the outcome on the known covariates and the fitted values of the mismeasured covariate from the first stage[2] (Hardin and Carroll, 2003b). We use the BMI of the oldest child and the squared BMI of the oldest child as instruments. Cawley and Meyerhoefer (2012) used the BMI of the oldest child due to higher rates of non-response among younger children, and did not find significant differences in estimates across different sets of instruments.

In all models we control for gender, race/ethnicity (white, black, Hispanic, other race), respondent age (indicator variables for whether age in years is 20-34, 35-44, 45-54, or 55-64), education level (no high school diploma, high school graduate, some college,

[1]We conduct modified Park tests to confirm our choice of conditional variance, in particular that the variance is proportional to the square of the conditional mean. We find $\lambda = 1.77 - 2.02$ across our samples and is precisely estimated. We also perform a modified Hosmer–Lemeshow test by regressing prediction error from each model on deciles of the distribution of predicted expenditures. We fail to reject the null-hypothesis that the decile coefficients are jointly equal to zero for all subpopulations, indicating the choice of link function is appropriate.

[2]We estimate the IV-GLM in STATA using the **qvf** command in (Hardin et al., 2003). The command fits the GLM using iteratively re-weighted least squares (IRLS). Hardin and Hilbe (2012) (2012) detail the steps of the IRLS algorithm.

bachelors degree or higher), census region (northeast, midwest, south, or west), whether the respondent lives in an MSA, household composition (number of household members age 0-5 years, 6-17, 18-64, and 65 or older), whether the survey information was self-reported as opposed to proxy reported, fixed effects for year, the gender of the oldest child, and the age of the oldest child in months. For subgroup analyses the set of regressors is modified to drop irrelevant control variables.

3.3 Data

The Medical Expenditure Panel Survey (MEPS) is a comprehensive, nationally representative survey of the U.S. civilian non-institutionalized population.[1] In the MEPS, families are surveyed five times during a two year period about their medical care utilization and expenditures. For each family (the responding unit in the MEPS), a single individual is the primary respondent. For most families in the MEPS, the mother is the primary respondent. We can identify the primary respondent in each family (usually a parent) and we use restricted-use biological linkage variables to match parents to their biological children. Heights and weights are not measured in the MEPS. The primary respondent typically reports the heights and weights of everyone in the reporting unit. This means that the primary respondent self-reports his or her height and weight, and heights and weights for their spouse and children are generated from proxy-reports.[2]

We use data from the 2000-2010 household component of the MEPS and inflate all expenditures in each year to 2010 dollars. We limit the sample to adults between the ages 20 and 64 with biological children between the ages of 11 and 20 years old; We exclude pregnant women and observations with missing BMI. We exclude parents whose oldest child is younger than 11 years of age due to high rates non-response in height and weight. We exclude nine observations with implausibly high BMI > 80, and one observation with annual expenditures above \$500,000. We are only interested in modeling the effect of

[1] To account for the complex survey design of the MEPS, we use the method of balanced repeated replications to estimate standard errors in all models, which implements clustering at the PSU-level, stratification, and weighting.

[2] The exception to this is when all adult members of the household are present during the interview, in which case each adult self-reports their height and weight.

obesity, and drop 412 underweight individuals, bringing our final estimation sample to 17,533 men and 25,475 women.

In the MEPS, medical expenditures and the source of payment are collected directly from households as well as from the households medical care providers for every medical event. MEPS respondents are also asked whether their medical visits or other events are related to any specific medical conditions. These responses are then professionally coded using the *International Classification of Diseases*, Ninth Revision (ICD-9), and subsequently collapsed to into 259 clinically relevant medical conditions using the Clinical Classification System (CCS) developed by the Agency for Healthcare Research and Quality (Agency for Healthcare Research and Quality (AHRQ), 2007). In addition to total medical expenditures, we estimate the impact of obesity on expenditures by all third party payers (typically, public and private insurers), and also expenditures by all payers on specific categories of care: inpatient, outpatient, prescription drugs, and other (which includes dental, vision, home health care services, and medical equipment but excludes spending on over-the-counter medications).

MEPS data are collected through a stratified multi-stage probability design, which we account for in the calculation of the standard errors for our marginal effects. In particular, we use the method of balanced repeated replications to estimate standard errors in our non-IV and IV two-part GLM models. This method accounts for clustering at the PSU-level, stratification, and weighting.

We correct for reporting error in BMI by using the National Health and Nutrition Examination Survey (NHANES) as a validation dataset. The NHANES is a nationally representative survey of adults and children that combines interviews and physical examinations (CDC / National Center for Health Statistics, 2014). In the NHANES, adults and children aged 16 or older self-report their weight, and all survey respondents have their heights and weights measured[1]. The NHANES does not contain medical expenditures, but does share with the MEPS rich covariates.

The continuous version of the survey data are released in two-year waves (survey-

[1]Survey participants are not aware that they will be weighed until after they self-report their height and weight.

cycles), beginning in 1999. We construct the validation dataset by appending six survey cycles of the data, from the 1999-2000 survey cycle to the 2009-2010 survey cycle. Like the MEPS, the NHANES provides weights to account for the complex survey design (including oversampling), survey non-response, and post-stratification (CDC / National Center for Health Statistics, 2013). We construct the appropriate survey weights for a 12 year span of the data (1999-2010).[1]

3.4 Results

3.4.1 Descriptive Statistics

Descriptive statistics of variables used in our analysis are presented in Table 3.1 for men and Table 3.2 for women. (The samples are limited to parents of biological children.) In our sample, 79% of men and 88% of women have positive medical expenditures during the periods they are surveyed in the data. The high rate of adults with no expenditures is our justification for using a two-part model. Conditional on having medical expenditures, men have $3,094.94, and women have $3,592.08 in average annual medical expenditures.

The average BMI, calculated from proxy-reported or self-reported height and weight, is 28.46 for men, and 27.86 for women. After imputing BMI using correction equations from the NHANES, we estimate that the true average BMI is 29.49 for men and 29.53 for women. Comparing the means reveals that both mens' and womens' BMIs are under-reported, with womens' BMI being under-reported to a greater extent. After correcting for mis-reporting we find that the true prevalence of obesity in our sample is 39% for men and 41% for women. These are larger than the prevalence rates in the NHANES of 35.5% for men and 35.8% for women. (Flegal et al., 2012). It is possible that this difference is due to our sample of adults with biological children having a different underlying rate of obesity than the entire population, although we would expect parents to have better underlying health compared to adults without children. It may also be that our

[1]We extrapolate the method to construct weights when combining survey cycles for 12 years of data using the 10 year method in task 2 in CDC / National Center for Health Statistics (2013). Our code is displayed in Appendix A.

imputation method over-states the degree to which individuals under-report. We find similar mis-reporting of the BMI of the oldest child, which is the result of proxy-reporting by the survey responding parent.

3.4.2 Self- and Proxy-Reporting Error

Figure 3.1 compares the distribution of reported BMI in the NHANES to the measured BMI of adult females. It is clear in the NHANES, where women self-report and are subsequently measured, that the distribution of reported BMI has less mass in the right tail, and greater mass in the overweight and healthy weight BMI range. In Figure 3.2, which displays the distribution of measured and self-reported BMI for men in the NHANES, There is a spike in the mass of the distribution of self-reports just to the left of the cutoff between overweight and obese. This distribution shift due to mis-reporting is less pronounced than for women.

A necessary assumption for imputation is that reporting behavior is the same in both the validation data and the principal sample. We compare the distribution of imputed BMI in the MEPS to the reported BMI in order to indirectly observe reporting error in the MEPS as well as compare the distribution shifting in the MEPS after imputation to the under-reporting in the NHANES. Figure 3.3 shows the distribution of women's reported BMI in the MEPS and the distribution of imputed BMI. The pattern of misreporting is similar to that observed in Figure 3.1, although there appears to be more mass shifted from the right tail to the non-obese portion of the distribution. Figure 3.4 shows the distribution of men's reported BMI in the MEPS and the distribution of imputed BMI. In the MEPS, very few men self-report, so we interpret Figure 3.4 as under-reporting by their proxy, (who is usually their spouse) instead of men under-reporting their own weight as in the NHANES. However, we still find indirect evidence that spouses do similarly under-report their husbands BMI, even if we cannot directly corroborate proxy-reporting with observed behavior in the NHANES. In Figure 3.5 we observe under-reporting in the parental proxy-reports of their oldest child's BMI, showing that the instruments are themselves likely under-reported.

We use the imputed BMIs to generate sample measurement errors. We then compare the sample errors to the imputed BMIs in order to indirectly observe parental self and proxy reporting behavior. Figure 3.6 plots womens' imputed BMI against the additive error term in BMI generated by self (and some proxy) reporting. The smoothness of this relationship is due to the regression based imputation. Thus we can best interpret these as estimates of the mean additive error for a given BMI level. The distinct curves are each associated with each race category used in the prediction. We observe a general trend where as womens' BMI increases, they under-report their BMI by a greater amount on average. (Observations with negative values of u are under-reported.) The trend breaks down for the most extreme BMIs above 50, which represents fewer than 2% of all observations. Most BMIs are under-reported. Figure 3.7 plots mens' imputed BMI against the additive error term in BMI generated by proxy (and some self) reporting. We find a very similar trend of under-reporting as in women. The imputed relationships are non-linear, suggesting that the increased flexibility of the imputation method is necessary to model the relationship between observed and true BMI.

3.4.3 Impact of Obesity on Medical Expenditures

We begin by estimating a non-IV two-part model with the same specification as our main model for the full sample, as well as for certain subpopulations: men, women, white, nonwhite, those with private insurance, those with Medicaid, and the uninsured. The results are presented in the left column of Table 3.3. Each estimated coefficient represents the marginal effect of a one unit increase in BMI on annual medical expenditures (Standard errors are below in parentheses). From our non-IV model we find that, on average, gaining an extra unit of BMI raises medical care costs by $74.35 per year.[1] All the marginal effects of a one unit increase in BMI are statistically significant across subgroups, and are under $100. The non-IV results suggest that the effect of elevated BMI is larger for men ($84.22) than for women ($74.41), larger for whites ($89.17) than non-

[1]Appendix table B1 contains the effects of overweight and obese from the non-IV model. We find from the non-IV model that moving from normal weight to obese is associated with a $884 increase in annual medical care costs. This is similar to the non-IV baseline found in Cawley and Meyerhoefer (2012) of $656.

whites ($47.95), and larger for privately insured ($75.41) relative to those with Medicaid ($65.97).

Column 2 of Table 3.3 contain the estimated marginal effects from the IV-GLM model, where both adult BMI and the BMI of their oldest child are the reported BMI from the MEPS. The effect sizes are over twice as large as the point estimates in column 1. We expect using IV-GLM to increase the effect sizes since instrumenting controls for bias in the cross-sectional model due to the endogeneity of BMI. The increase in effect size is akin to the increase found in Cawley and Meyerhoefer (2012), except we find statistically significant effects of BMI on medical care costs for men. This is possibly due to increased statistical power, as we are pooling over more years of data, as opposed to indicating structural change in the effect of BMI on men's medical expenditures over time. After instrumenting we find that the effect of a one BMI unit increase on annual medical expenditure is similar across subpopulations, except for the medicaid covered and the uninsured, which are $101.45 and $104.39 respictively. Like the non-IV estimates, we do not find any statistically significant impact of elevated BMI on expenditures among adults enrolled in Medicaid.

Columns 3 and 4 contain estimates from the RCIV and ARCIV estimators described in Chapter 1. The RCIV estimator uses predicted child and adult BMI in place of the mismeasured values from the MEPS. The RCIV estimates are smaller than the IV estimates. This is consistent with non-classical measurement error in BMI, which can cause downward bias in the RCIV estimator. Column 4 contains the ARCIV estimator in which only the instrument is replaced with it's fitted value. The ARCIV estimates are in very similar to the IV estimates. This is corroborated by the simulation results in Chapter 1, which show that the IV and ARCIV estimates should be very similar.

In the simulation results from chapter 1, we can use the sample covariances to explain the difference between the RCIV and the ARCIV estimates when ARCIV is consistent. Column 5 in table 3.3 contains the ARCIV estimates multiplied by δ from (1.26). We find that this relationship holds as the estimates in column 5 are very close to the RCIV estimates in Column 3. When both BMI and child BMI are systematically under-reported,

the RCIV estimator will be downward. We do find evidence of under-reporting in both adult and child BMI in the MEPS, and conclude that the RCIV estimator is biased downward. We use the ARCIV estimator as the preferable estimator. We acknowledge that these estimates are very similar to the IV model, and under most plausible measurement error scenarios will be identical to the IV estimator.

We find that on average, a one BMI unit increase is associated with a $189.31 increase in annual medical expenditures. The estimated effect is larger for women ($208.18) than for men ($166.09), larger for non-whites ($210.58) than whites ($175.20), and lower than average for privately insured ($163.00) relative to the full sample, but still larger than the effect for the uninsured ($105.06). We still do not find any significant effect of elevated BMI on expenditures among Medicaid enrollees. These effects are significantly larger than those from the non-IV model, suggesting that endogeneity and measurement error play a significant role in attenuating the non-IV estimates.

It is of health policy interest to know the marginal effect of changing weight categories; in particular, the effect of moving from normal weight to overweight or to obese on annual medical expenditures. Table 3.4 presents to marginal effects of overweight and obese on annual medical expenditures. Columns 2 and 3 are estimated by predicting expenditures at the mean BMI within weight classes (healthy weight, overweight, obese) and then taking the difference. We find that moving from healthy weight obese increases total expenditures by $2,383.94. Comparing the estimated effects of overweight and obesity implies that there are non-linearities in the impact of elevated BMI, and that these trends may vary by subpopulation. We observe that the impact of obesity is larger for women than for men.

The impact of obesity in column 3 is lower than the $3,060.38 (2010 dollars) average impact of obesity found in Cawley and Meyerhoefer (2012) as well as the $3,508 average effect in Cawley et al. (2014).[1] We cannot directly compare our estimates in column 3 to these binary effects because they directly estimated a model using a binary treatment effect for obesity. For comparison, we estimate the same method in column 4, and find a

[1]The $2,741 effect in Cawley and Meyerhoefer (2012) is in 2005 dollars. We inflate their estimated effect to 2010 dollars using the CPI.

$3,212.62 increase in annual medical expenditures associated with becoming obese, which is similar to the effect found in Cawley and Meyerhoefer (2012). We also use the binary treatment effect model to estimate a marginal effect of obesity of $2,456.39 for men and $3,840.45 for women.

We prefer to use the estimates in column 2, but report the results using a binary treatment effect in order to compare the more current estimates here to prior research. The estimation method in columns 2 and 3 has some advantages over replacing continuous BMI in the model with an indicator variable for overweight or obese. The main benefit is that we avoid adding categorical variables to ensure that the comparison group is only healthy weight (as opposed to comparing obese to non-obese). Adding additional categorical variables would require additional instruments. An additional concern is that even after imputation, there is probably still idiosyncratic error in the BMI prediction (with mean zero). This is not a problem as such error is essentially classical in nature, and will not bias estimates in an IV model. However, such random additive error will become non-classical when continuous BMI is transformed into a binary indicator, as a mismeasured 1 can only become a 0, and vice versa. In addition to being non-classical error, the error is negatively correlated with the true regressor, and will be negatively correlated with the IV. This can generate exactly the upward bias that imputation is used to alleviate. Empirically, we cannot know the exact degree of error introduced in the binary treatment effect model. We compute the sample correlation between the instrument and the difference in the indicator variables for obesity before and after imputation, and find a negative correlation, but it is near zero and unlikely to result in serious bias. (The sample correlation between this error term and error in the IV is larger at -.04.) Further, the IV estimate of the effect of obesity where the categorical variable is both endogenous and mismeasured is only usable to generate bounds for the true effect. (Frazis and Loewenstein, 2003)

The marginal effects in Tables 3.3 and 3.4 predict the impact of changes in BMI at the mean BMI, or the mean BMI within a weight class. We demonstrate the non-linear relationship between BMI and costs implied in Table 3.4 by predicting medical expendi-

tures across the BMI distribution, allowing a more flexible relationship by including BMI squared as an additional regressor. Figure 3.8 shows the predicted relationship between BMI and annual expenditures for the full sample (men and women pooled). Figures 3.9 and 3.10 show the predicted relationships for men and women respectively. We observe similar patterns as found in Cawley and Meyerhoefer (2012); a J-shaped curve for the pooled sample with a flatter J-shape for women and a U-shaped curve for men, however, our cost curves are estimated with greater precision, and are generally flatter than those found in Cawley and Meyerhoefer (2012). This may be due to our corrections for upward bias from measurement error reducing the magnitude of the marginal effects, and in turn the predictions, particularly in the extreme values of BMI.

In Figure 3.9, we see that expenditures fall sharply as BMI passes above 20, rising slowly at first as BMI increases through the overweight and obese categories. Expenditures rise sharply again only as BMI moves into Grade III obesity (BMI ≥ 40), suggesting that expenditures are relatively similar for men who are overweight as well as obese. In Figure 3.10, the J-shaped curve shows a more gradual decline in expenditures as women move into healthy weight from underweight, and a quicker increase in expenditures as BMI increases into the overweight and obese ranges. Though the non-linear relationship between BMI and predicted expenditures is more gradual for women, predicted expenditures for overweight women are larger than for healthy weight, and predicted expenditures for obese women are larger than those for overweight women. The overlaid distributions show that, although a considerable number of adults are overweight or obese, the highest predicted expenditures are among the few individuals with BMIs well in excess of the cutoff for obese. These few individuals are responsible for a disproportionate share of obesity-related medical costs, and have a large influence on the size of the marginal effects of obesity.

3.4.4 Economic Impact of Obesity

Table 3.5 contains the results from estimating the impact of BMI on third-party expenditures using imputed parent and child BMI. Third-party expenditures represent

a possible channel for obesity-related externalities. Some of the medical costs of obesity may be borne by other enrollees in private insurance pools or by taxpayers in the form of higher expenditures by the Medicaid program. We find that elevated BMI and obesity may be associated with substantial externalities. A one unit increase in BMI raises third-party annual expenditures by $158.18 in the pooled sample, which is 84% of the total effect. Obesity raises third-party annual expenditures by $1,9183.50 , which is 84% of the total effect. The impact of obesity on third-party medical expenditures is significantly higher for women ($2,375.54) than men ($1,559.37). We find a lower than average impact of obesity on third-party medical expenditures among those who are privately insured ($1,762.68, 87% of the total effect), yet do not find a significant impact of BMI on third-party payments among Medicaid enrollees.

To better understand the channels through which obesity affects medical expenditures, we estimate the impact of BMI on specific categories of medical expenditures. Table 3.6 contains the marginal effects of BMI, overweight, and obese on spending on inpatient care, ambulatory care (outpatient services), prescription drugs, and other care (vision, dental, home health, and medical equipment) using imputed adult and child BMI. We find that on average obesity increases spending on inpatient care, ambulatory care, and prescription drugs by roughly the same amount, with the effect being largest for ambulatory care ($884.05). When we estimate separate effects by gender, we find that for women (men), a one unit increase in BMI increases medical spending by $81.54 ($51.35, not significant) for inpatient care, $67.54 ($62.90) for ambulatory care, $62.29 ($55.91) for prescription drugs, and has a negligible impact on spending on other medical care. For women (men), obesity raises medical spending by $1,067.50 ($547.62, not significant) for inpatient care, $932.46 ($714.84) for ambulatory care, and $846.42 ($637.97) for prescription drugs. We find that the sum of the effects for the pooled sample and for women are similar to the total effects in Table 3.4. The sum of the significant effects for men fall short of the estimated total effect, suggesting that there may be a larger effect of BMI on inpatient spending than the point estimate would suggest, but we are unable to precisely estimate it.

We use the IV estimates to estimate the aggregate cost of obesity. Table 3.7 contains estimates of annual medical care expenditures (in 2010 USD) associated with obesity.[1] Under the (admittedly, strong) assumption that the effect of obesity in our subpopulation generalizes to the full non-institutionalized population of adults aged 18 and older, we scaled the costs in the subpopulation used to estimate our model up to the entire adult population by multiplying the subpopulation aggregate costs by the ratio of the US population of adults to the US population of adults with biological children in 2010, or \$US 59.5 billion *(233.7 million/33.9 million) = \$410.2 billion in 2010. Using the same method, we estimate that third-party payments accounts for \$360.5 billion (88%) of the average annual costs of obesity in 2010.

3.5 Discussion

We estimate the effect of obesity on adult medical expenditures using data from the 2000–2010 Medical Expenditures Panel Survey (MEPS). We employ an adjusted conditional expectation IV estimator to correct measurement error in the instrument by imputing measured BMI values into the MEPS from the National Health and Nutrition Examination Survey. Imputation allows us to indirectly identify adults self- and proxy-reporting behavior with regard to their own BMI. Our indirect observations correspond with the findings from empirical studies in two key ways. First, we find evidence of systematic under-reporting, and second, under-reporting is more pronounced for women than for men.

We estimate an non-IV model and find small but significant estimates of the impact of obesity similar to those found in Finkelstein et al. (2009). Unlike this and other studies, we subset to parents of children aged 11-20 due to availability of instruments. Parents may have different underlying health or access to care on average relative to the full US population that will cause our estimates to differ from prior studies. In all non-IV studies of the impact of obesity, we cannot interpret estimated coefficients as causal effects, and

[1]Using the MEPS sampling weights, we determined the population of adults aged 20–64 to be 233.7 million in 2010.

expect that the smaller magnitude of these effects are the result of attenuating bias due to the endogeneity of weight, as well as measurement error in BMI.

We use imputed BMI to correct for reporting error when estimating the effect of obesity on medical expenditures. After correcting for reporting error in the instrument, child BMI, we find that The impact of obesity is $2,383.94. This is lower than the $3,060.38 found by Cawley and Meyerhoefer (2012) and the $3,508 found by Cawley et al. (2014). This estimate in not comparable to our main results because they used a binary treatment effects model for the impact of obesity. For comparison, we use the same specification and find a $3,212.62 increase in annual medical expenditures associated with becoming obese, which is very similar to the effect in Cawley and Meyerhoefer (2012). We prefer to estimate the marginal effect of obesity using continuous BMI as the main regressor. A binary treatment effect model warrants a different interpretation from our estimated marginal effects. Models using binary indicators for obese do not include additional indicators for overweight or underweight. Thus their effects are marginal effects of obese compared to all other BMI classes. Estimates using the binary indicator are best interpreted as the average difference in expenditures associated with being obese compared to any other weight status. Our estimates explicitly uses healthy weight as the control group, and represent the marginal effect of becoming overweight or obese compared to healthy weight. This method essentially compares means within categories. The mean expenditures among he obese likely obscures the significantly higher spending in the right tail, and will likely lead to our method understating the true effect of moving from healthy weight to obese. Comparing both estimates can provide bounds around the true effect. Despite the econometric concerns, researchers may prefer an aggregate estimate of expenditures associated with obesity or to aggregate our estimate of the marginal effect of obesity depending on the policy question they are studying. Thus we provide both sets of estimates using corrected BMI.

There are some limitations to our modeling approach. Like previous research using instrumental variables, the validity of the IV depends on an untestable exclusion restriction. There is a large behavioral genetics literature that supports the genetic linkage between

the weight of biological relatives, and little support for shared environmental effects that are correlated with weight. But genes which influence weight may be inherited alongside genes that also affect demand for medical care. We cannot observe genetic information in the MEPS and acknowledge this possible limitation. Further, BMI is an imperfect measure of fatness. BMI is strongly correlated with more accurate measures of obesity such as body fat or waist circumference (Burkhauser and Cawley, 2008; ONeill, 2015). However, BMI is the only measure of fatness available in the MEPS and the NHANES.

The consistency of our estimates also depend on the validity of the assumptions necessary for imputation. There is no way to observe the true correction equations estimated in the NHANES. We can however rely on the survey design and sample weights for both the MEPS and the NHANES to be random samples of the same population. We also compare estimated misreporting to the behavior observed in prior studies as well as empirically in the NHANES and find that they are similar patterns of under-reporting. We do not explicitly control for variance due to imputation. However, we know from chapter 1 that the ARCIV estimator essentially reduces to the IV estimator (when the is no differential measurement error). Further, the calibration equations have extraordinarily high $R^2 = 0.999$, which means the validation contributes little additional variance. Thus we use the IV-GLM standard errors, as they are very close to the true standard errors.

Despite these limitations, we make an important contribution by highlighting the role of measurement error and modeling in estimating the impact of obesity on medical expenditures. By comparing the imputed BMI to those reported in the MEPS, we can corroborate the under-reporting behavior found in other studies. Our estimates suggest that non-IV estimates of the impact of obesity understate the true effect. We ultimately find that the more recent IV estimates are likely consistent estimates of the true effect by comparing them to error corrected estimates. However, IV estimation is not robust to all possible measurement errors, and the use of validation data can improve the accuracy of point estimation, as well as improve the bounding of true effects when point estimation is not adequate.

Table 3.1: Descriptive Statistics for Men with Biological Children

Variables	**Mean**	**S.D.**	**Min**	**Max**
Has positive medical expenditures	0.79	0.41	0	1
Annual Medical Expenditures*	$3,088.50	7,955.55	1.05	265,048
BMI (MEPS)	28.46	5.06	18.5	70.6
BMI (Imputed)	29.49	5.58	19.03	59.30
Obesity (MEPS)	0.31	0.46	0	1
Obesity (Imputed)	0.39	0.49	0	1
Oldest Child is Female	0.47	0.50	0	1
Oldest Child's Age in months	190.66	30.85	132	239
Oldest Child's BMI (MEPS)	22.48	4.91	6.1	72.7
Oldest Child's BMI (Imputed)	23.86	5.74	11.60	52.23
Hispanic	0.15	0.36	0	1
Black	0.16	0.36	0	1
Other race	0.05	0.22	0	1
Age is 35 - 44	0.41	0.49	0	1
Age is 45 - 55	0.44	0.50	0	1
Age is 55 - 64	0.03	0.18	0	1
Self-reported	0.26	0.44	0	1
High school graduate	0.32	0.47	0	1
Some college	0.22	0.41	0	1
Bachelor's degree	0.17	0.38	0	1
BA plus	0.12	0.33	0	1
Married	0.90	0.30	0	1
People in the household aged 0 - 5	0.17	0.48	0	5
People in the household aged 6 - 17	1.59	1.05	0	8
People in the household aged 18 - 64	2.45	0.82	1	9
People in the household aged 65+	2.45	0.82	0	3
Midwest census region	0.23	0.42	0	1
South census region	0.34	0.47	0	1
West census region	0.24	0.43	0	1
Residence in MSA	0.82	0.39	0	1
Year 2000	0.09	0.28	0	1
Year 2001	0.09	0.29	0	1
Year 2002	0.10	0.29	0	1
Year 2003	0.09	0.29	0	1
Year 2004	0.09	0.29	0	1
Year 2005	0.09	0.29	0	1
Year 2006	0.09	0.29	0	1
Year 2007	0.09	0.29	0	1
Year 2008	0.09	0.28	0	1
Year 2009	0.09	0.28	0	1
Year 2010	0.09	0.28	0	1

Notes: Data: MEPS 2000 – 2010. N = 17,533 (12,907 have positive expenditures). All entries are in 2010 dollars.

Table 3.2: Descriptive Statistics for Women with Biological Children

Variables	**Mean**	**S.D.**	**Min**	**Max**
Has positive medical expenditures	0.88	0.33	0	1
Annual Medical Expenditures*	$3,591.28	8,360.67	1.02	326,153
BMI (MEPS)	27.86	6.54	18.50	78.3
BMI (Imputed)	29.53	7.19	18.50	67.04
Obesity (MEPS)	0.31	0.46	0	1
Obesity (Imputed)	0.41	0.49	0	1
Oldest Child is Female	0.49	0.50	0	1
Oldest Child's Age in months	191.12	30.71	132	239
Oldest Child's BMI (MEPS)	22.75	5.09	6.1	73.5
Oldest Child's BMI (Imputed)	24.13	5.91	11.64	52.57
Hispanic	0.16	0.37	0	1
Black	0.20	0.40	0	1
Other race	0.05	0.21	0	1
Age is 35 - 44	0.50	0.50	0	1
Age is 45 - 55	0.35	0.48	0	1
Age is 55 - 64	0.04	0.19	0	1
Self-reported	0.83	0.38	0	1
High school graduate	0.32	0.47	0	1
Some college	0.26	0.44	0	1
Bachelor's degree	0.16	0.37	0	1
BA plus	0.10	0.29	0	1
Married	0.72	0.45	0	1
People in the household aged 0 - 5	0.17	0.46	0	6
People in the household aged 6 - 17	1.55	1.03	0	8
People in the household aged 18 - 64	2.28	0.89	1	9
People in the household aged 65+	0.04	0.82	0	3
Midwest census region	0.23	0.42	0	1
South census region	0.36	0.48	0	1
West census region	0.23	0.42	0	1
Residence in MSA	0.82	0.38	0	1
Year 2000	0.08	0.27	0	1
Year 2001	0.09	0.29	0	1
Year 2002	0.09	0.29	0	1
Year 2003	0.09	0.29	0	1
Year 2004	0.09	0.29	0	1
Year 2005	0.09	0.29	0	1
Year 2006	0.09	0.29	0	1
Year 2007	0.09	0.29	0	1
Year 2008	0.09	0.29	0	1
Year 2009	0.09	0.29	0	1
Year 2010	0.09	0.29	0	1

Notes: Data: MEPS 2000 – 2010. N = 25,475 (21,386 have positive expenditures). All entries are in 2010 dollars.

Figure 3.1: Distribution of Womens' BMI in the NHANES

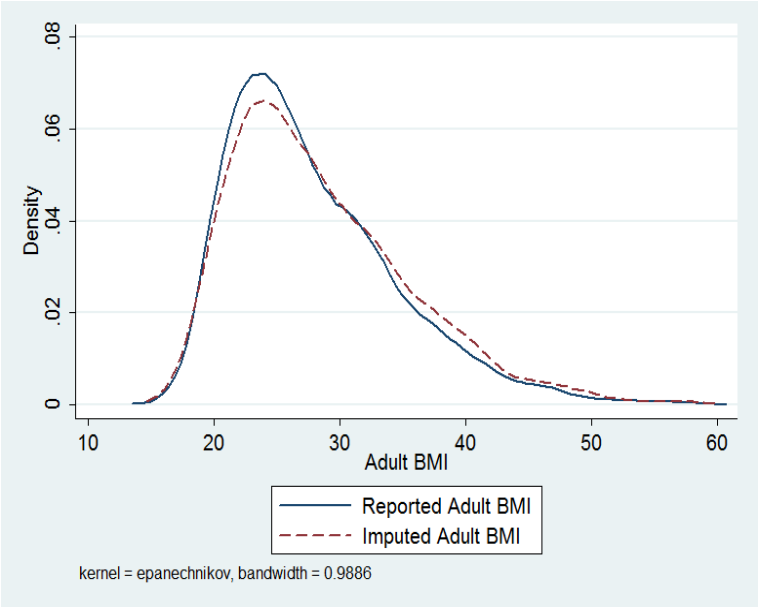

Figure 3.2: Distribution of Mens' BMI in the NHANES

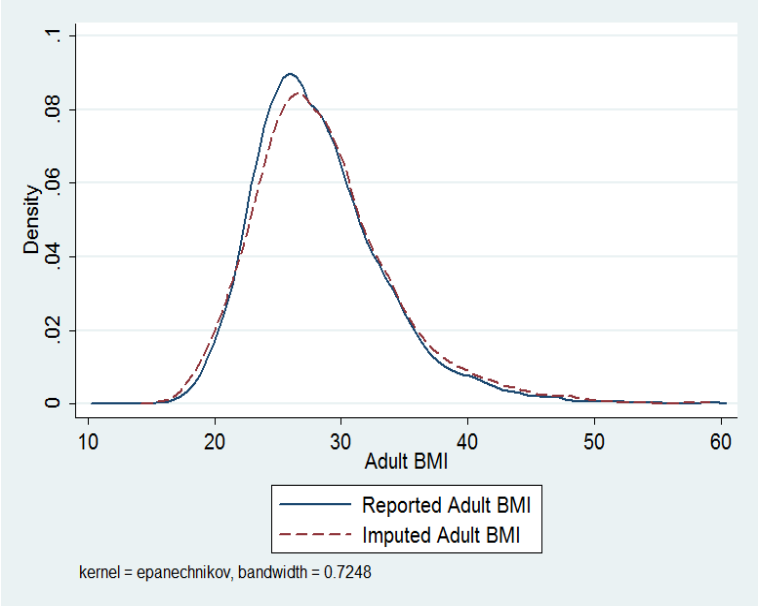

Figure 3.3: Distribution of Womens' BMI in the MEPS

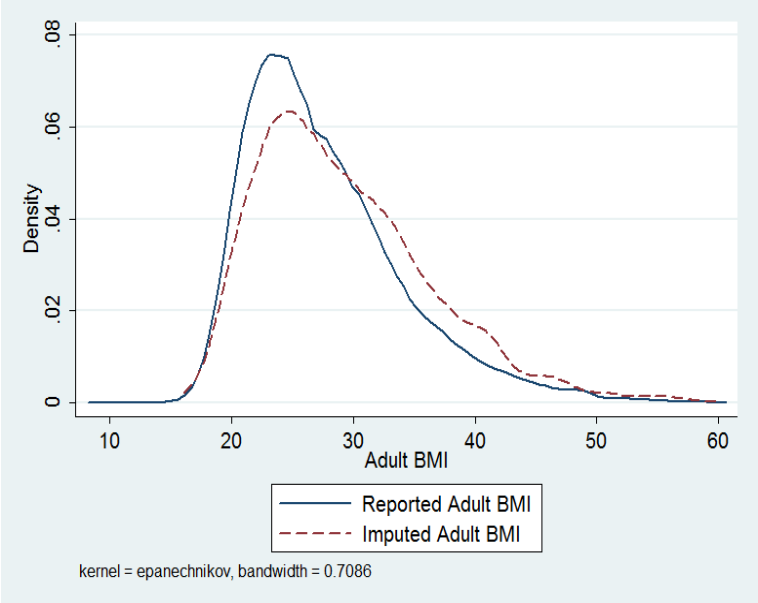

Figure 3.4: Distribution of Mens' BMI in the MEPS

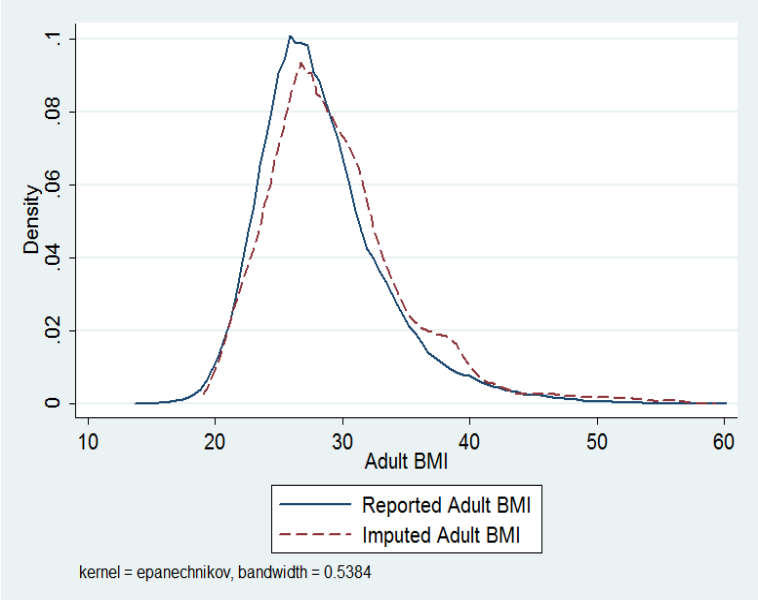

Figure 3.5: Distribution of The Oldest Child's BMI in the MEPS

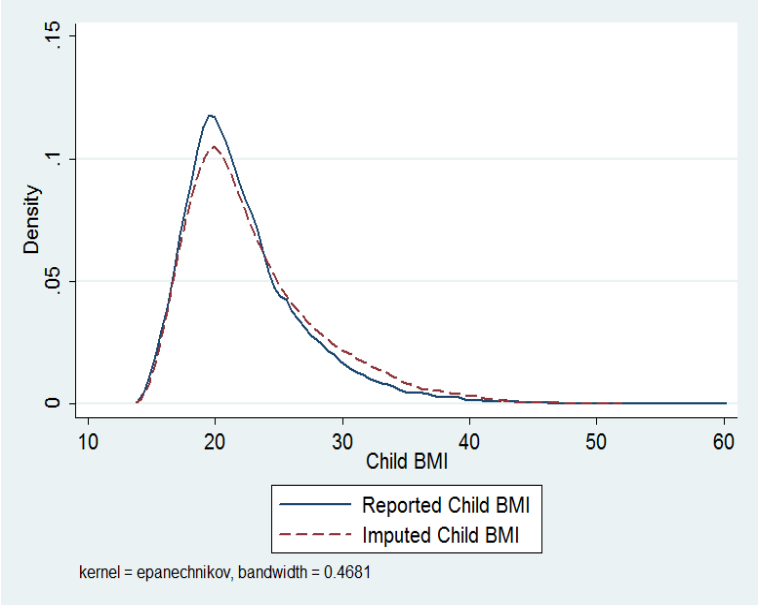

Figure 3.6: Reporting Error in Womens' BMI

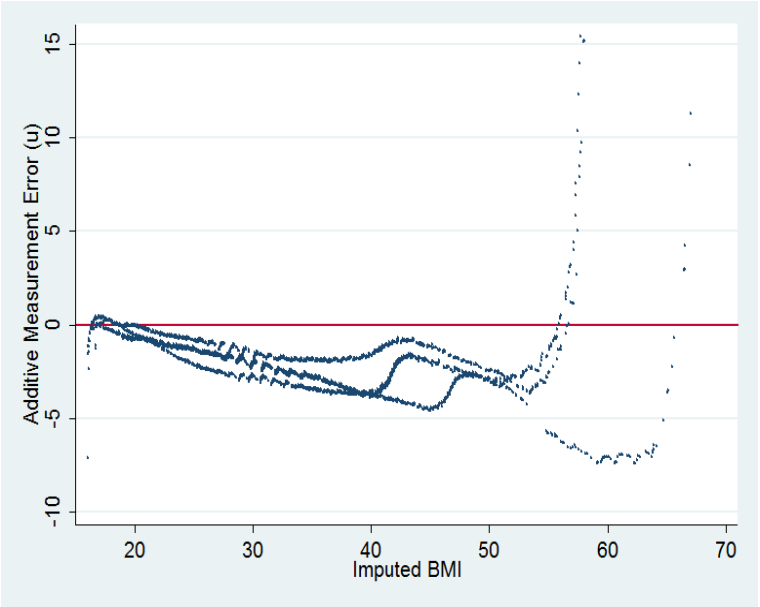

Figure 3.7: Reporting Error in Mens' BMI

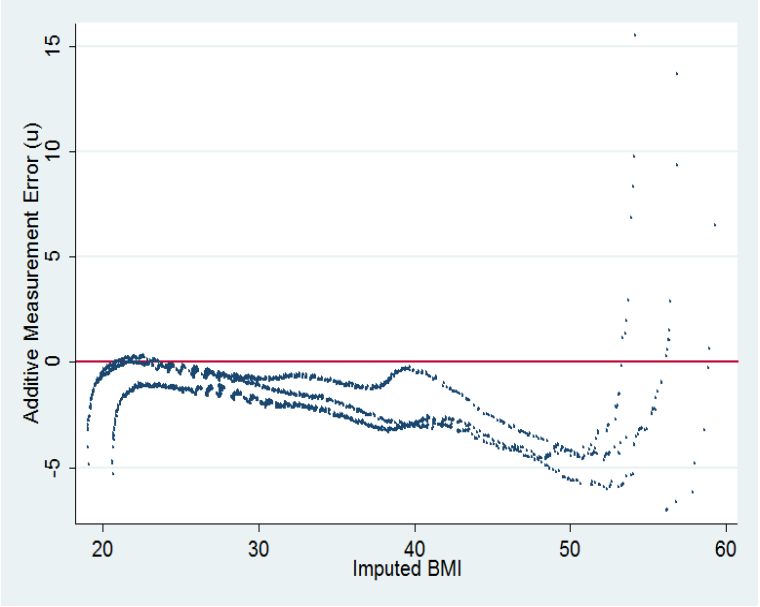

Table 3.3: Marginal Effects of BMI on Annual Medical Expenditures

Estimator:	Non-IV	IV	RCIV	ARCIV	ARCIV$\times\hat{\delta}$
Column	(1)	(2)	(3)	(4)	(5)
Total	74.35***	189.57***	189.31***	173.24***	166.31
N =43,008	(8.96)	(40.01)	(38.83)	(35.31)	(.879)
Men	84.22***	185.64***	166.09***	153.53**	142.80
N =17,533	(14.33)	(78.67)	(72.14)	(66.89)	(.86)
Women	74.41***	194.40***	208.18***	190.11***	185.81
N = 25,475	(10.71)	(38.69)	(40.58)	(36.83)	(.891)
White	89.17***	183.04***	175.20***	161.34***	162.27
N = 21,037	(11.77)	(54.52)	(52.38)	(48.15)	(.926)
Non-White	47.95***	206.81***	210.58***	187.81***	187.25
N = 21,971	(10.36)	(57.21)	(57.46)	(49.93)	(.889)
Private Insurance	75.41***	165.06***	163.00***	148.65***	142.96
N = 28,418	(10.45)	(36.53)	(34.35)	(31.37)	(.877)
Medicaid	65.97***	101.45	102.69	96.57	93.58
N = 5,050	(16.53)	(95.03)	(95.46)	(90.11)	(.911)
Uninsured	20.15***	104.39**	105.06**	96.08**	93.11
N = 8,461	(6.96)	(45.3)	(46.27)	(40.84)	(.886)
Source of BMI:	MEPS	MEPS	MEPS	NHANES	MEPS
IV:	—	MEPS	NHANES	NHANES	NHANES

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parenthesis. First-stage F-statistics in brackets. All entries are in 2010 dollars. For column 5, $\hat{\delta}$ displayed in curly brackets.

Table 3.4: Marginal Effects of Obesity on Annual Medical Expenditures

Estimator:	ARCIV	ARCIV	ARCIV
Population	BMI	Obese	Obese†
Population	BMI	Obese	Obese*
Total	189.31***	2,383.94***	3,212.62***
	(38.83)	(461.22)	(690.88)
Men	166.09***	1,869.42***	2,456.39***
	(72.14)	(767.11)	(1,145.63)
Women	208.18***	2,832.78***	3,840.45***
	(40.58)	(525.52)	(811.13)
White	175.20***	2,244.16***	3,046.23***
	(52.38)	(656.9)	(976.38)
Non-White	210.58***	2,434.78***	3,230.20***
	(57.46)	(541.56)	(860.96)
Private Insurance	163.00***	2,042.83***	2,710.33***
	(34.35)	(423.13)	(617.53)
Medicaid	102.69	1,461.34	1,724.35
	(95.46)	(1,307.59)	(1,644.63)
Uninsured	105.06**	1,291.86***	1,825.61**
	(46.27)	(486.22)	(811.09)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parenthesis. All entries are in 2010 dollars. Effects of moving form normal weight to obese are estimated by taking the difference of predicted expenditures at mean BMI within the respective weight class. † Marginal effect of obesity estimated by regressing expenditures on a binary indicator equal to one if BMI$\geq$ 30.

Figure 3.8: Predicted Expenditures By BMI - Men and Women Pooled

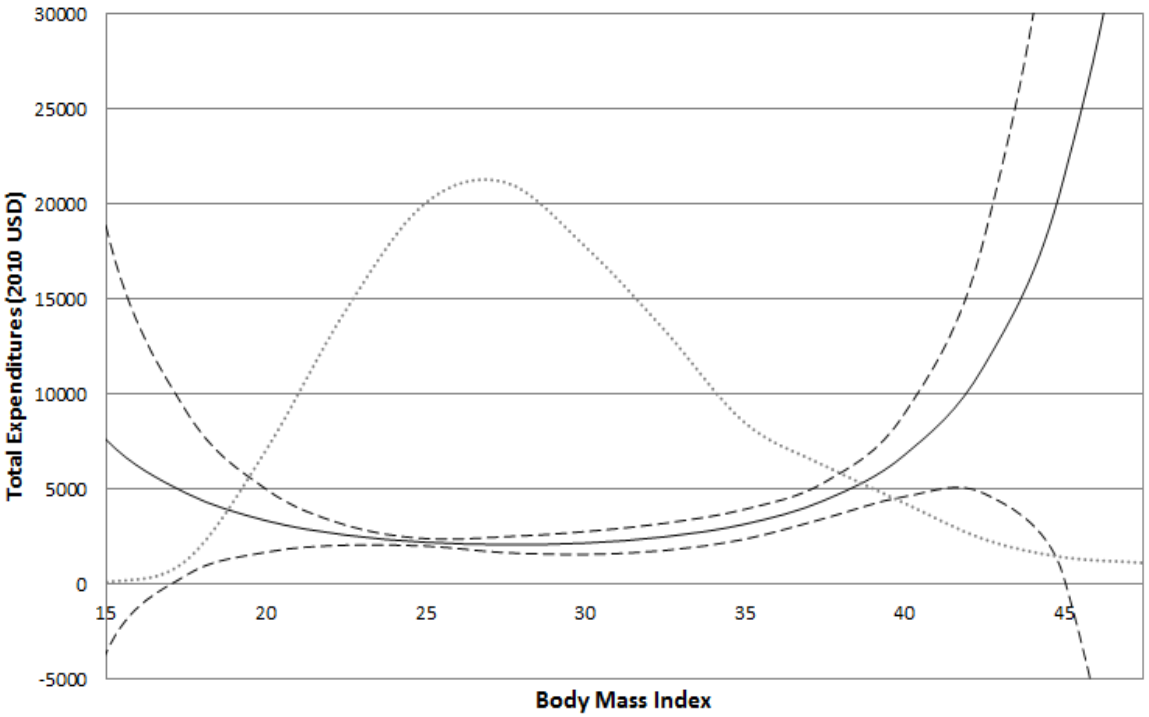

Fig 3.8. Predicted relationship between BMI and annual medical expenditures for all adults with biological children. Notes: Data: MEPS 2000–2010. Expenditures are in 2010 USD. Dashed lines represent the 90% confidence interval, which has been adjusted for the complex design of the MEPS. Medical expenditures are denoted by the solid line, while the distribution of individuals in the population is indicated by the dotted line. BMI is predicted in the MEPS using correction equations calibrated using the NHANES.

Figure 3.9: Predicted Expenditures By BMI - Men

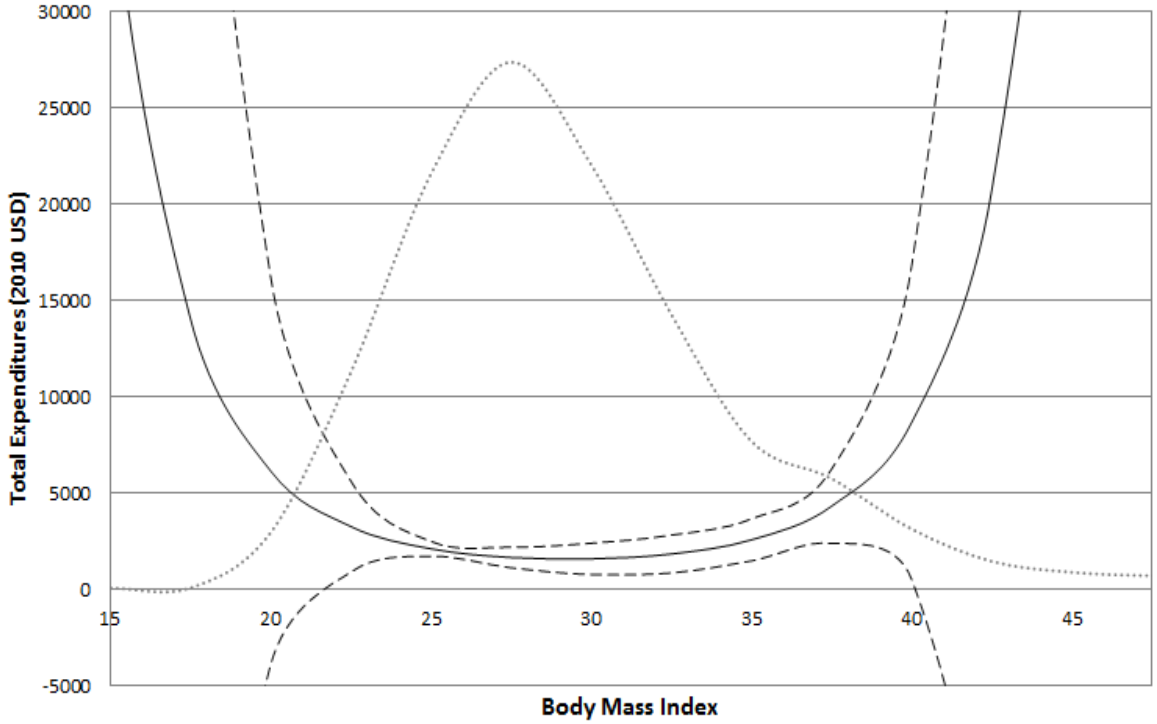

Fig 3.9. Predicted relationship between BMI and annual medical expenditures for men with biological children. Notes: Data: MEPS 2000–2010. Expenditures are in 2010 USD. Dashed lines represent the 90% confidence interval, which has been adjusted for the complex design of the MEPS. Medical expenditures are denoted by the solid line, while the distribution of individuals in the population is indicated by the dotted line. BMI is predicted in the MEPS using correction equations calibrated using the NHANES.

Figure 3.10: Predicted Expenditures By BMI - Women

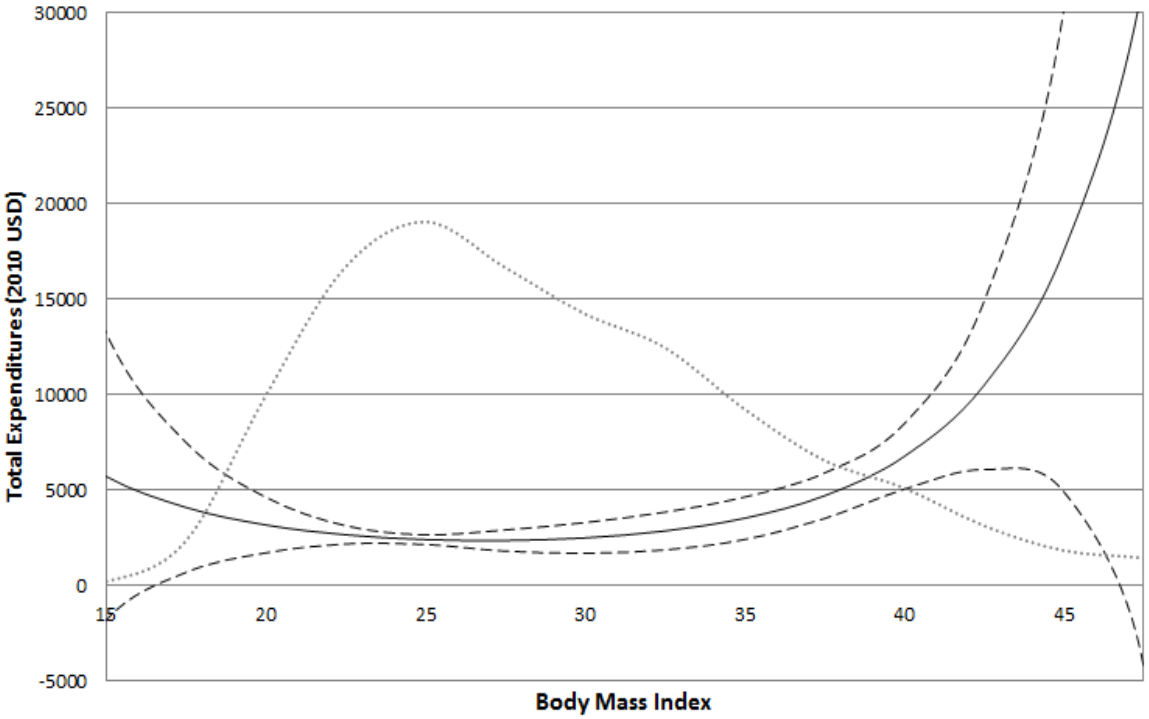

Fig 3.10. Predicted relationship between BMI and annual medical expenditures for women with biological children. Notes: Data: MEPS 2000–2010. Expenditures are in 2010 USD. Dashed lines represent the 90% confidence interval, which has been adjusted for the complex design of the MEPS. Medical expenditures are denoted by the solid line, while the distribution of individuals in the population is indicated by the dotted line. BMI is predicted in the MEPS using correction equations calibrated using the NHANES.

Table 3.5: Marginal Effects of BMI, Overweight and Obese on Third-Party Expenditures

Estimator:	ARCIV	ARCIV
Population	BMI	Obese
Total	158.18***	1,983.50***
	(34.37)	(403.1)
Men	138.52***	1,559.37***
	(62.18)	(656.83)
Women	175.85***	2,375.54***
	(38.95)	(493.65)
White	146.63***	1,873.44***
	(44.55)	(553.25)
Non-White	177.33***	2,047.33***
	(53.94)	(506.43)
Private Insurance	140.94***	1,762.68***
	(30.95)	(379.08)
Medicaid	92.26	1,321.18
	(88.17)	(1,216.2)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parenthesis. All entries are in 2010 dollars. Effects of moving form normal weight to obese are estimated by taking the difference of predicted expenditures at mean BMI within the respective weight class.

Table 3.6: Marginal Effects of BMI, Overweight and Obese on Categories of Expenditures

Men and Women Pooled			
Category of Care	BMI	Overweight	Obese
Inpatient	66.40***	201.66***	801.13***
	(23.19)	(36.52)	(237.79)
Ambulatory	69.21***	274.44***	884.05***
	(18.04)	(53.68)	(224.79)
Prescription Drugs	59.68***	209.66***	754.59***
	(9.8)	(19.15)	(111.39)
Other	-3.54	-16.89	-45.59
	(3.32)	(16.76)	(42.)
Men			
Category of Care	BMI	Overweight	Obese
Inpatient	51.35	145.48**	547.62
	(40.77)	(78.49)	(382.46)
Ambulatory	62.90*	224.49***	714.84**
	(33.44)	(82.91)	(363.24)
Prescription Drugs	55.91***	183.96***	637.97***
	(14.11)	(24.49)	(145.62)
Other	-1.58	-6.11	-18.45
	(5.35)	(23.45)	(61.46)
Women			
Category of Care	BMI	Overweight	Obese
Inpatient	81.54***	246.29***	1,067.50***
	(29.49)	(42.16)	(321.83)
Ambulatory	67.54***	286.84***	932.46***
	(15.12)	(51.33)	(207.14)
Prescription Drugs	62.29***	225.34***	846.42***
	(11.43)	(23.46)	(139.17)
Other	-4.98	-25.38	-67.59
	(4.27)	(22.96)	(55.98)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parenthesis. All entries are in 2010 dollars. Effects of moving form normal weight to overweight and obese are estimated by taking the difference of predicted expenditures at mean BMI within the respective weight class.

Table 3.7: Annual Medical Costs of Obesity for Adults Aged 20 - 64 with Biological Children

Year	Total Expenditures	Third-Party Expenditures	Population
2000	29.0 (18.2, 39.9)	24.1 (14.3, 33.9)	30,627,982
2001	39.7 (24.7, 54.8)	32.7 (19.6, 45.7)	33,968,508
2002	40.7 (25.3, 56.2)	34.2 (20.4, 48.0)	35,476,220
2003	40.1 (25.4, 54.9)	33.0 (20.2, 45.8)	34,686,341
2004	45.9 (28.6, 63.1)	38.4 (23.2, 53.5)	34,145,908
2005	44.7 (27.2, 62.1)	37.0 (21.5, 52.5)	34,667,712
2006	45.3 (28.8, 61.7)	37.4 (23.0, 51.8)	34,578,514
2007	47.7 (29.5, 65.9)	41.3 (24.5, 58.0)	34,395,320
2008	46.8 (30.2, 63.5)	39.6 (24.7, 54.5)	34,198,229
2009	55.7 (36.2, 75.2)	49.0 (30.7, 67.3)	34,530,901
2010	59.5 (38.5, 80.5)	52.3 (32.7, 71.9)	33,962,749
2000 - 2010	45.0 (29.0, 61.0)	38.1 (23.7, 52.4)	34,112,580
average	44.1 (28.5, 59.7)	37.5 (23.5, 51.4)	34,112,580

Notes: Data: MEPS 2000–2010. All expenditures are in billions of 2010 dollars. 90% Confidence intervals in parentheses are adjusted for the complex design of the MEPS. Marginal effects of obesity are estimated by regressing expenditures on a binary indicator equal to one if BMI$\geq$ 30. Estimates are aggregated using obesity indicators where BMI is corrected for reporting error.

References

CDC / National Center for Health Statistics (2013). Specifying weighting parameters. 38, 80, 112

CDC / National Center for Health Statistics (2014). About the national health and nutrition examination survey. 38, 79

Agency for Healthcare Research and Quality (AHRQ) (2007). Clinical classification software (ccs) for icd-9-cm fact sheet. 79

Bhattacharya, J. and N. Sood (2005). Health Insurance and the Obesity Externality. NBER Working Paper No. 11529. 28, 70

Biro, F. M. and M. Wien (2010). Childhood obesity and adult morbidities. *American Journal of Clinical Nutrition 91*(5), 1499S–15005S. 26

Bound, J., C. Brown, and N. Mathiowetz (2001). Measurement error in survey data. *Handbook of Econometrics 5*, 3705–3843. 2, 3, 4, 5, 6, 7, 29, 33, 35, 71, 75

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450. 10

Brettschneider, A.-K., U. Ellert, and A. Schaffrath Rosario (2012). Comparison of bmi derived from parent-reported height and weight with measured values: Results from the german kiggs study. *International Journal of Environmental Research and Public Health 9*(2), 632–647. 32

Brown, H. S., A. Perez, Y.-P. Li, D. Hoelscher, S. Kelder, and R. R. (2007). The cost-effectiveness of a school-based overweight program. *International Journal of Behavioral Nutrition and Physical Activity 4*(1), 47. 51

Burkhauser, R. V. and J. Cawley (2008). Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research. *Journal of Health Economics 27*(2), 519–529. 28, 40, 50, 90

Burkhauser, R. V., J. Cawley, and M. D. Schmeiser (2009). The timing of the rise in u.s. obesity varies with measure of fatness. *Economics & Human Biology 7*(3), 307 – 318. 69

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective* (2 ed.). Chapman & Hall/CRC. 5, 8, 9, 14, 30, 72

Cawley, J. (2000). An instrumental variables approach to measuring the effect of body weight on employment disability. *Health Services Research 5*(5 Pt 2), 1159–1179. 3, 4, 9, 11, 18, 31, 35, 36, 73, 75

Cawley, J. (2004). The impact of obesity on wages. *The Journal of Human Resources 39*(1), 451–474. 4, 9, 28, 29, 31, 35, 36, 71, 73, 75

Cawley, J. and C. Meyerhoefer (2012). The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics 31*(1), 219–230. 2, 3, 4, 14, 16, 28, 31, 43, 70, 71, 73, 77, 82, 83, 84, 85, 86, 89

Cawley, J., C. Meyerhoefer, A. Biener, M. Hammer, and N. Wintfeld (2014). Savings in medical expenditures associated with reductions in body mass index among us adults with obesity, by diabetes status. *PharmacoEconomics*. 16, 31, 43, 70, 73, 77, 84, 89

Centers for Disease Control and Prevention (CDC) (2014, July). About BMI for children and teens. `http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html`. 53

Chen, X., H. Hong, and D. Nekipelov (2011). Nonlinear models of measurement errors. *Journal of Economic Literature 49*(4), 901–37. 40

Courtemanche, C., J. C. Pinkston, and J. Stewart (2014). Adjusting body mass for measurement error with invalid validation data. *IZA Discussion Paper No. 8009*. 11, 18, 35, 36, 71, 75, 76

Dietz, W. H. (1998). Health consequences of obesity in youth: Childhood predictors of adult disease. *Pediatrics 101*(518). 26

Dietz, W. H. and T. N. Robinson (2005). Overweight children and adolescents. *The New England Journal of Medicine 352*(1), 2100 – 2109. 26

Dixon, J. B. (2010). The effect of obesity on health outcomes. *Molecular and Cellular Endocrinology 316*(2), 104 – 108. 69

Dutton, D. and L. McLaren (2014). The usefulness of "corrected" body mass index vs. self-reported body mass index: comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using canadian population-based data. *BMC Public Health 14*(1), 430. 35, 75

Finkelstein, E. A., I. C. Fiebelkorn, and G. Wang (2003). National medical spending attributable to overweight and obesity: How much, and who's paying? *Health Affairs W3-219-26*. 69

Finkelstein, E. A. and J. G. Trogden (2008). Public health interventions for addressing childhood overweight: Analysis of the business case. *American Journal of Public Health 93*(3), 411415. 27, 39, 48

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz (2009). Annual medical spending attributable to obesity: Payer-and service-specific estimates. *Health Affairs 28*(5), 822–831. 2, 14, 51, 69, 71, 77, 88

Flegal, K. M., M. D. Carroll, B. K. Kit, and C. L. Ogden (2012). Prevalence of obesity and trends in the distribution of body mass index among us adults, 1999-2010. *JAMA 307*(5), 491–497. 69, 80

Fontaine, K. R. and S. J. Bartlett (2000). Access and use of medical care among obese persons. *Obesity Research 8*(5), 403–406. 28, 70

Frazis, H. and M. A. Loewenstein (2003). Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. *Journal of Econometrics* (17), 151–178. 40, 85

Freedman, D. S., L. K. Khan, M. K. Serdula, C. L. Ogden, and W. H. Dietz (2006). Racial and ethnic differences in secular trends for childhood bmi, weight, and height. *Obesity 14*(2), 301–308. 27

Freedman, L. S., D. Midthune, R. J. Carroll, and V. Kipnis (2009). A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Stat Med. 27*(25), 5195–5216. 22

Fuller, W. A. (1987). *Measurement Error Models* (1 ed.). Hoboken, NJ: John Wiley & Sons, Inc. 5

Gillum, R. and C. T. Sempos (2005). Obesity and disability, time to act. *Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the Third National Health and Nutrition Examination Survey 4*(27). 29

Goodman, E., B. R. Hinden, and S. Khandelwal. Accuracy of teen and parental reports of obesity and body mass index. *Pediatrics*. 32

Gorber, S. C., M. Tremblay, D. Moher, and B. Gorber (2007). A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews 8*, 307326. 3, 32, 71, 74

Grilo, C. M. and M. E. Pogue-Geile (1991). The nature of environmental influences on weight and obesity: A behavior genetic analysis. *Psychological Bulletin 110*(3), 520–537. 30, 31, 72, 73

Guo, S. S. and W. C. Chumlea (1999). Tracking of body mass index in children in relation to overweight in adulthood. *The American Journal of Clinical Nutrition 70*(1), 145–148. 26

Guo, Y. and R. J. Little (2011). Regression analysis with covariates that have heteroscedastic measurement error. *Statistics in Medicine 30*(18), 22782294. 7

Haberstick, B., J. Lessem, M. McQueen, J. Boardman, C. Hopfer, A. Smolen, and J. Hewitt (2010). Stable genes and changing environments: body mass index across adolescence and young adulthood. *Behavioral Genetics 40*(4), 495–504. 30, 31, 72, 73

Han, J. C., D. A. Lawlor, and S. Y. S. Kimm (2010). Childhood obesity. *Lancet 375*(9727), 1737–1748. 26

Hardin, J. W. and R. J. Carroll (2003a). Measurement error, glms, and notational conventions. *The Stata Journal*, 1–12. 15, 40, 77

Hardin, J. W. and R. J. Carroll (2003b). Variance estimation for the instrumental variables approach to measurement error in generalzed liner models. *The Stata Journal*, 1–10. 40, 77

Hardin, J. W. and J. M. Hilbe (2012). *Generalized Linear Models and Extensions* (3 ed.). Stata Press. 15, 40, 77, 111

Hardin, J. W., H. Schmiediche, and R. J. Carroll (2003). Instrumental variables, bootstrapping, and generalized linear models. *The Stata Journal 3*(4), 351–360. 15, 16, 40, 77

Hu, F. B. (2008). *Obesity Epidemiology*. New York: Oxford University Press. 69

Inoue, A. and G. Solon (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics 92*(3), 557–561. 8

Jones, A. (2000). Health econometrics. In J. C. A. and P. N. J. (Eds.), *Handbook of Health Economics* (1 ed.), Volume 1, Chapter 06, pp. 265–344. 14, 39, 76

Kline, B. and J. L. Tobias (2008). The wages of bmi: Bayesian analysis of a skewed treatment-response model with nonparametric endogeneity. *Journal of Applied Econometrics 23*(6), 767–793. 31, 73

Langford, E., N. Schwertman, and M. Owens (2001). Is the property of being positively correlated transitive? *The American Statistician 55*(4), pp. 322–325. 33, 74

Latner, J., A. Stunkard, and G. Wilson. (2005). Stigmatized students: age, sex, and ethnicity effects in the stigmatization of obesity. *Obesity Research 13*(7), 1226–31. 33

Lee, L. F. and J. H. Sepanski (1995). Estimation of linear and nonlinear errors-in-variables models using validation data. *Journal of the American Statistical Association 9*(429), 130–140. 3, 8, 12, 14, 35, 36, 75

Lindeboom, M., P. Lundborg, and B. van der Klaauw (2010). Assessing the impact of obesity on labor market outcomes. *Economics and Human Biology* (8), 309–319. 31, 73

Lundahl, A., K. M. Kidwell, and T. D. Nelson (2014). Parental underestimates of child weight: A meta-analysis. 3, 32, 42, 48

Lyles, R. H. and L. L. Kupper (1997). A detailed evaluation of adjustment methods for multiplicative measurement error in linear regression with applications in occupational epidemiology. *Biometrics 53*(3), 1008–1025. 8

Maes, H. H. M., M. C. Neale, , and L. J. Eaves (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics 27*(4), 325–351. 30, 31, 73

Mander, A. and D. Clayton (1999, January). HOTDECK: Stata module to impute missing values using the hotdeck method. Statistical Software Components, Boston College Department of Economics. 115

Manning, W. G. and J. Mullahy (2001). Estimating log models: to transform or not to transform? *Journal of Health Economics* (20), 461–494. 14, 17, 39, 77

Monheit, A. C., J. P. Vistnes, and J. A. Rogowski (2009). Overweight in adolescents: Implications for health expenditures. *Economics & Human Biology 7*(1), 55 – 63. 2, 27, 39, 48

National Center for Health Statistics (NCHS) (2014, September). Prevalence of overweight and obesity among children and adolescents: United states, 1963–1965 through 2011–2012. `http://www.cdc.gov/nchs/data/hestat/obesity_child_11_12/obesity_child_11_12.htm`. 53

NIH (1998, Sep). NHLBI Obesity Education Initiative Expert Panel on the Identification, Evaluation, and Treatment of Obesity in Adults (US). Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report. Bethesda (MD): National Heart, Lung, and Blood Institute. 27

O'Connor, D. P. and J. J. Gugenheim (2011). Comparison of measured and parents' reported height and weight in children and adolescents. *Obesity 19*(5), 1040–1046. 32

Ogden, C. L., M. D. Carroll, B. K. Kit, and K. M. Flegal (2012). Prevalence of obesity and trends in body mass index among us children and adolescents, 1999-2010. *JAMA 307*(5), 483–490. 26

Ogden, C. L., K. M. Flegal, M. D. Carroll, and C. L. Johnson (2002). Prevalence and trends in overweight among us children and adolescents, 1999-2000. *JAMA 288*(14), 1728–1732. 26

O'Neill, D. and O. Sweetman (2013). The consequences of measurement error when estimating the impact of obesity on income. *IZA Journal of Labor Economics 2*(1), 3. 3, 4, 8, 22, 29, 71

ONeill, D. (2015). Measuring obesity in the absence of a gold standard. *Economics & Human Biology 17*(0), 116 – 128. 40, 50, 90

Pelone, F., M. L. Specchia, M. A. Veneziano, S. Capizzi, S. Bucci, A. Mancuso, W. Ricciardi, and A. G. de Belvis (2012). Economic impact of childhood obesity on health systems: a systematic review. *Obesity Reviews 13*(5), 431–440. 27

Price, R. A. and I. I. Gottesman (1991). Body fat in identical twins reared apart: Roles for genes and environment. *Behavior Genetics 21*(1), 1–7. 30, 73

Puhl, R. M. and C. A. Heuer (2010). Obesity stigma: important considerations for public health. *health 24*, 252. 33

Reed, D. and R. A. Price (1998). Estimates of the heights and weights of family members: accuracy of informant reports. *International Journal of Obesity 22*, 827–835. 32, 74

Rowland, M. L. (1990). Self-reported weight and height. *American Journal of Clinical Nutrition 52*(6), 1125–1133. 29

Shiely, F., K. Hayes, I. J. Perry, and C. C. Kelleher (2013). Height and weight bias: The influence of time. *PLOS ONE 8*(1), e54386. 32

Skelton, J. A., S. R. Cook, P. Auinger, J. D. Klein, and S. E. Barlow (2009). Prevalence and trends of severe obesity among us children and adolescents. *Acad Pediatr 9*(5), 322329. 53

Smith, G. D., J. A. Sterne, A. Fraser, P. Tynelius, D. A. Lawlor, , and F. Rasmussen (2009). The association between bmi and mortality using offspring bmi as an indicator of own bmi: large intergenerational mortality study. *British Medical Journal 339:b5043*. 30, 72

Stock, J. H., J. H. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics 20*(4), 518. 10, 30, 72

Stommel, M. and C. Schoenborn (2009). Accuracy and usefulness of bmi measures based on self-reported weight and height: findings from the nhanes & nhis 2001-2006. *BMC Public Health 9*(1), 421. 3, 71, 74

Stunkard, A. J., T. I. Srensen, C. Hanis, T. W. Teasdale, R. Chakraborty, W. J. Schull, and F. Schulsinger (1986). An adoption study of human obesity. *New England Journal of Medicine 314*(4), 193–198. 30, 73

Srensen, T. I. and A. J. Stunkard (1993). Does obesity run in families because of genes? an adoption study using silhouettes as a measure of obesity. *Acta PsychiatricaScandanavia 370*(Suppl), 67–72. 30, 73

Thorpe, K. E., C. S. Florence, D. H. Howard, and P. Joski (2004). The impact of obesity on rising medical spending. *Health Affairs*. 69

Trasande, L. (2010). How much should we invest in preventing childhood obesity? *Health Affairs 29*(3), 372378. 51

Trasande, L. and S. Chatterjee (2009). The impact of obesity on health service utilization and costs in childhood. *Obesity 17*(9), 1749–1754. 27, 69

Trasande, L. and B. Elbel (2012). The economic burden placed on healthcare systems by childhood obesity. *Expert Review of Pharmacoeconomics & Outcomes Research 12*(1), 39–45. 27

Trasande, L., Y. Liu, G. Fryer, and M. Weitzman (2009). Effects of childhood obesity on hospital care and costs, 1999-2005. *Health Affairs 28*(4), 751–760. 27, 28, 69

Villanueva, E. V. (2001). The validity of self-reported weight in us adults: a population based cross-sectional study. *BMC Public Health 1*(11). 29, 32, 74

Wang, L. Y., Q. Yang, R. Lowry, and H. Wechsler (2003). Economic analysis of a school-based obesity prevention program. *Obesity Research 11*(11), 1313–1324. 51

Wardle, J., S. Carnell, C. M. Haworth, and R. Plomin (2008). Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *The American Journal of Clinical Nutrition 87*, 398–404. 30, 31, 72, 73

Whitlock, E. P., E. A. O'Connor, S. B. Williams, T. L. Beil, and K. W. Lutz (2010). Effectiveness of weight management interventions in children: A targeted systematic review for the USPSTF. *Pediatrics 125*(2), e396–e418. 52

Appendix A

A.1 IRLS for the Log-Gamma Model

Hardin and Hilbe (2012) derive the general IRLS algorithm. To avoid confounding notation with names for covariates, we define weighting matrix Ξ, and a vector of pseudodata ζ such that,

$$\Xi = g'(\mu)^{-2}V(\mu)^{-1}, \tag{A.1a}$$

$$\zeta = g'(\mu)(y-\mu)+\eta. \tag{A.1b}$$

The IRLS algorithm first uses initial values of μ and η to calculate Ξ and ζ. Second, β is chosen to minimize

$$\sum_{i=1}^{n} \Xi_i(\zeta_i - X_i\beta)^2. \tag{A.2}$$

The solution $\beta^{(r)}$ is then used to evaluate $\eta^{(r+1)}$ and the procedure is reiterated.

We state the full IRLS algorithm in terms of the Log-Gamma model:

IRLS Algorithm for Log-Gamma Model

Set initial values

1. $\mu^{\circ} = (y + \bar{y})/2$
2. $\eta^{\circ} = ln(\mu^{\circ})$
3. Set Dev_{old}
4. Set Dev_{new}
5. Set ΔDev

While $|\Delta Dev| >$ tolerance {

6. Calculate $\Xi = g'(\mu^{(r-1)})^{-2} \cdot V(\mu^{(r-1)})^{-1} = I$
7. Calculate $\zeta = g'(\mu^{(r-1)}) + \eta = (y_i - \mu_i^{(r-1)})/\mu_i^{(r-1)} + X\beta^{(r-1)}$
8. $\beta^r = (X'\Xi X)^{-1}(X'\Xi\zeta)$
9. $\eta^r = X\beta^r$
10. $\mu^r = g^{-1}(\eta^r) = exp(X\beta^r)$
11. $Dev_{old} = Dev_{new}$
12. $Dev_{new} = 2\sum\left\{\frac{y_i - \mu^r}{\mu^r} - ln\left(\frac{y_i}{\mu^r}\right)\right\}$
13. $\Delta Dev = Dev_{new} - Dev_{old}$
14. }

An interesting special case result of the Log-Gamma specification is that the weighting matrix Ξ devolves to the identity matrix I. Thus IRLS for the Log-Gamma model can be characterizes as simply as iterated least squares, where pseudodata $\zeta = (y_i - \mu_i)/\mu_i) + \eta$ is iteratively regressed on X. If the IRLS algorithm converges to a fixed β^r, then it is a MLE of β.

A.2 NHANES: Constructing 12-Year Sample Weights

The NHANES weight construction example code from CDC / National Center for Health Statistics (2013) is as follows:

Example 6: How to combine 10 years of data that include 1999-2000 through 2007-2008

Answer: You must use the 4-year weights provided for 1999-2002 (WTMEC4YR) with the 2-year weights for 2003-2004, 2005-2006 and 2007-2008 (WTMEC2YR) to create a 10-year weight variable (MEC10YR)..

For 10 years of data from 1999-2008 a weight should be constructed as:

```
if sddsrvyr=1 or sddsrvyr=2 then
MEC10YR = 2/5 * WTMEC4YR ; /* for 1999-2002 */

if sddsrvyr=3 or sddsrvyr=4 or sddsrvyr=5 then
MEC10YR = 1/5 * WTMEC2YR ; /* for 2003-2008 */
```

Again, future years of data can continue to be added using the same methods as above for combining cycles by taking the correct proportion of the 4-year and 2-year weights.

We extrapolate from the 10 year code the method to construct 12 year weight (MEC12YR) from 1999 - 2010

```
gen MEC12YR = .

replace MEC12YR = WTMEC4YR/3 if sddsrvyr==1

replace MEC12YR = (2/6)*WTMEC4YR if sddsrvyr==2

replace MEC12YR = (1/6)*WTMEC2YR if sddsrvyr==3

replace MEC12YR = (1/6)*WTMEC2YR if sddsrvyr==4

replace MEC12YR = (1/6)*WTMEC2YR if sddsrvyr==5

replace MEC12YR = (1/6)*WTMEC2YR if sddsrvyr==6
```

Appendix B

B.3 Additional Results

Table B1: Marginal Effects of BMI, Overweight and Obese on Third-Party Expenditures from Non-IV Model

Population	BMI	Overweight	Obese	Obese†
Total	69.18***	307.52***	884.68***	906.98***
	(9.05)	(36.3)	(116.08)	(113.65)
Men	80.39***	307.89***	916.96***	903.38***
	(14.18)	(46.74)	(161.02)	(162.06)
Women	68.33***	321.01***	935.21***	926.99***
	(10.73)	(46.64)	(149.02)	(137.66)
White	86.27***	382.79***	1,108.16***	1081.76***
	(11.77)	(46.58)	(152.69)	(152.2)
Non-White	40.49***	181.94***	513.65***	566.18***
	(10.76)	(43.21)	(135.61)	(131.88)
Private Insurance	73.60***	325.73***	922.42***	905.08***
	(10.29)	(41.42)	(129.94)	(129.65)
Medicaid	41.49**	209.37***	595.72**	654.15**
	(17.03)	(76.27)	(244.93)	(309.31)
Uninsured	19.64***	90.49***	248.80***	238.13**
	(6.97)	(29.66)	(88.89)	(101.89)
Source of BMI:	MEPS	MEPS	MEPS	MEPS
IV:	n/a	n/a	n/a	n/a

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars. Effects of moving form normal weight to overweight and obese are estimated by taking the difference of predicted expenditures at mean BMI within the respective weight class. †

B.4 Hotdeck Imputation: Impact of Youth Obesity

An alternative to regression based imputation is hotdeck imputation. Hotdeck imputation does not use correction equations to predict measured values of BMI, but instead stochastically chooses an observed BMI value in the validation sample and assigns it to an observation in the principal sample. Hotcdeck allows for this imputation to be carried out within exclusive strata defined by covariates shared in both datasets.[1] The advantage of hotdeck imputation is that it does not impose a functional form on the validation equation. This can allow for improved fitting of error corrected BMI when the relationship between true BMI and it's surrogate (typically observed BMI) is highly irregular.

Hotdeck imputation demands the same assumptions as regression based imputation. In particular, the matching surrogate must be transportable such that,

$$E(x|w_M, Z_M) = E(x|w_N, Z_N), \tag{A.3}$$

where w and covariates Z are used to define the strata.

In order for transportability to be satisfied empirically, the strata are defined as to finely divide the data, so that within a strata, we can assume that reporting behavior is homogeneous. We alternatively impute children within strata defined by gender, age (in years) 100 bins for each percentile of BMI.[2] For adults, we generate strata using gender, categories for age, and 100 bins for each percentile of BMI.

Table B2 contains the estimated coefficients from Table 2.7 where hotdeck imputation is the validation method. We find that the ARCIV estimates using hotdeck imputed BMI are close to those from regression based imputation in Table 2.7. The RCIV estimates are again close for the pooled sample, but differ significantly when estimated on separate subsamples by gender. Further, δ does not successfully link β_{ARCIV} to β_{RCIV}, suggesting that the hotdeck either fails to fully control for measurement error, or that it introduces additional error.

[1] Hotdeck is implemented using the STATA **hotdeck** command of Mander and Clayton (1999).

[2] We generated an additional imputed measure for children incorporating race categories, but there were large numbers of categories with no observations in them, and the estimation results using this imputation were unstable.

Table B2: Marginal Effects of Child BMI on Annual Medical Expenditures: Hotdeck Imputation

Estimator:	Non-IV	IV	RCIV	ARCIV	ARCIV$\times\hat{\delta}$
Column	(1)	(2)	(3)	(4)	(5)
Boys & Girls	17.45	92.66***	95.40***	95.61***	85.73
N= 27,002	(10.89)	(32.35)	(33.83)	(31.1)	(.899)
Boys	19.02	101.02**	100.30**	78.07*	90.82
N=13,718	(15.77)	(51.18)	(53.89)	(42.39)	(.906)
Girls	16.80	100.04***	105.27***	121.57***	94.81
N= 13,284	(14.31)	(36.31)	(38.12)	(36.8)	(.901)

Notes: Data: MEPS 2000–2010. *,**,*** indicate significance at 10%, 5%, 1% level respectively. BRR standard errors in parentheses. All entries are in 2010 dollars. Effects of moving form normal weight to overweight and obese are estimated by taking the difference of predicted expenditures at mean BMI within the respective weight class. (*) BMI are hotdeck imputed from NHANES are matched to observations in the MEPS using percentile rank of BMI. For column 5, $\hat{\delta}$ displayed in curly brackets.

ADAM I. BIENER

Lehigh University
College of Business and Economics
621 Taylor St.
Bethlehem, PA 18015

E-mail: aib210@lehigh.edu
Phone: (201) 819-8714
www.lehigh.edu/~aib210

EDUCATION

Lehigh University, Bethlehem, PA — (expected) September 2015
Ph.D. (Ph.D Candidate) Economics

Ramapo College of New Jersey, Mahwah, NJ — May 2010
B.A., Economics, Minors in Mathematics and Music

EMPLOYMENT

Research Economist — (Starting) August 2015
Center for Financing, Access and Cost Trends
Agency for Healthcare Research and Quality

Adjunct Professor — September 2013 - May 2015
Lehigh University

Teaching Assistant — September 2011 - May 2013
Lehigh University

RESEARCH PAPERS

"Savings in Medical Expenditures Associated with Reductions in Body Mass Index Among Adults with Obesity, by Diabetes Status"
J. Cawley, C. Meyerhoefer, A. Biener, M. Hammer, N. Wintfeld
Forthcoming in PharmacoEconomics,

REFEREED CONFERENCE PRESENTATIONS

AEA/ASSA Annual Meeting, Boston, MA. — January 2015
Contributed Papers in the Economics of Hospitals, Public Insurance, and Medical Expenditures
"Estimating the Medical Care Costs of Youth Obesity in the Presence of Proxy Reporting Error"
A. Biener, C. Meyerhoefer, J. Cawley

Obesity Week 2014 (The Obesity Society's Annual Scientific Meeting), Boston, MA. — November 2014
"Savings in Medical Expenditures Associated with Reductions in Body Mass Index Among Adults with Obesity, by Diabetes Status"
J. Cawley, C. Meyerhoefer, A. Biener, M. Hammer, N. Wintfeld

American Society of Health Economists — July 2014
5th Biennial Research Conference, Los Angeles, CA.
"The Medical Care Costs of Childhood Obesity" (poster)
A. Biener, C. Meyerhoefer, J. Cawley

Eastern Economic Association Annual Conference, New York, NY. — May 2013
"The Medical Care Costs of Childhood Obesity", A. Biener, C. Meyerhoefer, J. Cawley

OTHER CONFERENCE ACTIVITIES

Invited Attendee, NBER Summer Institute for Health Economics — July 2013

Session Chair, "The Economics of Obesity" Eastern Economic Association Annual Conference, New York, NY. May 2013

Discussant, Eastern Economic Association Annual Conference, New York, NY. May 2013

WORKS IN PROGRESS

"The Medical Care Costs of Youth Obesity"
A. Biener, C. Meyerhoefer, J. Cawley

"Bias due to Measurement Error in Regression Calibrated Estimates of the Costs of Adult Obesity"
A. Biener, C. Meyerhoefer, J. Cawley

MEDIA CITATIONS

BloombergBusiness, "Obesity is Hurting the U.S. Economy in Surprising Ways" March 5, 2015
The Incidental Economist, "What are the Health Care Costs of Obesity?" July 15, 2015

REFEREE EXPERIENCE

Bulletin of Economic Research April 2015
Health Economics July 2013, June 2014, December 2014, July 2015
International Journal of Obesity June 2014

TEACHING EXPERIENCE

Lehigh University
Instructor, Statistical Methods Summer 2014, Fall 2014, Spring 2015
Instructor, Intermediate Microeconomics Fall 2013, Spring 2014, Fall 2014
Research Assistant, Lehigh University / LVHN EHR Evaluation Summer 2012, Summer 2013
Teaching Assistant, Money and Banking and Financial Systems Spring 2013
Teaching Assistant, Principles of Economics Fall 2011, Spring 2012, Fall 2012

HONORS AND AWARDS

Visiting scholar to the Institute on Health Economics, Health Behaviors, and Disparities, Cornell University, Ithaca, NY. August 2014

Presidential Fellowship, Lehigh University Fall 2010 - Spring 2011

Sebastian J. Raciti endowed scholarship for undergraduates majoring in economics, Ramapo College 2009

SKILLS

STATA, SAS, LaTeX

REFERENCES

Chad D. Meyerhoefer

Lehigh University & NBER
621 Taylor St.
Bethlehem, PA 18015
Tel. (610) 758-3445
Fax (610) 758-4677
chm308@lehigh.edu

John H. Cawley

Cornell University & NBER
2312 MVR Hall
Ithaca, NY 14853
Tel: (607) 255-0952
johncawley@cornell.edu

James Dearden

Lehigh University
621 Taylor St.
Bethlehem, PA 18015
Tel. (610) 758-5129
Fax. (610) 758-4677
jad8@lehigh.edu