American Journal of Epidemiology
Vol. 146, No. 10
Printed in U.S.A.
Copyright © 1997 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

Effect of Measurement Error on Energy-Adjustment Models in Nutritional Epidemiology

Victor Kipnis,[1] Laurence S. Freedman,[1] Charles C. Brown,[1] Anne M. Hartman,[2] Arthur Schatzkin,[3] and Sholom Wacholder[4]

The use and interpretation of energy-adjustment regression models in nutritional epidemiology has been vigorously debated recently. There has been little discussion, however, regarding the effect of dietary measurement error on the performance of such models. Contrary to conventional assumptions invoked in the standard treatment of the effect of measurement error in regression analysis, reporting errors in dietary studies are usually biased, correlated with true nutrient intakes and with each other, heteroscedastic, and nonnormally distributed. Methods developed in this paper allow for this more complex error structure and are therefore more appropriate for dietary data. For practical illustration, these methods are applied to data from the Women's Health Trial Vanguard Study. The results demonstrate considerable shrinkage in the magnitude of the estimated main exposure effect in energy-adjustment models due to attenuation of the true effect and contamination from the effect of an adjusting covariate. In most cases, this shrinkage causes a sharply reduced statistical power of the corresponding significance test in comparison with measurement without error. These results emphasize the need to understand the measurement error properties of dietary instruments through validation/calibration studies and, where possible, to correct for the impact of measurement error when applying energy-adjustment models. *Am J Epidemiol* 1997;146:842–55.

energy intake; epidemiologic methods; measurement error; models, statistical; nutrient intake; regression analysis

There has been considerable interest and extensive discussion recently concerning application of energy-adjustment models in nutritional epidemiology. In these regression models relating nutrient intake to disease, the effect of a particular nutrient of interest is adjusted for total energy intake or intake of other nutrients. The interpretation and use of these models has been vigorously debated (1–7); however, there has been little discussion regarding the effect of dietary measurement error on their performance.

Dietary measurement is subject to substantial error that can have a profound impact on assessment of the effect of an exposure on disease (8–17). This has important implications for the design of nutritional epidemiologic studies (8, 13) and for their analysis and interpretation (15–17). In regression analysis, covariate measurement error leads to nonconsistent estimators of the regression coefficients (18, 19). It has long been appreciated in the applied literature that, in a univariate model, measurement error tends to bias the estimated regression coefficient toward zero (the attenuation effect). It may be less well known that this result is based on the "classical" assumption that error is independent of the true covariate. In general, when measurement error is correlated with the true covariate, the estimated regression coefficient can also be biased upward or have an opposite sign, as has been demonstrated in the statistical (20) and epidemiologic (16) literature. The effect of measurement error in multiple regression analysis can be much more complex (18, 21). Most of the published work studied this effect under at least some of the restrictive assumptions which constitute the classical measurement error model.

The application of energy-adjustment models is more complex than the standard treatment of the impact of measurement error, for several reasons. First,

Received for publication October 24, 1994, and in final form June 4, 1997.

Abbreviations: ARE, asymptotic relative efficiency; WHT, Women's Health Trial.

[1] Biometry Branch, Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, MD.
[2] Applied Research Branch, Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, MD.
[3] Cancer Prevention Study Branch, Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, MD.
[4] Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
Reprint requests to Dr. Victor Kipnis, Division of Cancer Prevention and Control, National Cancer Institute, Executive Plaza North, Suite 344, 6130 Executive Blvd., MSC 7354, Bethesda, MD 20892-7354.

there is usually a strong positive correlation between two covariates in the model (e.g., fat intake and total energy intake), both of which are measured with error. Second, errors are generally biased and may depend on true nutrient intakes (underreporting is more pronounced among those with higher nutrient intake). This dependence usually results in negative correlations between errors and true nutrient intakes. Third, the variance of errors is commonly large relative to the variance of true intakes, and it may be heteroscedastic. Fourth, there is usually a strong positive correlation between errors related to two covariates (under- and/or overreporting of different nutrient intakes tend to occur together). Fifth, both true and reported nutrient intakes generally have skewed nonnormal distributions.

There is a need, therefore, to consider this more complex error structure in order to address our main question: How does dietary measurement error affect the results of analyses using energy-adjustment models? Below we develop methods and provide formulas for the effect of generally structured measurement error on multiple regression analysis. Then, in order to illustrate the practical effect of dietary measurement error on energy-adjustment models, we apply these formulas to data from the Women's Health Trial (WHT) Vanguard Study. In particular, we demonstrate the impact of dietary error on both estimation of the main exposure effect and testing of its significance. Finally, we discuss the implications of measurement error for the design, analysis, and interpretation of nutritional epidemiologic studies.

MODELS AND METHODS

Energy-adjustment models

We consider below three alternative energy-adjustment models (5): the standard model,

$$E(y|F_t, T_t) = \beta_{0S} + \beta_{1S}F_t + \beta_{2S}T_t, \quad (1)$$

the residual model,

$$E(y|R_t, T_t) = \beta_{0R} + \beta_{1R}R_t + \beta_{2R}T_t, \quad (2)$$

and the energy partition model,

$$E(y|F_t, N_t) = \beta_{0P} + \beta_{1P}F_t + \beta_{2P}N_t. \quad (3)$$

In expressions 1–3, y denotes the disease outcome, F_t is the true long-term usual intake of the macronutrient of interest (e.g., fat), N_t is the true intake of "other" macronutrients (e.g., nonfat), and T_t is the true total energy intake. The subscript t is used to denote true as opposed to reported intake. We assume that the variables F_t, N_t, and T_t are measured in kilocalories per day. The variable R_t in model 2 is the true "energy-adjusted intake" of nutrient F_t—that is, the residual from the linear least squares approximation of F_t by T_t:

$$R_t = F_t - \alpha_0 - \alpha_1 T_t, \quad (4)$$

where

$$\alpha_0 = E(F_t) - \alpha_1 E(T_t),\ \alpha_1 = \frac{\sigma_{F_t}}{\sigma_{T_t}}\rho_{F_tT_t}.$$

Throughout this paper, we use the symbols σ_u and ρ_{uv} to denote the standard deviation of a random variable u and the correlation coefficient of random variables u and v, respectively.

For continuous variables, models 1–3 are mathematically equivalent and could be viewed as different reparameterizations of the same model (2, 3, 5). They lead to the same likelihood, and the regression coefficients of one model can be expressed as linear combinations of the coefficients of any other model. In this paper, we concentrate on estimating and testing the substitution and addition effects of the main exposure variable F_t (5). The former is the effect of substituting 1 kcal of nutrient F_t for 1 kcal of other macronutrients (β_{1S} in the standard model or β_{1R} in the residual model); the latter is the effect of adding 1 kcal of nutrient F_t to the diet without changing the intake of other macronutrients (β_{1P} in the partition model).

Below, we assume for simplicity that the three energy-adjustment models represent linear regressions of a continuous disease variable y on the corresponding covariates. However, as is discussed in more detail in the last section of this paper, our results apply, at least qualitatively, to logistic regression when y is a dichotomous variable.

True regression model

In the rest of this section, we consider energy-adjustment models 1–3 in their general form of the multiple linear regression

$$y = E(y|x_1, x_2) + \epsilon = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon. \quad (5)$$

We assume that x_1 is the main exposure variable, x_2 is an adjusting covariate (e.g., in the standard model $x_1 = F_t$ and $x_2 = T_t$), and ϵ is a disturbance term independent of x_1 and x_2.

Measurement error model

Because of measurement error, instead of the true covariates x_1 and x_2 we observe their respective surrogates z_1 and z_2. We assume that error is nondifferential with regard to the outcome variable y; i.e., z_1 and z_2 contribute no information about y beyond what

Material may be protected by copyright law (Title 17, U.S. Code)

is available in x_1 and x_2. Most of the literature on the effect of measurement error in regression analysis utilizes an additive error model,

$$z_i = x_i + e_i,\ i = 1, 2, \tag{6}$$

with at least some of the following conventional assumptions:

1. Errors are (conditionally) unbiased, i.e., $E(e_i|x_1, x_2) = 0$, $i = 1, 2$.
2. Errors do not correlate with true covariates, i.e., $\rho_{x_i e_j} = 0$, $i, j = 1, 2$.
3. Errors do not correlate with each other, i.e., $\rho_{e_1 e_2} = 0$.
4. Errors are homoscedastic, i.e., $\mathrm{Var}(e_i|x_1, x_2) = \sigma^2_{e_i}$, $i = 1, 2$.
5. True and observed covariates have a joint multivariate normal distribution.

Additive representation 6, together with assumptions 2–4, is usually called the classical measurement error model. The standard assessment of the effect of measurement error is based on the classical error model, sometimes with one or both of the remaining assumptions (11–15, 17, 22). Cochran (20) relaxes restriction 2, allowing error e_i to be correlated with the true corresponding covariate x_i, $i = 1, 2$.

Rosner et al. (23), among others, consider a more general measurement error model. Instead of making explicit assumptions about error structure, they assume that the regression of each of the true variables, x_i, on the observed covariates, z_1 and z_2, is linear, i.e.,

$$x_i = E(x_i|z_1, z_2) + \xi_i = \phi_{0i} + \phi_{1i}z_1 + \phi_{2i}z_2 + \xi_i,\quad i = 1, 2, \tag{7}$$

where ξ_i is independent of z_1 and z_2. Following Carroll et al. (21), we call expression 7 the linear regression calibration model. Although this model relaxes classical constraints 1–3, its underlying assumption about the linearity and homoscedasticity of regression 7 remains rather restrictive. In most practical cases, this assumption is equivalent to the requirement that the true and observed covariates be jointly normally distributed (20).

Popular methods of measuring dietary intake, such as food frequency questionnaires, commonly involve errors that are biased, correlated with true values and with each other, heteroscedastic, and nonnormally distributed (22, 24–26). As a result, none of the conventional assumptions 1–5 usually holds, and both the classical and the linear regression calibration models do not adequately describe the more complex structure of dietary measurement error. Below we consider the general case of nondifferential measurement error, relaxing all of the conventional assumptions.

Naive regression model

Unlike the true model 5, the regression of the response variable y on the observed covariates z_1 and z_2 will, in general, be nonlinear and/or heteroscedastic. Ignoring the impact of measurement error on model specification leads to the so called "naive" model,

$$y = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \delta, \tag{8}$$

representing the best (in the mean squared error sense) linear approximation of the regression $E(y|z_1, z_2)$. Following Cramer (27), we call model 8 the mean square linear regression of y on z_1 and z_2. The least squares theory for the mean square linear regression, which turns out to be somewhat different from the familiar linear regression case, is outlined in Appendix 1.

Bias of the estimated main exposure effect

As is shown in Appendix 1, the estimated main exposure effect $\hat{\gamma}_1$ is asymptotically (asy) normally distributed with the asymptotic mean

$$\mathrm{asy}E(\hat{\gamma}_1) \equiv \gamma_1 = A\beta_1 + C\beta_2, \tag{9}$$

where

$$A = \frac{g\sigma_{x_1 \cdot z_2}}{\sigma_{z_1 \cdot z_2}}\rho_{x_1 z_1 \cdot z_2} = \frac{\sigma_{x_1}}{\sigma_{z_1}}\left(\frac{\rho_{x_1 z_1} - \rho_{x_1 z_2}\rho_{z_1 z_2}}{1 - \rho^2_{z_1 z_2}}\right) \tag{10}$$

and

$$C = \frac{\sigma_{x_2 \cdot z_2}}{\sigma_{z_1 \cdot z_2}}\rho_{x_2 z_1 \cdot z_2} = \frac{\sigma_{x_2}}{\sigma_{z_1}}\left(\frac{\rho_{x_2 z_1} - \rho_{x_2 z_2}\rho_{z_1 z_2}}{1 - \rho^2_{z_1 z_2}}\right). \tag{11}$$

Material may be protected by copyright law (Title 17 U.S. Code)

Here, for random variables u, v, and w, $\sigma_{u \cdot w}$ and $\rho_{uv \cdot w}$ denote the standard deviation of u and the (partial) correlation coefficient of u and v, respectively, adjusted for w.

As follows from formula 9, $\hat{\gamma}_1$ is a biased estimator of the true main exposure effect β_1. In contrast to the univariate regression case, the bias is not fully multiplicative but also contains an additive component. The multiplicative component of the bias is due to factor A, which we call the attenuation coefficient, and its additive component is due to factor C, which we call the contamination coefficient.

As is shown in Appendix 1, for the classical measurement error model, $0 < A < 1$, so that factor A always attenuates the contribution of the true coefficient β_1, just as in the univariate case. An attenuation factor close to 1 indicates minimum attenuation, whereas a factor A close to 0 leads to maximum attenuation. When a less restrictive error structure is allowed, it is also possible to have the contribution of β_1 reversed ($A < 0$) and/or biased upward or deattenuated ($|A| > 1$), but we will retain the name of factor A as an attenuation coefficient.

The contamination factor C determines how much the true coefficient for the adjusting covariate, β_2, contributes to the bias of the estimated main exposure effect. Note that a contamination coefficient close to zero means minimum contamination, and that contamination increases proportionally to $|C|$.

To better understand the combined impact of the attenuation and contamination factors on the estimated main exposure effect, we rewrite expression 9 as

$$\text{asy}E(\hat{\gamma}_1) = \begin{cases} C\beta_2, & \beta_1 = 0, \\ (A + Cr_\beta)\beta_1 = D\beta_1,\ r_\beta = \beta_2/\beta_1, & \beta_1 \neq 0, \end{cases} \tag{12}$$

where factor D may be called the distortion coefficient. For a nonzero main exposure effect, overall distortion may vary from attenuation of the true effect when $-A/C \leq r_\beta \leq (1 - A)/C$ to its deattenuation when $r_\beta > (1 - A)/C$, and to its reversal when $r_\beta < -A/C$.

Test of significance of the main exposure effect

Ignoring measurement error leads to the naive test of significance of the main exposure effect which is based on model 8 fitted to the reported data. As is shown in Appendix 2, the level of the naive test against the true null hypothesis $\beta_1 = 0$ is generally greater than the nominal level; therefore, the naive test does not control the type I error rate. However, since the naive and true tests are based on different data (with and without measurement error, respectively), this fact alone does not uniquely define the relation between the two tests regarding their statistical power.

A convenient comparative power characteristic is the asymptotic relative efficiency (ARE) of the naive test with respect to the true one. The interpretation of the ARE is as follows. If n_r is the number of observations required by the naive test to obtain the same power against the same alternative as the true test based on n_t observations, then the ARE is asymptotically equal to the ratio $n_t : n_r$. Thus, to achieve the same power, the sample size n_t should be divided by the ARE. Retaining the same number of observations leads to a relative reduction in power of the naive test when ARE < 1, or an increase in power when ARE > 1.

As is shown in Appendix 2, the ARE is a monotonically decreasing function of the squared multiple correlation coefficient R^2 for the true model 5. When R^2 tends toward 1 (an almost functional relation between the response and covariates), the ARE approaches its minimum value of zero. When R^2 tends toward 0 (a very weak association), the ARE approaches its maximum value, given by formula A15.

Relation to earlier work

As is shown in Appendix 1, for the classical measurement error model, our expression for the naive regression coefficients reduces to the familiar results reported by other authors (12, 14, 17, 20, 22). Moreover, despite our loose assumptions, the estimated main exposure effect $\hat{\gamma}_1$ has asymptotically the same mean, given by formulas 9–11, as would be obtained under more restrictive measurement error model 7. Therefore, we have generalized the previously published results to the broad-spectrum situation, in which measurement errors may be biased, correlated with true covariates and with each other, heteroscedastic, and nonnormally distributed. It should be noted, however, that expression A6 for the asymptotic variance of $\hat{\gamma}_1$ in the general case differs from formula A7 for the linear regression calibration model 7. The latter defines a smaller value and therefore will account for only part of the variation of $\hat{\gamma}_1$ if the regression of the true covariates on observed covariates is, in fact, nonlinear and/or heteroscedastic.

EXAMPLE

Data set

We illustrate the developed methods with data from the WHT Vanguard Study (28). This trial was carried

Material may be protected by copyright law (Title 17, U.S. Code)

out from 1985 through 1988 in three US cities (Cincinnati, Ohio; Houston, Texas; and Seattle, Washington). We use data on 86 women aged 45–69 years who were assigned to the nonintervention group and completed food frequency questionnaires and food records 6, 12, and 24 months from entry into the study. The average nutrient intake from 12 days of food records (three 4-day periods over 18 months) is used as the "true" usual intake. The nutrient intake from the Block food frequency questionnaire (29) at the end of the reporting period is taken as the "reported" intake. We consider dietary fat intake as the main exposure variable of interest.

Using the subscripts r and t to denote reported intake and true intake, respectively, the measurement error is defined as the difference between these intakes. For example, the error in reporting of fat intake is given by $e_F = F_r - F_t$. Table 1 shows descriptive statistics for the true and reported fat, nonfat, and total energy intakes and the corresponding measurement errors. The correlations shown in tables 2–4 demonstrate that the classical assumptions do not hold. There are strong positive correlations among reporting errors for fat energy (e_F), nonfat energy (e_N), and total energy (e_T) intakes, indicating that under- and/or overreporting of fat, nonfat, and total energy tend to occur together. There is a moderate negative correlation between the nonfat reporting error (e_N) and the true nonfat intake (N_t) and between the total energy reporting error (e_T) and the true total energy intake (T_t), respectively. There are smaller negative correlations between the fat reporting error (e_F) and each of the true intakes of fat, nonfat, and total energy. These negative correlations indicate that the women in the study tended to underreport high values for nutrient intakes and/or overreport low values, particularly for nonfat and total energy intakes. Switching to the residual method does not improve the situation. Table 4 shows a significant negative correlation between each of the errors e_R and e_T and the true residual fat intake R_t.

In addition, the joint distribution of the true and reported intakes is nonnormal ($p < 0.0001$ using the Shapiro-Wilk test), which makes the linear regression calibration model for measurement error rather implausible, at least without appropriate transformations of nutrient intake data to make their distribution more symmetric. Thus, the more complex structure of dietary measurement error requires the use of the general theory described in the previous section.

TABLE 1. Mean values for true and reported intakes of macronutrients and corresponding reporting errors: The Women's Health Trial Vanguard Study, 1985–1986

Variable	Notation	Mean
True fat	F_t	600.7 (175.3)*
True residual fat†	R_t	0.0 (101.4)
True nonfat	N_t	991.9 (201.1)
True total energy	T_t	1,592.6 (316.6)
Reported fat	F_r	588.8 (301.1)
Reported residual fat‡	R_r	0.0 (136.6)
Reported nonfat	N_r	952.4 (299.4)
Reported total energy	T_r	1,541.2 (534.8)
Fat error	e_F	−11.9 (262.5)
Residual fat error	e_R	0.0 (117.4)
Nonfat error	e_N	−39.5 (274.6)
Total energy error	e_T	−57.3 (482.2)

* Numbers in parentheses, standard deviation.
† $R_t = F_t - \alpha_0 - \alpha_1 T_t$, $\alpha_0 = -119.2$, $\alpha_1 = 0.452$.
‡ $R_r = F_r - \alpha_0' - \alpha_1' T_r$, $\alpha_0' = -184.9$, $\alpha_1' = 0.502$.

TABLE 2. Correlations among true and reported covariates and corresponding reporting errors in the partition model: The Women's Health Trial Vanguard Study, 1985–1986

	F_t	N_t	F_r	N_r	e_F	e_N
F_t	1.0	0.413	0.497	0.213	−0.097	−0.071
N_t		1.0	0.200	0.454	−0.047	−0.237
F_r			1.0	0.586	0.815	0.724
N_r				1.0	0.531	0.721
e_F					1.0	0.613
e_N						1.0

TABLE 3. Correlations among true and reported covariates and corresponding reporting errors in the standard model: The Women's Health Trial Vanguard Study, 1985–1986

	F_t	T_t	F_r	T_r	e_F	e_T
F_t	1.0	0.816	0.497	0.399	−0.097	−0.093
T_t		1.0	0.402	0.454	−0.083	−0.153
F_r			1.0	0.891	0.815	0.724
T_r				1.0	0.756	0.811
e_F					1.0	0.893
e_T						1.0

TABLE 4. Correlations among true and reported covariates and corresponding reporting errors in the residual model: The Women's Health Trial Vanguard Study, 1985–1986

	R_t	T_t	R_r	T_r	e_R	e_T
R_t	1.0	0.0	0.547	0.050	−0.227	−0.227
T_t		1.0	−0.005	0.454	−0.006	−0.153
R_r			1.0	0.0	0.691	0.003
T_r				1.0	−0.043	0.811
e_R					1.0	−0.044
e_T						1.0

Scenarios

As follows from expressions 9 and A15, both the estimated main exposure effect and the maximum ARE of the corresponding significance test depend on the true regression coefficients. Because the WHT study was terminated in its pilot phase, no data on

Material may be protected by copyright law (Title 17 U.S. Code)

disease variables exist, so the true regression parameters are unknown. For this reason, to understand the effect of measurement error on estimating and significance testing, we have constructed five hypothetical sets of true regression coefficients for models 1–3. Table 5 displays these five scenarios.

Consider for a moment the relative risk regression model (14)

$$\log \mathrm{RR}(F_t, N_t) = \beta_0 + \beta_{1P}F_t + \beta_{2P}N_t, \qquad (13)$$

where RR denotes the disease relative risk. In model 13, $\exp(\beta_{1P})$ represents the proportional change in risk per unit of change in F_t while controlling for N_t, with the analogous interpretation for $\exp(\beta_{2P})$. Let ΔF_t and ΔN_t represent the change in intakes of fat and nonfat, respectively, from the middle of the first quartile (i.e., the 12.5 percentile) to the middle of the fourth quartile (the 87.5 percentile) of their corresponding distributions. For the WHT data, $\Delta F_t = 399.6$ kcal and $\Delta N_t = 469.7$ kcal. The relative risks for the corresponding addition effects of fat and nonfat are then equal to $\exp(\Delta F_t\beta_{1P})$ and $\exp(\Delta N_t\beta_{2P})$, respectively. Each scenario is based on the specified relative risks, $\mathrm{RR}(\Delta F_t)$ and $\mathrm{RR}(\Delta N_t)$, for these two addition effects, as shown in columns 2 and 3 of table 5. The regression slopes in model 13 are then uniquely defined as

$$\beta_{1P} = \log \mathrm{RR}(\Delta F_t)/\Delta F_t,\ \beta_{2P} = \log \mathrm{RR}(\Delta N_t)/\Delta N_t.$$

These coefficients, presented in columns 4 and 5 of table 5, were adopted for partition model 3. The corresponding regression coefficients for standard and residual models 1 and 2, displayed in columns 6–9 of table 5, were calculated according to the following formulas (5):

$$\beta_{1S} \equiv \beta_{1R} = \beta_{1P} - \beta_{2P},\ \beta_{2S} = \beta_{2P},$$

$$\beta_{2R} = \alpha_1\beta_{1P} + (1 - \alpha_1)\beta_{2P},\ \alpha_1 = \frac{\sigma_{F_t}}{\sigma_{T_t}}\rho_{F_tT_t}.$$

RESULTS

Attenuation and contamination factors

We have used formulas 10 and 11 to calculate the attenuation and contamination factors for the estimated fat effect in the three energy-adjustment models (models 1–3) using the WHT data. The necessary second moments of the joint distribution of the true and reported intakes are displayed in tables 1–4. As is shown in table 6, the attenuation and contamination factors for the three energy-adjustment models are quite similar. The attenuation factor is approximately one third to two fifths, and the contamination factor is −0.068 to −0.011.

Effect of measurement error on estimated fat coefficient

For the five scenarios described in the previous section, table 7 contains the true values of the addition (β_{1P}) and substitution ($\beta_{1S} = \beta_{1R}$) fat effects and the asymptotic mean values of their estimates based on the reported data. We also show in this table the values of the distortion factor D, calculated according to formula 12, and the maximum ARE of the naive significance test, calculated according to formula A15.

For all five scenarios, table 7 demonstrates a substantial downward bias in the estimated substitution fat effect (the relative bias is between −66 percent and −59 percent) and a dramatic reduction in power (the maximum ARE is between 0.21 and 0.31) for testing its significance. The situation with the addition fat effect for scenarios 1–3 is somewhat similar. The estimates based on the reported data are biased downward, with the relative bias between −73 percent and −61 percent, and a corresponding reduction in power for testing the significance of the effect. In scenario 4, the estimated addition effect has the least relative bias of −32 percent. The maximum ARE of the corresponding significance test is 1.06; i.e., the naive test

TABLE 5. Five different scenarios and corresponding values of the true regression coefficients (per 1,000 kcal) for the three energy-adjustment models: The Women's Health Trial Vanguard Study, 1985–1986

Scenario	Fat RR*,†	Nonfat RR‡	Partition model		Standard model		Residual model	
			β_{1P}	β_{2P}	β_{1S}	β_{2S}	β_{1R}	β_{2R}
1	3.00	0.33	2.75	−2.34	5.09	−2.34	5.09	−0.04
2	3.00	1.00	2.75	0.0	2.75	0.0	2.75	1.24
3	3.00	3.00	2.75	2.34	0.41	2.34	0.41	2.53
4	1.20	0.33	0.46	−2.34	2.80	−2.34	2.80	−1.08
5	1.20	3.00	0.46	2.34	−1.88	2.34	−1.88	1.49

* RR, relative risk.

† Relative risk due to the addition of fat intake from the 12.5 percentile to the 87.5 percentile of the fat distribution (399.6 kcal), while keeping nonfat intake constant.

‡ Relative risk due to the addition of nonfat intake from the 12.5 percentile to the 87.5 percentile of the nonfat distribution (469.7 kcal), while keeping fat intake constant.

Material may be protected by copyright law (Title 17, U.S. Code)

TABLE 6. Attenuation and contamination factors for the estimated fat coefficient in the three energy-adjustment models: The Women's Health Trial Vanguard Study, 1985–1986

Factor	Partition model	Standard model	Residual model
Attenuation	0.331	0.401	0.406
Contamination	−0.068	−0.011	−0.011

can have slightly greater power to find a significant effect than the test based on data without measurement error. For scenario 5, the mean estimated addition effect is close to zero and has the sign opposite that of the true value, with almost total loss of power of the corresponding significance test.

DISCUSSION

Although there has been much discussion about the use and meaning of different energy-adjustment models, little consideration has been given to the practical performance of such models in the presence of dietary measurement error. Several questions come to mind. First, how robust are the estimated substitution and addition effects against realistic levels of dietary measurement error? Does adjustment for total energy intake or for intake of other nutrients exacerbate the degree of attenuation usually found with a univariate model, or distort the estimate in some other way? Second, what impact does this distortion have on the statistical power to find a significant effect? Third, given the more complex structure of dietary measurement error, how much in error would the classical approach have been? We used the WHT data to explore these questions.

Robustness to measurement error

In our example, none of the scenarios yields estimated substitution and/or addition effects that appear to be robust to measurement error. Table 6 shows somewhat more favorable (leading to less distortion) attenuation and contamination coefficients for the standard and residual models than for the partition model. However, since the overall bias also depends on the ratio of the true regression coefficients (expression 12), the estimated substitution effect was not always less biased than the estimated addition effect in our scenarios.

In dietary studies, reported intake of a nutrient is usually more correlated with the true intake of the same nutrient than with a different nutrient, and these correlations are positive. Using general notations, this can be written as

$$\rho_{x_i z_i} > \rho_{x_i z_j} > 0,\ i, j = 1, 2;\ i \neq j. \qquad (14)$$

It then follows from formula 10 that the attenuation coefficient will commonly be positive. Although this is not assured for all data sets, in our example the attenuation coefficient is less than 1 and is much greater in magnitude than the contamination coefficient for all three energy-adjustment models. Hence, for a relatively small (in magnitude) ratio of the true regression coefficients, r_β, the attenuation coefficient dominates in determining the distortion factor and causes it to attenuate the true main exposure effect. This can be observed for all five scenarios for the estimated substitution effect and for scenarios 1–3 for the estimated addition effect (table 7).

However, in comparison with the univariate model without any adjusting covariate, the degree of overall attenuation in these cases is not exacerbated. The attenuation coefficient for the univariate model relating disease to crude (unadjusted) fat intake in the WHT data is equal to 0.29, somewhat smaller (and therefore causing more attenuation) than most values of the distortion coefficient obtained with the energy-adjustment models. This may seem surprising, since the addition of the second poorly measured covariate to the regression would intuitively be expected to worsen the problems of estimation. Indeed, we can show that if the measurement errors were "classical"—i.e., uncorrelated with the true variables and with each other—overall attenuation would get worse. A special aspect of the dietary data, however, is the strong positive correlation between errors in reported

TABLE 7. The true fat effect (β_1), its mean estimated value ($E(\hat{\gamma}_1)$), the distortion factor (D), and the maximum asymptotic relative efficiency (maxARE) of the naive significance test for five scenarios: The Women's Health Trial Vanguard Study, 1985–1986

Scenario	Addition effect (partition model)				Substitution effect (standard/residual model)			
	β_1	$E(\hat{\gamma}_1)$	D	max ARE	β_1	$E(\hat{\gamma}_1)$	D	max ARE
1	2.75	1.07	0.39	0.35	5.09	2.07	0.41	0.30
2	2.75	0.91	0.33	0.26	2.75	1.10	0.40	0.29
3	2.75	0.75	0.27	0.17	0.41	0.14	0.34	0.21
4	0.46	0.31	0.68	1.06	2.80	1.15	0.41	0.31
5	0.46	−0.01	−0.02	0.001	−1.88	−0.78	0.41	0.31

Material may be protected by copyright law (Title 17 U.S. Code)

nutrient intakes—e.g., 0.61 between errors in reported fat and nonfat, respectively, in the WHT data (table 3). This strong correlation is responsible for the greater attenuation coefficient in the energy-adjustment models than in the crude model, which, for relatively small absolute values of ratio r_β, leads to overall smaller attenuation of the estimated main exposure effect.

The contribution of contamination from an adjusting covariate to distortion of the main exposure effect becomes substantial when ratio r_β is comparable with, or greater than, A/C in magnitude. Because of inequality 14, the contamination factor will usually be substantially smaller in magnitude than the attenuation factor. Contamination, therefore, becomes noticeable only for relatively large absolute values of ratio r_β, i.e., when an adjusting covariate has a much greater effect on disease than the main exposure variable. Examples are shown in scenarios 4 and 5 for the estimated addition effect. In scenario 4, the two covariates have opposite effects on disease, and since C is negative, contamination reduces the degree of overall attenuation of the estimated fat coefficient. In scenario 5, both covariates affect disease in the same direction. As a result, contamination seriously further distorts the estimated fat coefficient, causing it to become negligible in magnitude and of a different direction than the true fat effect.

In general, we can show that the contamination factor is a monotonically decreasing function of the correlation between the reporting errors. For some data sets, this correlation may be somewhat smaller than in our example, causing C to be positive. More often, however, we would expect a rather strong positive correlation between the errors, leading to a negative contamination factor. When the correlation between errors approaches 1, the contamination factor tends to be closer in magnitude to the attenuation factor. As a result, even for moderate absolute values of ratio r_β, the estimated main exposure effect can have bias of arbitrary size and direction.

Statistical power

Shrinkage in the magnitude of the estimated regression coefficients, observed in all of our examples, is a problem not only for proper estimation of a nutrient-disease association but also for the power to detect a significant nutrient effect in an epidemiologic study. As follows from formula A15, the maximum ARE of the naive test is a product of two factors. The first factor is related to the ratio of the variances of the estimated main exposure effects using true and reported data, respectively. It is generally greater than 1 for dietary studies. In the WHT data, this factor is equal to 2.34 for the partition model and 1.81 for the standard or residual model. The second factor is the squared distortion coefficient, D^2. If D is relatively small in magnitude (causing rather substantial shrinkage of the true regression coefficient), the maximum ARE will be less than 1; i.e., the naive test will always be less powerful than the test based on data without measurement error. Except for testing of the addition effect in scenario 4, this is demonstrated in all of our examples, where rather severe shrinkage in magnitude of the true main exposure effect leads to approximately a threefold or greater reduction in the maximum ARE (table 7). Scenario 4 is different because, due to considerable contamination from the adjusting covariate, the distortion coefficient for the addition effect there is relatively close to 1. Thus, the extra power in this case is actually an artifact due to contamination from the relatively large protective addition effect assumed for nonfat. This artificial extra power of the naive significance test will be observed every time when the distortion coefficient is close to or greater than 1 in magnitude.

Comparison with classical measurement error

Our results on the impact of measurement error on energy-adjustment models are determined, to a substantial degree, by the specific structure of dietary data and, in particular, strong positive correlations between the true and reported nutrient intakes and between the reporting errors. Table 8 displays the attenuation and contamination coefficients obtained in the WHT data if we ignore these correlations and base our calculations on formulas A9 and A10 for the classical measurement error model. Table 9 demonstrates the differences among the true fat effect and its estimates under the general and classical assumptions for the three energy-adjustment models and five scenarios. Note that the standard and residual models no longer produce the same estimated substitution effect in the "classical" case. Using definition 4 of the energy-adjusted nutrient intake, we can show that the classical assumptions would not hold for the residual model, even if they were true for the standard model. Ignoring this fact and applying classical formulas to both models leads to different estimated substitution effects.

TABLE 8. Attenuation and contamination factors for the estimated fat coefficient in the three energy-adjustment models under classical measurement error assumptions: The Women's Health Trial Vanguard Study, 1985–1986

Factor	Partition model	Standard model	Residual model
Attenuation	0.295	0.263	0.427
Contamination	0.097	0.339	0.0

TABLE 9. The true fat effect (β_1) and its mean estimated values under general ($E(\hat{\gamma}_1)$) and classical ($E(\hat{\gamma}_{1C})$) measurement error assumptions for the three energy-adjustment models in five scenarios: The Women's Health Trial Vanguard Study, 1985–1986

Scenario	Partition model			Standard model			Residual model		
	β_1	$E(\hat{\gamma}_1)$	$E(\hat{\gamma}_{1C})$	β_1	$E(\hat{\gamma}_1)$	$E(\hat{\gamma}_{1C})$	β_1	$E(\hat{\gamma}_1)$	$E(\hat{\gamma}_{1C})$
1	2.75	1.07	0.58	5.09	2.07	0.55	5.09	2.07	2.80
2	2.75	0.91	0.81	2.75	1.10	0.72	2.75	1.10	1.52
3	2.75	0.75	1.04	0.41	0.14	0.90	0.41	0.14	0.23
4	0.46	0.31	−0.09	2.80	1.15	−0.06	2.80	1.15	1.54
5	0.46	−0.01	0.36	−1.88	−0.78	0.30	−1.88	−0.78	−1.04

It is clear from table 9 that the "classical" estimates behave very differently from their general counterparts based on more realistic assumptions. Sometimes they demonstrate more severe shrinkage of the true effect than actually takes place (scenarios 1 and 2 for the partition and standard models). In other cases, they mask real attenuation of the estimated effect (scenario 3 for the partition model or scenarios 1–5 for the residual model), show a false change of direction (scenario 4 for the partition model and scenarios 4 and 5 for the standard model), or exaggerate the true effect when in fact it is attenuated (scenario 3 for the standard model).

Limitations

One should note that our results are based on the WHT data set, with its attendant particularities and certain limitations. For instance, our data included information on only 86 volunteer women; the "true" nutrient intakes were assumed to be the averages of data from the three 4-day food records; and the reported intakes reflected one particular dietary instrument, the Block food frequency questionnaire. Nevertheless, the mean values, standard deviations, and correlations of the joint distribution of the true and reported macronutrient intakes (tables 1–4) appear to be rather typical of nutritional epidemiologic studies. They reflect the major features of dietary data, such as a substantial positive correlation between different nutrient intakes, strong positive correlations between the true and reported nutrient intakes and between the reporting errors, etc. As a result, our findings should be qualitatively rather similar to those obtained in other data sets.

The results presented in this paper are based on linear regression models, which are appropriate when the disease variable is continuous. In many applications in nutritional epidemiology, the disease variable is dichotomous (e.g., disease status), and logistic regression is then the method of choice. Published results on application of the regression calibration method (21, 23, 30), as well as some simulations of our own based on the presented approach, suggest that when measurement error is moderate and/or the main exposure effect is not too strong, logistic regression produces results qualitatively similar to those obtained with linear regression.

Conclusion

Dietary assessment is the foundation of a nutritional epidemiologic study. As was demonstrated here, dietary measurement error can have a profound impact on the results of a study. Contrary to a univariate regression under classical measurement error assumptions, there do not seem to be simple rules with which to make even qualitative statements about the direction and magnitude of the bias of the estimated effects in energy-adjustment models. Inconsistencies in the conclusions among several nutritional studies (such as those relating fat intake to breast cancer) may, of course, be explained in terms of the different populations sampled or different methodologies applied; but the controversial results could be also due, at least partly, to differences in the magnitude and pattern of measurement error associated with the dietary instrument used. Moreover, even consistent evidence about nutrient risk factors found in different studies could be misleading because of bias caused by errors of a similar structure (17).

It is therefore essential to understand the measurement error properties of the instruments that are being used. Calibration/validation studies of dietary instruments can be useful for estimating the attenuation and contamination factors for energy-adjustment models. In practice, when estimated regression coefficients are obtained from the analysis of an epidemiologic study, these factors can provide information for evaluation and comparison of different dietary instruments. In addition, the estimated attenuation and contamination factors can be used in an effort to correct the estimated effects, as described, for example, by Rosner et al. (23). Our theoretical results indicate that this type of bias correction is robust, at least in large-enough samples, to departures from assumptions about the linearity and homoscedasticity of the regression calibration model. Correction of standard errors and confidence

Material may be protected by copyright law (Title 17, U.S. Code)

intervals proposed by others (23, 30) should be modified, however, to accommodate expression A6 for the variance of the estimated regression coefficients in the more general case. Besides these efforts to understand and cope with dietary measurement error, there is, of course, a continuing need to improve the instruments of dietary assessment.

REFERENCES

1. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. Am J Epidemiol 1986;124:17–27.
2. Pike MC, Bernstein L, Peters RK. Re: "Total energy intake: implications for epidemiologic analyses." (Letter). Am J Epidemiol 1989;129:1312–13.
3. Howe GR. The first author replies. (Letter). Am J Epidemiol 1989;129:1314–15.
4. Pike MC, Peters RK, Bernstein L. Re: "Total energy intake: implication for epidemiologic analyses." (Letter). Am J Epidemiol 1993;137:811–12.
5. Kipnis V, Freedman LS, Brown CC, et al. Interpretation of energy adjustment models for nutritional epidemiology. Am J Epidemiol 1993;137:1376–80.
6. Wacholder S, Schatzkin A, Freedman LS, et al. Can energy adjustment separate the effects of energy from those of specific macronutrients? Am J Epidemiol 1994;140:848–55.
7. Mackerras D. Energy adjustment: the concepts underlying the debate. J Clin Epidemiol 1996;49:957–62.
8. Walker AM, Blettner M. Comparing imperfect measures of exposure. Am J Epidemiol 1985;121:783–90.
9. Freudenheim JL, Marshall JR. The problem of profound mismeasurement and the power of epidemiological studies of diet and cancer. Nutr Cancer 1988;11:243–50.
10. Greenland S. The effect of misclassification in the presence of covariates. Am J Epidemiol 1980;112:564–9.
11. Kupper LL. Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. Am J Epidemiol 1984;120:643–8.
12. Liu K. Measurement error and its impact on partial correlation and multiple linear regression analyses. Am J Epidemiol 1988; 127:864–74.
13. Freedman LS, Schatzkin A, Wax Y. The impact of dietary measurement error on planning sample size required in a cohort study. Am J Epidemiol 1990;132:1185–95.
14. Armstrong BG. The effect of measurement errors on relative risk regressions. Am J Epidemiol 1990;132:1176–84.
15. Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease relationships and methods of correction. Annu Rev Public Health 1993;14:69–93.
16. Wacholder S. When measurement errors correlate with truth: surprising effects of nondifferential misclassification. Epidemiology 1995;6:157–61.
17. Aickin M, Ritenbaugh C. Analysis of multivariate reliability structures and the induced bias in linear model estimation. Stat Med 1996;15:1647–61.
18. Fuller WA. Measurement error models. New York, NY: John Wiley and Sons, Inc, 1987.
19. Stefanski LA. The effect of measurement error on parameter estimation. Biometrika 1985;72:583–92.
20. Cochran WG. Errors of measurement in statistics. Technometrics 1968;10:637–66.
21. Carroll RJ, Ruppert D, Stefanski LA. Measurement error in nonlinear models. London, England: Chapman and Hall Ltd, 1995.
22. Armstrong BG, Whittemore AS, Howe GR. Analysis of case-control data with covariate measurement error: application to diet and colon cancer. Stat Med 1989;8:1151–63.
23. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. Am J Epidemiol 1990;132:734–45.
24. Hartman AM, Brown CC, Palmgren J, et al. Variability in nutrient and food intakes among older middle-aged men: implications for design of epidemiologic and validation studies using food recording. Am J Epidemiol 1990;132:999–1012.
25. Willett W. Nutritional epidemiology. New York, NY: Oxford University Press, 1990.
26. Carroll RJ, Stefanski LA. Approximate quasi-likelihood estimation in models with surrogate predictors. J Am Stat Assoc 1990;85:652–63.
27. Cramer H. Mathematical methods of statistics. Princeton, NJ: Princeton University Press, 1974.
28. Henderson MM, Kushi LH, Thompson DJ, et al. Feasibility of a randomized trial of a low-fat diet for the prevention of breast cancer: dietary compliance in the Women's Health Trial Vanguard Study. Prev Med 1990;19:115–33.
29. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 1986;124:453–69.
30. Kuha J. Corrections for exposure measurement error in logistic regression models with an application to nutritional data. Stat Med 1994;13:1135–48.

APPENDIX 1

Statistical Theory for the Multiple Stochastic Regression with Covariates Subject to Nondifferential Measurement Error

Consider the linear stochastic regression model

$$y = E(y|\boldsymbol{x}) + \epsilon = \beta_0 + \boldsymbol{\beta}'\boldsymbol{x} + \epsilon, \tag{A1}$$

where y is the response variable, $\boldsymbol{x} = (x_1 \ldots x_k)'$ is the $(k \times 1)$ vector of random covariates, $\boldsymbol{\beta} = (\beta_1 \ldots \beta_k)'$ is the $(k \times 1)$ vector of regression slopes, β_0 is the intercept, and ϵ is the random disturbance term independent of $\boldsymbol{x}$. Assume that the true covariates are not directly observable, and let $\boldsymbol{z} = (z_1 \ldots z_k)'$ be the $(k \times 1)$ vector of surrogate covariates containing measurement error. Assume that error is nondifferential with respect to the response variable, i.e.,

$$f(y|\boldsymbol{x}, \boldsymbol{z}) = f(y|\boldsymbol{x}), \tag{A2}$$

where f denotes the appropriate probability (density) function.

General Measurement Error

Unlike the true model, A1, the regression of the response variable y on the surrogate covariates z, $E(y|z)$, will, in general, be nonlinear and/or heteroscedastic. Ignoring the impact of measurement error and formally following specification A1 leads to the naive model

$$y = \gamma_0 + \boldsymbol{\gamma}'z + \delta, E(\delta) = 0, E[\delta z] = 0, \tag{A3}$$

representing the best (in the mean squared error sense) linear approximation of the regression $E(y|z)$. Model A3 is called the mean square linear regression (27), with parameters defined by

$$\boldsymbol{\gamma} = \mathrm{Var}^{-1}[z]\mathrm{Cov}[z, y], \gamma_0 = E(y) - \boldsymbol{\gamma}'E[z]. \tag{A4}$$

Substituting y from A1 into equation A4 and taking assumption A2 into account, we have

$$\boldsymbol{\gamma} = \boldsymbol{\Lambda\beta}, \boldsymbol{\Lambda} = \mathrm{Var}^{-1}[z]\mathrm{Cov}[z, \boldsymbol{x}], \tag{A5}$$

where $\boldsymbol{\Lambda}$ is the matrix of the regression slopes in the mean square multivariate linear regression of the true covariates on observed covariates:

$$\boldsymbol{x} = \boldsymbol{\lambda}_0 + \boldsymbol{\Lambda}'z + \boldsymbol{\nu}, E[\boldsymbol{\nu}] = 0, E[z\boldsymbol{\nu}'] = 0.$$

In the case of two covariates, corresponding to the energy-adjustment models considered in the paper, matrix $\boldsymbol{\Lambda}$ can be expressed as

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{x_1z_1\cdot z_2} & \lambda_{x_2z_1\cdot z_2} \\ \lambda_{x_1z_2\cdot z_1} & \lambda_{x_2z_2\cdot z_1} \end{pmatrix},$$

where

$$\lambda_{x_iz_j\cdot z_k} = \frac{\sigma_{x_i\cdot z_k}}{\sigma_{z_j\cdot z_k}}\rho_{x_iz_j\cdot z_k} = \frac{\sigma_{x_i}(\rho_{x_iz_j} - \rho_{x_iz_k}\rho_{z_jz_k})}{\sigma_{z_j}(1 - \rho^2_{z_jz_k})}$$

denotes the (partial) regression coefficient for z_j, adjusted for z_k, in the mean square linear regression of x_i on z_j and z_k, $i, j, k = 1, 2, j \neq k$.

Note that the parameters γ_0, $\boldsymbol{\gamma}$ of the mean square linear regression, A3, are defined by the same formulas, A4, as if the regression of y on z were linear, i.e., $E(y|z) = \gamma_0 + \boldsymbol{\gamma}'z$. However, the least squares estimates of the regression parameters in model A3 have different properties than in the familiar linear, homoscedastic regression. With n subjects in the sample, the least squares estimator $\hat{\boldsymbol{\gamma}}^* = (\hat{\gamma}_0\hat{\boldsymbol{\gamma}}')'$ satisfies the following (conditionally) unbiased estimating equation:

$$\sum_{i=1}^{n} z_i^*(y_i - \widehat{\boldsymbol{\gamma}^*}'z_i^*) = \mathbf{0},$$

where $z_i^* = (1\ z_{i1}\ \ldots\ z_{ik})'$, $i = 1, \ldots, n$; $\boldsymbol{\gamma}^* = (\gamma_0\ \boldsymbol{\gamma}')'$. As a result, under certain regularity conditions, $\sqrt{n}(\hat{\boldsymbol{\gamma}}^* - \boldsymbol{\gamma}^*)$ converges in distribution to $\mathbb{N}(\mathbf{0}; E^{-1}[z^*z^{*\prime}]E[z^*\delta^2z^{*\prime}]E^{-1}[z^*z^{*\prime}])$. In particular, the least squares slopes estimator $\hat{\boldsymbol{\gamma}}$ is asymptotically (asy) normally distributed with the mean

$$\mathrm{asy}E[\hat{\boldsymbol{\gamma}}] = \boldsymbol{\gamma} = \mathrm{Var}^{-1}[z]\mathrm{Cov}[z, y]$$

and the variance-covariance matrix

$$\mathrm{asyVar}[\hat{\boldsymbol{\gamma}}] = \frac{1}{n}\mathrm{Var}^{-1}[z]E[(z - E[z])\delta^2(z - E[z])']\mathrm{Var}^{-1}[z].$$

Expressing δ from equations A1 and A3 as

$$\delta = \boldsymbol{\beta}'(\boldsymbol{x} - E[\boldsymbol{x}]) + \epsilon - \boldsymbol{\gamma}'(z - E[z]),$$

Material may be protected by copyright law (Title 17 U.S. Code)

we may obtain

$$\text{asyVar}[\hat{\boldsymbol{\gamma}}] = \frac{\sigma_\epsilon^2}{n}\text{Var}^{-1}[z] + \frac{1}{n}\text{Var}^{-1}[z]W\,\text{Var}^{-1}[z], \tag{A6}$$

where

$$W = E[(z - E[z])\{\boldsymbol{\beta}'(\boldsymbol{x} - E[\boldsymbol{x}]) - \boldsymbol{\gamma}'(z - E[z])\}^2(z - E[z])'].$$

Linear Regression Calibration Measurement Error Model

If the regression of the true covariates on observed covariates is linear and homoscedastic, the regression of y on z is also linear and homoscedastic. Then, according to the standard theory, the mean of $\hat{\boldsymbol{\gamma}}$ is equal to $\boldsymbol{\gamma}$ given by the same expression as before,

$$E[\hat{\boldsymbol{\gamma}}] = \boldsymbol{\gamma} = \text{Var}^{-1}[z]\text{Cov}[z, y],$$

but its asymptotic variance-covariance matrix is given by the usual least squares formula,

$$\text{asyVar}[\hat{\boldsymbol{\gamma}}] = \frac{\sigma_\delta^2}{n}\text{Var}^{-1}[z], \tag{A7}$$

which is different from expression A6.

Classical Measurement Error Model

Under the classical assumptions, errors are additive, i.e.,

$$z = \boldsymbol{x} + \boldsymbol{e},$$

where $\boldsymbol{e} = (e_1 \ldots e_k)'$, and do not correlate with each other and with the true covariates,

$$\text{Cov}(x_i, e_j) = 0, \text{Cov}(e_i, e_j) = 0, i, j = 1, \ldots, k, i \neq j.$$

As a result, matrix $\boldsymbol{\Lambda}$ in expression A5 reduces to

$$\boldsymbol{\Lambda} = (\text{Var}[\boldsymbol{x}] + \text{Var}[\boldsymbol{e}])^{-1}\text{Var}[\boldsymbol{x}]. \tag{A8}$$

In the case of two covariates, denote by

$$R_i = \frac{\sigma_{x_i}^2}{\sigma_{z_i}^2} = \frac{\sigma_{x_i}^2}{\sigma_{x_i}^2 + \sigma_{e_i}^2}, i = 1, 2,$$

the coefficient of reliability of z_i as a measure of x_i, and by

$$R_{i\cdot j} = \frac{\sigma_{x_i\cdot z_j}^2}{\sigma_{z_i\cdot z_j}^2} = \frac{\sigma_{x_i}^2(\sigma_{x_j}^2 + \sigma_{e_j}^2) - \sigma_{x_i}^2\sigma_{x_j}^2\rho_{x_ix_j}^2}{(\sigma_{x_i}^2 + \sigma_{e_i}^2)(\sigma_{x_j}^2 + \sigma_{e_j}^2) - \sigma_{x_i}^2\sigma_{x_j}^2\rho_{x_iqx_j}^2}, i, j = 1, 2, i \neq j,$$

the (partial) coefficient of reliability of z_i adjusted for z_j. Then the elements of matrix A8 are

$$\lambda_{x_iz_j\cdot z_k} = \begin{cases} R_{i\cdot j}, & i = j, i \neq k; \\ \rho_{x_ix_j}\dfrac{\sigma_{x_i}}{\sigma_{x_j}}R_j(1 - R_{i\cdot j}), & i = k, i \neq j. \end{cases}$$

Thus, general formulas 10 and 11 for the attenuation and contamination coefficients reduce to

$$A = R_{1\cdot 2} \tag{A9}$$

and

$$C = \rho_{x_2x_1}\frac{\sigma_{x_2}}{\sigma_{x_1}}R_1(1 - R_{2\cdot 1}). \tag{A10}$$

Material may be protected by copyright law (Title 17, U.S. Code)

Since $0 < R_{1\cdot 2} < 1$, it follows from formula A9 that

$$0 < A < 1,$$

so in the classical case factor A is, indeed, the attenuation coefficient. As follows from formula A10, in the classical case the contamination coefficient has the sign of the correlation coefficient between the two true covariates, so $C > 0$ when $\rho_{x_1x_2} > 0$ and $C < 0$ when $\rho_{x_1x_2} < 0$.

APPENDIX 2

Asymptotic Relative Efficiency of the Naive Test of Significance of the Main Exposure Effect

Consider linear stochastic regression model 5 fitted to data without measurement error. The asymptotic maximum likelihood ratio test of the null hypothesis $H_0{:}\beta_1 = 0$ against the two-sided alternative $H_1{:}\beta_1 \neq 0$ requires rejecting H_0 if

$$T_{\widehat{\beta_1}} = \frac{\widehat{\beta_1}^2}{\text{asyVar}(\widehat{\beta_1})} > \chi^2_{1,1-\alpha},$$

where $\chi^2_{1,1-\alpha}$ denotes the upper alpha quantile of the central chi-square distribution with 1 degree of freedom. The naive test based on model 8 fitted to the reported data rejects H_0 if

$$T_{\widehat{\gamma_1}} = \frac{\widehat{\gamma_1}^2}{\text{asyVar}(\widehat{\gamma_1})} > \chi^2_{1,1-\alpha}.$$

For any fixed β_1, we have

$$T_{\widehat{\beta_1}} \sim \chi^2_1\left(\frac{\beta_1^2}{\text{asyVar}(\widehat{\beta_1})}\right) \tag{A11}$$

and

$$T_{\widehat{\gamma_1}} \sim \chi^2_1\left(\frac{\gamma_1^2}{\text{asyVar}(\widehat{\gamma_1})}\right), \tag{A12}$$

where $\chi^2_1(a)$ denotes the noncentral chi-square distribution with 1 degree of freedom and the noncentrality parameter a. Since, according to formula 12, $\beta_1 = 0$ implies $\gamma_1 = C\beta_2$, the noncentrality parameter of the distribution of $T_{\widehat{\gamma_1}}$ under H_0 is greater than 0 for $\beta_2 \neq 0$. Thus,

$$\Pr(T_{\widehat{\gamma_1}} > \chi^2_{1,1-\alpha} \mid H_0) > \alpha;$$

that is, the level of the naive test is generally greater than the nominal level α.

From the standard linear regression theory for model 5,

$$\text{asyVar}[\hat{\boldsymbol{\beta}}] = \frac{\sigma_\epsilon^2}{n}\text{Var}^{-1}[\boldsymbol{x}],$$

so

$$\text{asyVar}(\widehat{\beta_1}) = \frac{\sigma_\epsilon^2}{n\sigma_{x_1}^2(1-\rho_{x_1x_2}^2)}. \tag{A13}$$

From formula A6 for mean square linear regression 8,

$$\text{asyVar}(\widehat{\gamma_1}) = \frac{\sigma_\epsilon^2 + H}{n\sigma_{z_1}^2(1-\rho_{z_1z_2}^2)}, \tag{A14}$$

Material may be protected by copyright law (Title 17, U.S. Code)

where

$$H = \frac{E([\beta_1 \mathring{x}_1 + \beta_2 \mathring{x}_2 - \gamma_1 \mathring{z}_1 - \gamma_2 \mathring{z}_2]^2 [\sigma_{z_2} \mathring{z}_1 - \sigma_{z_1} \rho_{z_1 z_2} \mathring{z}_2]^2)}{\sigma_{z_1}^2 \sigma_{z_2}^2 (1 - \rho_{z_1 z_2}^2)},$$

$$\mathring{x}_i = x_i - E(x_i),\ \mathring{z}_i = z_i - E(z_i),\ i = 1, 2.$$

For the true test based on n_t observations, let $n_r = n_r(n_t)$ be the number of observations required by the naive test to achieve the same power against the same alternative $\beta_1 \neq 0$. As follows from substituting expressions A13 and A14 in formulas A11 and A12 and taking expression 12 into account, n_r satisfies the following condition:

$$\frac{n_t}{n_r} = \frac{\sigma_{z_1}^2 (1 - \rho_{z_1 z_2}^2)}{\sigma_{x_1}^2 (1 - \rho_{x_1 x_2}^2)} \left(\frac{D^2 \sigma_\epsilon^2}{\sigma_\epsilon^2 + H} \right).$$

The asymptotic relative efficiency (ARE) of the naive test compared with the true one is defined by

$$\mathrm{ARE}(T_{\widehat{\gamma_1}} : T_{\widehat{\beta_1}}) = \lim_{n_t \to \infty} \frac{n_t}{n_r(n_t)} = \frac{\sigma_{z_1}^2 (1 - \rho_{z_1 z_2}^2)}{\sigma_{x_1}^2 (1 - \rho_{x_1 x_2}^2)} \left(\frac{D^2}{1 + \frac{H}{\sigma_\epsilon^2}} \right).$$

Obviously, the ARE is a monotonically increasing function in σ_ϵ^2 or, equivalently, a monotonically decreasing function in R^2, the multiple squared correlation coefficient for the true model 5. Thus, the ARE has the following minimum and maximum values:

$$\mathrm{minARE} = \mathrm{ARE}(\sigma_\epsilon^2 \to 0 / R^2 \to 1) = 0$$

and

$$\mathrm{maxARE} = \mathrm{ARE}(\sigma_\epsilon^2 \to \infty / R^2 \to 0) = \frac{\sigma_{z_1}^2 (1 - \rho_{z_1 z_2}^2)}{\sigma_{x_1}^2 (1 - \rho_{x_1 x_2}^2)} D^2. \tag{A15}$$

Material may be protected by copyright law (Title 17, U.S. Code)