Appendix D
Estimating Impacts of Tiering on Nutrient Measures

The central purpose of the analysis was to estimate the effect of lower reimbursement for Tier 2 providers on the nutrient composition of the meals and snacks they offer. This analysis was carried out within strata defined by the meal (breakfast, lunch, morning snack, and afternoon snack) and by the age of the children to whom the meal is offered (1-2, 3-5, and 6-12). Within each stratum the unit of analysis was the provider, with the provider's value for any nutrient calculated as the amount of the nutrient offered at the specified meal to children of the specified age, averaged over all of the days on which the provider reported serving that meal to children of that age (typically 5 days).

The sample comprised providers who completed the menu survey: Tier 2 providers in 1999 and all providers in 1995. Regression models were estimated for all nutrients in all strata. Logistic regression was used for dichotomous nutrient measures (e.g., whether a provider offered lunches to 3-5-year-olds that supplied at least one-third of the RDA for food energy). Sample weights were incorporated using SAS-callable SUDAAN (PROC REGRESS and PROC RLOGIST).

Estimation of tiering effects is complicated by the fact that the amount of any nutrient that the provider offers—i.e., the dependent variable—is determined by what foods the provider places on the menu and by what portion sizes the provider offers. The foods on the menu were recorded independently by each provider. Portion sizes, however, were observed for only a subsample of providers and were therefore imputed on the basis of the characteristics of the provider, the meal, the provider's location, and other factors described above (Appendix C). In estimating the standard error of the tiering effect, it is necessary to take into account the variance associated with the portion size estimation process as well as the variance in the model of nutrient measures itself. This appendix describes the method used to estimate this combined variance.

The appendix contains four sections. The first describes the basic problem, the second uses two simplified examples to illustrate the main features of our estimation approach, and the final two sections describe the estimators that were actually used, their errors of estimate, and the way in which these errors were estimated.

Introduction

We set out to estimate models of the form:

(1) $$E\left(y_i \middle| x_i\right) = x_i'\beta$$

where

y_i = the amount of some specified nutrient included in the meals offered by the *i*th provider (e.g., the percent of RDA for food energy in lunches for children aged 3-5), and

x_i = a vector of provider characteristics, including a dummy for the observation year, a measure of the provider's household income relative to the Federal poverty guideline, and the percent of children in the provider's census block group in 1990 living in households with incomes at or below 185 percent of the poverty guideline.

The coefficient of interest is the coefficient for the observation year, which is interpreted as estimating the effect of the difference in reimbursement plus the effect of any general changes in CACFP providers' menus or portion sizes between 1995 and 1999. The other covariates represent two of the three criteria used in assigning providers to Tier 1 or Tier 2. (The third factor, the low-income status of the provider's elementary school attendance area, was not measured in available data.)

The value of y_i is computed based on the nutrient content of various foods that were served by providers. Nutrient content is measured relative to the nutrient being considered. It may be expressed in either absolute or relative terms (e.g., grams of fat or percent of RDA for calcium). Total nutrient content reflects the nutrients per unit and the portion sizes of all of the foods offered by a provider. We observe menus for several types of meals (breakfast, lunch, snacks) offered to children of various ages on each of several days of the week. Portion sizes and relative nutrients per unit of foods offered will differ depending on the age of the child and the type of meal. In addition, nutrient content is sometimes analyzed in terms of subsets of meals—for example, lunches or meals offered to children in a certain age range.

The equation for y_i is:

(2)

$$y_i = \sum_{t \in T_0} \sum_{j=1}^{m} \left(w_{tj} a_{itj} d_{itj} \right)$$

where

t = an index running over the total set of meals, which is defined by all possible combinations of a meal's type (breakfast, lunch, snack), the day of the week on which it was served, and the age of the child to whom it was served;

T_0 = the set of meals included in the measure;

m = the total number of foods;

w_{tj} = the amount of the nutrient that is provided by a unit of the *j*th food when served in the *t*th meal;

d_{itj} = a variable that indicates whether the *i*th provider included the *j*th food in its menu for the *t*th meal; and

a_{itj} = the amount of the *j*th food offered by the *i*th provider if the *j*th food is included in the menu for the *t*th meal.

Estimation of Equation (1) was complicated by the fact that we observed menus for the entire sample, but only obtained information on portion sizes, (and thus only observed y) for a subsample. We could have estimated β using only the subsample for which we had complete information. We increased the precision of our estimates by using a two-stage procedure. The first stage used the subsample to estimate food amounts (including the effects of tier on amounts). These estimates, combined with the observed menus, allowed us to create estimated y's for the entire sample. The second stage estimated overall effects for the entire sample based on the estimated y's.

This two-stage procedure reduces the errors of estimate by using the information on menus for the entire sample. At the same time, it does require some additional steps to estimate standard errors, described in this appendix. The overall effects estimated in the second stage reflect the observed differences in menus weighted by the estimated portion amounts. The second stage regression estimates the error associated with chance fluctuations in observed menus. The potential error associated with errors in estimating portion amounts has to be estimated separately and added to the second stage regression estimate.

The estimators involved are discussed in the third and last sections of this appendix. These involve a variety of nonlinear functions and are not always transparent. The next section uses two simpler specifications to illustrate the basic structure of the procedure.

Two Illustrative Examples

Two examples illustrate the procedure used to estimate β. The first illustrates the basic approach of sequential estimation, using the subsample to estimate amounts and then the entire sample to estimate β. The second illustrates the use of auxiliary variables in estimating amounts.

Example 1: A Log-Linear Specification For A Multiplicative Model

The definition of y_i in Equation (2) is complicated by the fact that it involves sums over a large number of possible menu items. We can, of course, rewrite Equation (2) as a simple product:

(3) $$y_i = \bar{a}_i m_i$$

where

$\bar{a}_i$ = the average portion size (measured in nutrient units) served by the ith provider, and

m_i = the number of menu items offered by the ith provider.

We did not use this specification, because we would expect $\bar{a}_i$ to depend on the composition of the menu: a sequential procedure that used the subsample to estimate $\bar{a}_i$ and the entire sample to estimate m_i would throw away the information on the menu composition that is available for the entire sample. Even so, the specification of Equation (3) does provide a straightforward illustration of the estimation process.

Say that we had adopted the specification of Equation (3) and, in addition, replaced the specification of Equation (1) with a multiplicative model:

(4)

$$\begin{aligned}
\left\{\ln\left(\bar{a}_i\right)\right\} &= X\beta_a + \varepsilon_a \\
\left\{\ln\left(m_i\right)\right\} &= X\beta_m + \varepsilon_m \\
\left\{\ln\left(y_i\right)\right\} &\equiv \left\{\ln\left(\bar{a}_i\right)\right\} + \left\{\ln\left(m_i\right)\right\} \\
&= X\left(\beta_a + \beta_m\right) + \left(\varepsilon_a + \varepsilon_m\right) \\
&= X\beta + \varepsilon
\end{aligned}$$

where, as throughout this appendix, $\{\cdot\}$ is used to indicate a column vector of the indicated elements, so that

$\{\ln(y_i)\}$ = the column vector whose ith element is $\ln(y_i)$.

We have information on m_i for the entire sample and information on $\bar{a}_i$ for only a subsample. Given the specification of Equation (4), we could, of course, estimate each component of β separately:

(5)

$$\begin{aligned}
\hat{\beta}_a &= \left(X_1'X_1\right)^{-1} X_1'\left\{\ln\left(a_i\right)\right\} \qquad \hat{\beta}_a \sim \left[\beta_a,\ \sigma_a^2\left(X_1'X_1\right)^{-1}\right] \\
\hat{\beta}_m &= \left(X'X\right)^{-1} X'\left\{\ln\left(m_i\right)\right\} \qquad \hat{\beta}_m \sim \left[\beta_m,\ \sigma_m^2\left(X'X\right)^{-1}\right]
\end{aligned}$$

where the subscripted X indicates the matrix for the subsample of observations used in the first stage estimation (those for which amounts are observed). Assuming that the two errors are independent, our estimate of β would then be distributed as follows:

(6)

$$\hat{\beta} = \hat{\beta}_a + \hat{\beta}_m \sim \left[\beta,\ \ \sigma_a^2\left(X_1'X_1\right)^{-1} + \sigma_m^2\left(X'X\right)^{-1}\right]$$

Alternatively, we could obtain the same estimate of β by using the subsample estimate of β_a from Equation (5) to estimate $\ln(y_i)$ for the entire sample, and then regressing this estimated $\ln(y_i)$ on x. Form :

(7)

$$\left\{\ln\left(\tilde{y}_i\right)\right\} = \left\{\ln(m_i)\right\} + X\hat{\beta}_a$$

and then regress $\{\ln(\tilde{y}_i)\}$ on x:

(8)

$$\hat{\beta} = \left(X'X\right)^{-1} X'\left\{\ln\left(\tilde{y}_i\right)\right\} = \hat{\beta}_m + \hat{\beta}_a$$

The two procedures yield the same estimates of β. However, the regression of Equation (8) underestimates the error of estimate. The estimated error term for the second stage regression of $\ln(\tilde{y}_i)$ on x_i only includes ε_m:

$$\hat{\varepsilon}_2 \;=\; \left[I \;-\; X\left(X'X\right)^{-1}X'\right]\left\{\ln\left(\tilde{y}_i\right)\right\} \;=\; \left[I \;-\; X\left(X'X\right)^{-1}X'\right]\varepsilon_m \tag{9}$$

Accordingly, the estimated variance from the second stage regression will be:

$$\hat{V}_2\left(\hat{\beta}\right) \;=\; \hat{\sigma}_m^2\left(X'X\right)^{-1} \tag{10}$$

whereas the actual error from Equation (6) is:

$$V\left(\hat{\beta}\right) \;=\; \sigma_a^2\left(X_1'X_1\right)^{-1} \;+\; \sigma_m^2\left(X'X\right)^{-1} \tag{11}$$

We can estimate the actual variance by combining the regression estimate with the estimated error variance from the first stage regression of $\ln(\bar{a}_i)$ in Equation (5):

$$\hat{V}\left(\hat{\beta}\right) \;=\; \hat{V}_2\left(\hat{\beta}\right) \;+\; \hat{\sigma}_a^2\left(X_1'X_1\right)^{-1} \;=\; \hat{V}_2\left(\hat{\beta}\right) \;+\; \hat{V}\left(\hat{\beta}_a\right) \tag{12}$$

The size of the correction may be indicated by considering the case in which the subsample with information on amounts is a random sample of the overall sample. In this case, we would expect that

$$\left(X_1'X_1\right) \;\approx\; \frac{n_1}{n}\left(X'X\right) \tag{13}$$

so that we might approximate the variance of Equation (12) by:

$$\hat{V}\left(\hat{\beta}\right) \;\approx\; \left[1 \;+\; \left(\frac{n}{n_1}\right)\frac{\hat{\sigma}_a^2}{\hat{\sigma}_m^2}\right]\left[\hat{V}_2\left(\hat{\beta}\right)\right] \;=\; \left[\hat{\sigma}_m^2 \;+\; \left(\frac{n}{n_1}\right)\hat{\sigma}_a^2\right]\left(X'X\right)^{-1} \tag{14}$$

Example 2: Adding Auxiliary Variables

As described in the next section, the actual estimation procedure used auxiliary variables in the prediction of amounts. Our illustrative specification is now:

$$\begin{aligned}\left\{\ln\left(m_i\right)\right\} \;&=\; X\beta_m \;+\; \varepsilon_m \\ \left\{\ln\left(a_i\right)\right\} \;&=\; \begin{pmatrix}X & Z\end{pmatrix}\begin{pmatrix}\gamma_x \\ \gamma_z\end{pmatrix} \;+\; \varepsilon_a\end{aligned} \tag{15}$$

The z's are endogenous. Thus, in the specification for $\ln(y_i)$:

(16) $$\{\ln(y_i)\} = X\beta + \varepsilon$$

β reflects the combination of the direct effects of x and its effect in determining z:

(17) $$\beta = \beta_m + \gamma_x + G\gamma_z \quad \text{where } E(Z|X) = XG$$

The intermediate variable $\ln(\tilde{y})$ is now:

(18) $$\{\ln(\tilde{y}_i)\} = \{\ln(m_i)\} + (X \quad Z)\begin{pmatrix} \hat{\gamma}_x \\ \hat{\gamma}_z \end{pmatrix}$$

or, substituting from Equation (15) and Equation (17),

(19) $$\{\ln(\tilde{y}_i)\} = X\beta + \tilde{\varepsilon}$$

where

(20) $$\tilde{\varepsilon} = (X \quad Z)\begin{pmatrix} \hat{\gamma}_x - \gamma_x \\ \hat{\gamma}_z - \gamma_z \end{pmatrix} + \varepsilon_m + (Z - XG)\gamma_z$$

The estimator for β from the regression of $\ln(\tilde{y})$ on x is:

(21) $$\begin{aligned} \hat{\beta} &= (X'X)^{-1}X'\{\ln(\tilde{y}_i)\} \\ \hat{\beta} &\sim \left[\beta,\ (X'X)^{-1}X'Var(\tilde{\varepsilon})X(X'X)^{-1}\right] \end{aligned}$$

We can characterize the variance of $\tilde{\varepsilon}$ in terms of the components from Equation (20)[1]:

---

[1] In general, if we take the expectation of some variable, η, over two sets of variables, a and b, then

$$E(\eta) = E_a\left[E_{b|a}(\eta)\right]$$

and

$$Var(\eta) = Var_a\left[E_{b|a}(\eta)\right] + E_a\left[Var_{b|a}(\eta)\right]$$

(22)
$$V\left(\hat{\beta}\right) = \left(X'X\right)^{-1} X'\left[V_1\left(\tilde{\varepsilon}\right) + V_2\left(\tilde{\varepsilon}\right)\right]X\left(X'X\right)^{-1}$$

where

(23)
$$V_1\left(\tilde{\varepsilon}\right) = Var_{\hat{\gamma}}\left[E_{\left(\varepsilon_m, z|\hat{\gamma}\right)}\left(\tilde{\varepsilon}\right)\right] = Var_{\hat{\gamma}}\left[\begin{pmatrix} X & XG \end{pmatrix}\begin{pmatrix} \hat{\gamma}_x - \gamma_x \\ \hat{\gamma}_z - \gamma_z \end{pmatrix}\right]$$
$$V_2\left(\tilde{\varepsilon}\right) = E_{\hat{\gamma}}\left[Var_{\left(\varepsilon_m, z|\hat{\gamma}\right)}\left(\tilde{\varepsilon}\right)\right] = E_{\hat{\gamma}}\left[Var_{\left(\varepsilon_m, z|\hat{\gamma}\right)}\left(\varepsilon_m + \left(Z - XG\right)\gamma_z\right)\right]$$

The variance associated with the second component is estimated by the second stage ln($\tilde{y}$) regression (for which $\hat{\gamma}$ is fixed)[2]:

(24)
$$\hat{V}_2\left(\hat{\beta}\right) = \hat{\sigma}_2^2\left(X'X\right)^{-1}$$

The variance associated with the first component is given by:

(25)
$$V_1\left(\hat{\beta}\right) = \left(X'X\right)^{-1} X'\left[Var_{\hat{\gamma}}\left[\begin{pmatrix} X & XG \end{pmatrix}\begin{pmatrix} \hat{\gamma}_x - \gamma_x \\ \hat{\gamma}_z - \gamma_z \end{pmatrix}\right]\right]X\left(X'X\right)^{-1}$$
$$= \left[I, \quad G\right] Var\left(\hat{\gamma}\right)\begin{bmatrix} I \\ G' \end{bmatrix}$$

---

2 The error term in Equation (1) is given by:

$$\varepsilon = \varepsilon_m + \varepsilon_a + \left(Z - XG\right)\gamma_z$$

If ε_a is assumed to be homoskedastic, then assuming that ε is homoskedastic implies that the remaining error vector is also homoskedastic – that is,

$$\varepsilon_m + \left(Z - XG\right)\gamma_z \sim \left[0, \ \sigma_{az}^2 I\right]$$

This assumption is not as remarkable as it may seem. Whether or not specification of a homoskedastic error for Equation (16) seems reasonable is unlikely to depend on an assumption that there are no omitted variables. If a homoskedastic error for the amount equation also seems reasonable, then homoskedasticity of $[\varepsilon_m + (Z\text{-}XG)\gamma_z]$ is also reasonable.

The subsample regression provides an estimate of Var ($\hat{y}$), and we could use either the full sample or the subsample value of $(X'X)^{-1}X'Z$ to estimate G. Var ($\hat{y}$) is:

$$V\left(\hat{\gamma}\right) = \sigma_a^2 \begin{bmatrix} X_1'X_1 & X_1'Z_1 \\ Z_1'X_1 & Z_1'Z_1 \end{bmatrix}^{-1} \quad \textbf{(26)}$$

If we use the subsample to estimate G, then the estimate for Equation (25) becomes:

$$\begin{aligned} \hat{V}_1 &= \left[I, \quad \left(X_1'X_1\right)^{-1} X_1'Z_1\right] Var\left(\hat{\gamma}\right) \begin{bmatrix} I \\ Z_1'X_1\left(X_1'X_1\right)^{-1} \end{bmatrix} \\ &= \left(X_1'X_1\right)^{-1}\left[X_1'X_1, \quad X_1'Z_1\right] Var\left(\hat{\gamma}\right) \begin{bmatrix} X_1'X_1 \\ Z_1'X_1 \end{bmatrix} \left(X_1'X_1\right)^{-1} \\ &= \hat{\sigma}_a^2\left(X_1'X_1\right)^{-1} \end{aligned} \quad \textbf{(27)}$$

Our estimate for the variance of the estimated β is like that of Equation (12):

$$\hat{V}\left(\hat{\beta}\right) = \hat{V}_2\left(\hat{\beta}\right) + \hat{\sigma}_a^2\left(X_1'X_1\right)^{-1} \quad \textbf{(28)}$$

Note, however, that unlike Equation (12), the second term in Equation (28) is not equal to the estimate of Var($\hat{y}_x$):

$$\hat{V}\left(\hat{\gamma}_x\right) = \hat{\sigma}_a^2\left(X_1'\left(I - Z_1\left(Z_1'Z_1\right)^{-1} Z_1\right) X_1\right)^{-1} \quad \textbf{(29)}$$

This reflects the fact that the effects of the x's includes their effects on the expected values of the z's for each provider. We either need to calculate $(X_1'X_1)^{-1}$ or use the approximation of Equation (13).

Results for the Actual Specification

The actual estimation procedure parallels that described in the previous section. We specified the a_{itj} as a multiplicative function of provider, meal, and food characteristics. Estimation was carried out separately for each of 10 groups of items. Thus the final specification for items in the kth group was:

$$a_{itj} = \exp\left(z'_{kitj}\ \gamma_k\right)\left(\eta_{k1i}\eta_{k2ij}\right) \text{ for j} \in J_k \quad \textbf{(30)}$$

where

a_{itj} = the amount of the *j*th food offered by the *i*th provider if the *j*th food is included in the menu for the *t*th meal;

z_k = the vector of variables used in the portion size equation for the *k*th food group;

z_{kitj} = the value of the z_k vector for the *j*th food served by the *i*th provider for the *t*th menu;

η_{k1}, η_{k2} = provider level and item level multiplicative errors, respectively; and

J_k = the set of foods included in the *k*th group.

The nature of the *z's* is not essential to what follows. However, it may still be useful to note that the z_k vector includes provider characteristics (including tier), items relating to specific foods, such as external values for both the standard and usual portion sizes for the item given the age of the child to whom it was served, and items related to the type of meal, number of foods offered in the meal, number of meals served in a day. Some of the variables in the z_k vector have the same value for every food and menu served by the *i*th provider; some have values that only depend on the food; some have values that vary with food and the meal index.

It will simplify notation somewhat if we define an intermediate variable that combines estimated amounts and per unit nutrient weights. Define

(31) $$r_{itj} \;=\; w_{tj}\;\exp\left(z'_{kitj}\,\gamma_k \right)\;\text{ for j } \in \; J_k$$

where

r_{ijt} = the expected nutrient yield obtained from the *j*th food if it is served by the *i*th provider as part of the *t*th menu with characteristics z_{kitj};

w_{tj} = the amount of the nutrient provided by a unit of the *j*th food when served in the *t*th meal; and

γ_k and z_{kitj} = as in Equation (30).

Note that the expectation represented by r_{ijt} is conditional on the food being served in the circumstances represented by the *z*'s. The estimated nutritional content of foods from the *k*th group served by the *i*th provider is given by:

(32) $$\tilde{y}_{ik} \;=\; \sum_{t\,\in\,T_0}\sum_{j\,\in\,J_k} \hat{r}_{itj} d_{itj}$$

where

$\tilde{y}_{ik}$ = the estimated nutritional content of foods from the *k*th group served by the *i*th provider;

T_0 = the set of meals included in the measure (as discussed in connection with Equation (2);

$\hat{r}_{ijt}$ = the estimated value of r_{ijt}, obtained from Equation (31), using the estimated values of γ_k; and

d_{itj} = a dummy variable indicating that the *j*th food was part of the *t*th meal offered by the *i*th provider.

Likewise, the estimate of the amount supplied by the *i*th provider from all foods is:

(33)
$$\tilde{y}_i = \sum_k \tilde{y}_{ik}$$

Following the specification of Equation (1), β is estimated from the weighted regression of $\tilde{y}$ on a set of provider characteristics, x:

(34)
$$\hat{\beta} = \left(X'PX\right)^{-1} X'P\tilde{y}$$

where

P = a diagonal matrix of sampling weights.

Recall that y_i is defined by

(35)
$$y_i = \sum_{t \in T_0} \sum_{j=1}^{m} \left(w_{tj} a_{itj} d_{itj} \right)$$

The expected value of y_i given x_i can be developed in terms of a sequence of conditional expectations:

(36)
$$\begin{aligned} E\left(y_i \middle| x_i\right) &= E_{\left(d,z|x\right)}\left[E_{\left(a|d,z,x\right)}\left(\sum_{t \in T_0} \sum_j w_{tj} a_{itj} d_{itj} \right) \right] \\ &= E_{\left(d,z|x\right)}\left[\sum_{t \in T_0} \sum_j r_{itj} d_{itj} \right] \end{aligned}$$

Given consistent estimates of γ_k, $\hat{r}_{ijt}$ will be a consistent estimate of r_{ijt}, and $\tilde{y}_i$ will be a consistent estimate of $\Sigma\Sigma(r_{ijt}d_{ijt})$—the expected value of y_i given x_i, the d_{itj}'s, and the z_{kitj}'s (the bracketed expression in the second line of Equation (36)). The regression of $\tilde{y}_i$ on x_i will accordingly provide consistent estimates of β, and the estimated error of estimate from that regression will reflect random fluctuations in the d_{ijt}'s and z_{kitj}'s around their expected values given x_i. The variance of estimate of β will consist of this variance plus the variance associated with the error of estimate of $\hat{r}_{ijt}$:

(37)
$$V\left(\hat{\beta}\right) = V_1\left(\hat{\beta}\right) + V_2\left(\hat{\beta}\right)$$

where

$V(\hat{\beta})$ = the variance of $\hat{\beta}$;

$V_1(\hat{\beta})$ = the variance associated with the error of estimate of the $\hat{\gamma}_k$'s; and

$V_2(\hat{\beta})$ = the variance given the value of the $\hat{\gamma}_k$'s, which is the variance estimated in the second stage regression.

The asymptotic error of estimate of β that is associated with the error of estimate of $\hat{r}_{ijt}$ is obtained from the asymptotic approximation to the $\tilde{y}_k$'s. The first order approximation to $\tilde{y}_{ik}$ is given by:

(38)
$$\begin{aligned}\tilde{y}_{ik} &\approx \left(\tilde{y}_{ik}\big|_{\hat{\gamma}=\gamma}\right) + \left(\frac{\partial \tilde{y}_{ik}}{\partial \hat{\gamma}_{ik}}\bigg|_{\hat{\gamma}=\gamma}\right)\left(\hat{\gamma}_k - \gamma_k\right) + R_{ik} \\ &= \sum_{t\in T_0}\sum_j r_{itj} d_{itj} + \sum_{t\in T_0}\sum_j r_{itj} z'_{kitj}\left(\hat{\gamma}_k - \gamma_k\right) + R_{ik}\end{aligned}$$

where R_{ik} is a remainder term. Because $\hat{r}_{ijt}$ is a differentiable function of $\hat{\gamma}_k$, and $\hat{\gamma}_k$ is a consistent estimator of γ_k, we can ignore the remainder term in deriving the asymptotic distribution[3]. Dropping the remainder term from Equation (38) and rewriting it in terms of the vector, $\tilde{y}_{k} = \{\tilde{y}_{ik}\}$ gives:

(39)
$$\tilde{y}_k \approx \left(\tilde{y}_k\big|_{\hat{\gamma}=\gamma}\right) + \left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\bigg|_{\hat{\gamma}=\gamma}\right)\left(\hat{\gamma}_k - \gamma_k\right)$$

The asymptotic variance component of the estimate of β associated with the error of estimate for the $\hat{\gamma}_k$'s is estimated by:

(40)
$$\hat{V}_1 = \left(X'PX\right)^{-1} X'P\left[\sum_k \left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right) Var\left(\hat{\gamma}_k\right)\left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right)'\right] PX\left(X'PX\right)^{-1}$$

where

$Var(\hat{\gamma}_k)$ = the variance-covariance matrix of the estimated γ_k's, and

---

[3] Because $\hat{r}_{ijt}$ is a differentiable function of $\hat{\gamma}_k$, the behavior of $\tilde{y}_{ik}$ is dominated by the first two terms of Equation (38) when $(\hat{\gamma}_k$-$\gamma_k)$ is small (that is, as $(\hat{\gamma}_k$-$\gamma_k)$ goes to zero, the ratio, $R_{ik}/(\hat{\gamma}_k$-$\gamma_k)$ also goes to zero). Because $\hat{\gamma}_k$ is a consistent estimator of γ_k, we know that for large enough samples, $\hat{\gamma}_k$ will be arbitrarily close to γ_k with probability arbitrarily close to one.

$\left(\dfrac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right)$ = an n by s_k matrix of partial derivatives, where

n = the number of providers (observations) in the second stage equation used to estimate β;

s_k = the number of variables in the portion-size equations for the kth food group; and

the (i,h)th term for the matrix of partial derivatives is given by:

(41)
$$\left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right)_{ih} = \sum_{t \in T_0} \sum_{j \in k} \hat{r}_{itj} d_{itj} z_{kitjh}$$

where

z_{kitjh} = the value for the jth food served by the ith provider in the tth menu of the hth variable used to predict portion size for foods in the kth group.

Further Results for the Actual Specification

The dependent variable, y, sometimes involved ratios such as food energy from fat as a proportion of total food energy offered. In other cases, the dependent variable involved a dichotomous outcome, such as did or did not provide adequate amounts of some nutrient, for which effects were estimated using logits. Each of these is briefly discussed below.

Ratios. A ratio variable would be estimated as:

(42)
$$\tilde{y}_i = \frac{\tilde{y}_{Ai}}{\tilde{y}_{Bi}}$$

where the numerator and denominator followed the form of Equation (32) and Equation (33), but with nutrient weights and food indicator variables appropriate to the nutrients included in A or B, respectively. Thus, for example,

(43)
$$\hat{r}_{Aitj} = w_{Atj} \exp\left(z'_{kitj} \hat{\gamma}_k \right) \text{ for j } \in J_k$$

and

(44)
$$\tilde{y}_{Aik} = \sum_{t \in T_0} \sum_{j \in J_k} \hat{r}_{Aitj} d_{Aitj}$$

Note that the same $\hat{\gamma}_k$'s are used for both $\tilde{y}_{Aik}$ and $\tilde{y}_{Bik}$. This is likely to reduce the effect of errors in estimating $\hat{\gamma}_k$. For regressions involving such ratio dependent variables, the first stage error term is given by Equation (40), but the partial derivatives reflect the fact that $\tilde{y}_i$ is a ratio of terms involving $\hat{\gamma}_k$:

(45)
$$\frac{\partial \tilde{y}_{ik}}{\partial \hat{\gamma}_k} = \frac{1}{\tilde{y}_{Bi}} \frac{\partial \tilde{y}_{A\,ik}}{\partial \hat{\gamma}_k} + \frac{\tilde{y}_{Ai}}{\left(\tilde{y}_{Bi}\right)^2} \frac{\partial \tilde{y}_{B\,ik}}{\partial \hat{\gamma}_k}$$

so that :

(46)
$$\left(\frac{\partial \tilde{y}_{ik}}{\partial \hat{\gamma}_k}\right)_{ih} = \sum_{t \in T_0} \sum_{j \in J_k} \left\{\left[\frac{\hat{r}_{Aitj} d_{Aitj}}{\tilde{y}_{Bi}} - \frac{\tilde{y}_{Ai}}{\left(\tilde{y}_{Bi}\right)^2} \left(\hat{r}_{Bitj} d_{Bitj}\right)\right] z_{kitjh}\right\}$$

where

(47)
$$\tilde{y}_{Ai} = \sum_k \tilde{y}_{Aik} \text{ and } \tilde{y}_{Bi} = \sum_k \tilde{y}_{Bik}$$

Dichotomous Variables. Dichotomous variables were defined by expressions such as:

(48)
$$\tilde{\delta}_i = \begin{cases} 1 \text{ if } \tilde{y}_i \geq y_c \\ 0 \text{ otherwise} \end{cases}$$

where

$\tilde{y}_i$ = as in Equation (33); and

y_c = some criterion value.

Weighted logits were then estimated based on the specification:

(49)
$$\text{Prob}\left(\delta_i = 1 \middle| x_i\right) = F\left(x_i' \beta\right)$$

where $F(\cdot)$ is the logistic distribution function, and

(50)
$$\delta_i = \tilde{\delta}_i\Big|_{\hat{\gamma}=\gamma} = \begin{cases} 1 \text{ if } \mathrm{E}\left(y_i \middle| d, z\right) \geq y_c \\ 0 \text{ otherwise} \end{cases}$$

The FOC for the weighted logits are:

(51) $$X'P\left(\tilde{\delta} - \hat{F}\right) = 0$$

where

$\tilde{\delta} = \left\{\tilde{\delta}_i\right\}$; and

$\hat{F} = \left\{F\left(x_i'\hat{\beta}\right)\right\}$.

Using the first order approximation for *F*, the asymptotic expansion for the estimated β is:

(52) $$\left(\hat{\beta} - \beta\right) \approx \left(X'D_{PF}X\right)^{-1} X'P\left(\tilde{\delta}_i - F\left(x_i'\beta\right)\right)$$

where

(53) $$D_{PF} = DIAG\left\{P_i F\left(x_i'\beta\right)\left(1 - F\left(x_i'\beta\right)\right)\right\}$$

We want to estimate $F(x_i'\beta)$—the probability that $\delta_i = 1$. We in fact estimate:

(54) $$\begin{aligned} \text{Prob}\left(\tilde{\delta}_i \geq 1\right) &= \text{Prob}\left(\tilde{y}_i \geq y_c\right) \\ &\approx \text{Prob}\left(E\left(y_i \middle| d, z\right) \geq \left[y_c - \sum_k \left(\left(\frac{\partial \tilde{y}_{ik}}{\partial \hat{\gamma}_k}\right)\left(\hat{\gamma}_k - \gamma_k\right)\right)\right]\right) \\ &\approx F\left(x_i'\beta\right) - \left[F\left(x_i'\beta\right)\left(1 - F\left(x_i'\beta\right)\right)\sum \left(\left(\frac{\partial \tilde{y}_{ik}}{\partial \hat{\gamma}_k}\right)\left(\hat{\gamma}_k - \gamma_k\right)\right)\right] \end{aligned}$$

The additional error term is estimated by:

(55) $$\hat{V}_1 = \sum_k \left\{\left(X'D_{PF}X\right)^{-1} X'D_{PF}\left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right) Var\left(\hat{\gamma}_k\right)\left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right) D_{PF} X \left(X'D_{PF}X\right)^{-1}\right\}$$

where

$\left(\frac{\partial \tilde{y}_k}{\partial \hat{\gamma}_k}\right)$ = as in Equation (41).