WP.28
ENGLISH ONLY

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Bonn, Germany, 25-27 September 2006)

Topic (v): New and emerging methods

REGRESSION IMPUTATION WITH LINEAR EQUALITY CONSTRAINTS ON THE VARIABLES

Supporting Paper

Prepared by Jeroen Pannekoek, Statistics Netherlands

I. INTRODUCTION

1. The occurrence of missing values is a common problem for virtually all surveys. Estimation procedures have to take these missing values into account. The way that missing values are taken care of depends on whether all values for a unit are missing (referred to as unit missing values or unit non-response) or only the values for some variables are missing (item missing values or item non-response). Unit missing values are usually treated by weighting the responding units to compensate for the missing units, whereas item missing values are usually handled by imputation: the missing values are filled in by predictions obtained from a model, algorithm or rule that takes relevant characteristics of the partially responding unit into account.

2. Imputations for continuous variables are often based on a standard linear regression model (Little and Rubin, 1987). Let y denote a variable with missing values that requires imputation, then the model can be written as: $y_i = \mathbf{x}_i\beta + \varepsilon_i$, with y_i the value of the variable to be imputed for unit *i*, $\mathbf{x}_i$ the value of a vector with predictor variables for unit *i*, β a vector with regression coefficients and ε_i a random disturbance with expectation zero. The coefficient vector β can be estimated by least squares using the units with both y and **x** observed. For records with y missing and **x** observed, the missing values can be filled in or "imputed" by the predictions from the model: $\hat{y}_i = \mathbf{x}_i\hat{\beta}$

3. A problem with this type of imputation method occurs when certain relations between variables must hold. This is especially important for business surveys where the variables must satisfy many linear constraints. For instance: profit must equal turnover minus expenses. Expenses are broken down in details such as employment costs, taxes paid, costs of purchases and so on, and these details of expenses must add op tot the total expenses. Employment costs may again be broken down in wages, costs of training, costs of temporary personnel etc. and these details must add to the total employment costs. These constraints, restricting the sum of a number of (detail) variables to be equal to another (total) variable are called balance edits. In many business surveys there are dozens of such balance edits. Regression imputation leads to completed records that do not, in general, satisfy such balance edits. Tempelman (2004) shows that when all variables that appear in the constraints and are in some record non-missing are used for predicting the missing values in that record, these imputations will satisfy the constraints. For business surveys with many balance edits this means that almost all non-missing variables must be included as predictors. The models that are actually used in practice almost always

include just a few predictors, often from registers or previous surveys and thus lead to inconsistencies at the record level and possibly also between published aggregates.

4. An approach to remedy this situation is to add an "adjustment" to each imputed value, such that the adjusted imputed values are consistent with the balance edits. Linear programming techniques are typically used to find such adjustments (De Waal, 2003). Also, multiplicative adjustments can be used and raking algorithms can be applied to calculate these adjustments (Statistics Canada, 2005). In this paper we will investigate another approach to this problem. The idea is to make the adjustments part of the regression model and to estimate the adjustments and the regression coefficients simultaneously using standard least squares techniques. It will be shown that this approach will lead to extended regression models with the constraints on the imputed values as additional predictors. That the predictions from this extended model are consistent with the constraints follows immediately from standard properties of least squares estimation.

5. In section II it is shown how the balance edits lead to constraints on the missing values that the imputations should satisfy. It is also shown in this section how cases can be determined in which the constraints allow for only one possible solution for the missing values and how this solution (a logical or deductive imputation) can be obtained. In section III it is shown how regression models can be extended such that the model predictions are consistent with the constraints. This approach to consistent prediction is contrasted with the approach that calculates inconsistent imputations first and uses adjustment afterwards to obtain consistent imputations. In section IV, two variants of the standard linear regression model are treated. Firstly, imputation models that are estimated by *weighted* least squares, with ratio imputation as a special case. Secondly, a regression model for a log-transformed dependent variable whereas, of course, the constraints are formulated in terms of the original scale.

II. CONSTRAINTS ON THE MISSING VALUES AND DEDUCTIVE IMPUTATION

A. Formulation of the constraints

6. Consider a survey with p variables, $y_1, \dots, y_p$. Suppose that these variables must satisfy q balance edits. These constraints can be formulated as

$$\mathbf{R}'\mathbf{y}_i = \mathbf{0}, \tag{1}$$

with $\mathbf{y}_i$ the vector containing the p variables for unit i, $\mathbf{R}$ a $p \times q$ matrix with columns defining the constraints and $\mathbf{R}'$ it's transpose. The columns of $\mathbf{R}$ consists of zero's for variables not in the constraint and 1's and -1's for variables in the constraint.

7. Illustration

Suppose that there are 5 variables: *employee costs* (y_1), *other costs* (y_2), *total costs* (y_3), *turnover* (y_4) *and profit* (y_5). We then have

$$\mathbf{R}' = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 \end{bmatrix},$$

defining the two constraints: *employee costs* + *other costs* - *total costs* = 0 and *total costs – turnover = profit.*

B. Deductive imputation

8. If there are missing values in $\mathbf{y}_i$ it is sometimes possible to deduce the values of the missing variables from the observed variables and the constraints. In the illustration above this is obviously the case if, for instance, y_3 and y_4 are missing. For large numbers of constraints, however, it is not easy to see which missing values can be "solved". The following method makes it possible to automatically determine these deductive or logical imputations. This should be done before applying any model based imputation.

9. For some record with missing values, let o denote the observed variables and m denote the missing variables. After a permutation of elements, the vector $\mathbf{y}_i$ can be partitioned in an observed part and a missing part: $\mathbf{y}_i = (\mathbf{y}'_{oi}, \mathbf{y}'_m)'$. If $\mathbf{R}$ is partitioned conform the partitioning of $\mathbf{y}$ we can write

$$\begin{bmatrix}\mathbf{R}'_o & \mathbf{R}'_m\end{bmatrix}\begin{bmatrix}\mathbf{y}_{oi} \\ \mathbf{y}_{mi}\end{bmatrix} = \mathbf{0}, \tag{2}$$

and so

$$\mathbf{R}'_m \mathbf{y}_{im} = -\mathbf{R}'_o \mathbf{y}_{oi} = \mathbf{b}_i, \text{ say.} \tag{3}$$

10. Since $\mathbf{b}_i$ is known, this last expression is a system of linear equations in the missing values $\mathbf{y}_{im}$. The solutions to this system can be characterized by the following equation (Rao, 1973, page 25)

$$\mathbf{y}_{mi} = \mathbf{R}'^{-}_m \mathbf{b}_i + (\mathbf{R}'^{-}_m \mathbf{R}'_m - \mathbf{I})\mathbf{z}_i, \tag{4}$$

with $\mathbf{R}'^{-}_m$ a generalized inverse of $\mathbf{R}'_m$ and $\mathbf{z}_i$ an arbitrary vector. Note that (4) generates all possible solutions for $\mathbf{y}_{im}$ satisfying $\mathbf{R}'_m \mathbf{y}_{mi} = \mathbf{b}_i$ for any specific choice of generalized inverse. If $\mathbf{R}'_m$ is square and of full rank such that its regular inverse exists, $\mathbf{y}_{mi} = \mathbf{R}'^{-1}_m \mathbf{b}_i$: all missing values can be imputed deductively. Otherwise there are infinitely many solutions for $\mathbf{y}_m$ satisfying $\mathbf{R}'_m \mathbf{y}_{mi} = \mathbf{b}_i$. However, there may still be some (but not all) elements of $\mathbf{y}_m$ that are uniquely determined by the constraints. These elements will have the same value in each of the possible solutions for $\mathbf{y}_{mi}$ or equivalently, for each value of $\mathbf{z}$. This implies that the corresponding rows of $\mathbf{R}'^{-}_m \mathbf{R}'_m - \mathbf{I}$ must consist of zero's only. This is easy to verify and the corresponding elements $\mathbf{y}_{mi}$ can then be imputed deductively by the corresponding elements of $\mathbf{R}'^{-}_m \mathbf{b} = -\mathbf{R}'^{-}_m \mathbf{R}'_o \mathbf{y}_{oi}$, which is easy tot calculate.

III. A REGRESSION MODEL INCORPORATING CONSTRAINTS

A Formulation of the model

11. For the missing values that cannot be imputed deductively, a model based procedure can be used. In this section a regression model will be described that meets the requirement that the model predictions for the missing values, to be used as imputations, satisfy the constraints.

12. As predictors for $\mathbf{y}_{mi}$, variables can be used from $\mathbf{y}_o$ but also from other sources such as registers, previous surveys or the sampling frame. The vector with predictor variables for the j[th] component of $\mathbf{y}_{mi}$ will be denoted by $\mathbf{x}_{ij}$ (j=1…J). A linear regression model for unit i and variable j can then be written as $y_{mij} = \mathbf{x}'_{ij}\beta_j + \varepsilon_{ji}$. For all J variables for unit i we can combine these models as follows:

$$\begin{bmatrix} y_{mi1} \\ \vdots \\ y_{miJ} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_{i1}\beta_1 \\ \vdots \\ \mathbf{x}'_{iJ}\beta_J \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iJ} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_{1i} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{x}'_{Ji} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_J \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iJ} \end{bmatrix}$$

or (5)

$$\mathbf{y}_{mi} = \mathbf{X}_i\beta + \varepsilon_i$$

13. The coefficient vector β can be estimated by ordinary least squares using the records with both $\mathbf{y}_{mi}$ and $\mathbf{X}_i$ observed: the set of records with $i \in O$, say. For these records we can define the residuals $\mathbf{e}_i = \hat{\mathbf{y}}_{mi} - \mathbf{X}_i\hat{\beta}$ and obtain $\hat{\beta}$ by minimization of $\sum_i \mathbf{e}'_i\mathbf{e}_i$. The resulting predicted values will, however, not be consistent with the constraints, neither for the for the records with observed values for $\mathbf{y}_{mi}$ nor for the records with missing values for $\mathbf{y}_{mi}$ (the imputations). To remedy this situation the model can be

extended with adjustments $\mathbf{a}_i$ that are defined such that the adjusted predictions do satisfy the constraints. This results in the following adjusted predictions:

$$\tilde{\mathbf{y}}_{mi} = \mathbf{X}_i \tilde{\beta} + \tilde{\mathbf{a}}_i , \tag{6}$$

with $\mathbf{R}'_m \tilde{\mathbf{y}}_{mi} = \mathbf{b}_i$. Note that in (6) estimators are indicated by ~ rather then by ^ to stress the fact that the parameter estimates as well as the predicted values for model (6) are different from those of model (5).

14. Since in (6) there is one adjustment parameter $\mathbf{a}_i$ for each observed (and missing) value $\mathbf{y}_{mi}$, this model cannot be estimated without restrictions on these parameters. Therefore we will require that the adjustments are as small as possible, in a least squares sense. This means that the $\tilde{\mathbf{a}}_i$ minimize $\tilde{\mathbf{a}}'_i\tilde{\mathbf{a}}_i$ subject to $\mathbf{R}'_m \tilde{\mathbf{y}}_{mi} = \mathbf{b}_i$. The Lagrangian for this constrained minimization problem can be written as:

$$L = \tfrac{1}{2}\tilde{\mathbf{a}}'_i\tilde{\mathbf{a}}_i - \tilde{\alpha}'_i(\mathbf{R}'_m(\mathbf{X}_i\tilde{\beta} + \tilde{\mathbf{a}}_i) - \mathbf{b}) , \tag{7}$$

with $\tilde{\alpha}_i$ a q-vector (q is the number of constraints) with Lagrange multipliers. A condition for a minimum is

$\partial L / \partial \tilde{\mathbf{a}}_i = \tilde{\mathbf{a}}_i - \mathbf{R}_m \tilde{\alpha}_i = 0$ and so $\tilde{\mathbf{a}}_i = \mathbf{R}_m \tilde{\alpha}_i$.

Thus, we have for the consistent adjusted predictions:

$$\tilde{\mathbf{y}}_{mi} = \mathbf{X}_i \tilde{\beta} + \mathbf{R}_m \tilde{\alpha}_i , \tag{8a}$$

which is an estimated regression of the target variables on the predictors and the constraints. For a sample of n records we can write this model as

$$\begin{bmatrix} \tilde{\mathbf{y}}_{m1} \\ \vdots \\ \tilde{\mathbf{y}}_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{R}_m & & \mathbf{0} \\ \vdots & & \ddots & \\ \mathbf{X}_n & \mathbf{0} & & \mathbf{R}_m \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_n \end{bmatrix} , \tag{8b}$$

If (8) is estimated by OLS, the residual vector is orthogonal to each of the “explanatory variables”, in this case the predictors and constraints. Hence, the OLS estimators $\tilde{\beta}$ and $\tilde{\alpha}_i$ satisfy the normal equations:

$$\sum\nolimits_{i \in O} \mathbf{X}'_i(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0} , \tag{9}$$

and

$$\mathbf{R}'_m(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0} ,\ i \in O \quad . \tag{10}$$

Equation (10) implies $\mathbf{R}'_m\mathbf{y}_{mi} = \mathbf{R}'_m\tilde{\mathbf{y}}_{mi}$ which shows that the predictions satisfy the constraints since $\mathbf{R}'_m\mathbf{y}_{mi} = \mathbf{b}_i$ because the observed values satisfy the constraints.

B. Estimation of parameters and missing values.

15. Equations (9) and (10) can be used to obtain expressions for calculating $\tilde{\alpha}$ and $\tilde{\beta}$. Solving (10) for $\tilde{\alpha}$ gives

$$\tilde{\alpha}_i = (\mathbf{R}'_m\mathbf{R}_m)^{-1}\mathbf{R}'_m(\mathbf{y}_{mi} - \mathbf{X}_i\tilde{\beta}) . \tag{11}$$

Substituting this result in (9) and solving for $\tilde{\beta}$ results in

$$\tilde{\beta} = \left[\sum_{i\in O} \mathbf{X}_i'(\mathbf{I} - \mathbf{R}_m(\mathbf{R}_m'\mathbf{R}_m)\mathbf{R}_m')\mathbf{X}_i\right]^{-1}\left[\sum_{i\in O} \mathbf{X}_i'(\mathbf{I} - \mathbf{R}_m(\mathbf{R}_m'\mathbf{R}_m)\mathbf{R}_m')\mathbf{y}_{mi}\right]$$
$$= \left[\sum_{i\in O} \mathbf{X}_i'\mathbf{M}\mathbf{X}_i\right]^{-1}\left[\sum_{i\in O} \mathbf{X}_i'\mathbf{M}\mathbf{y}_{mi}\right], \text{ say.} \quad (12)$$

Note that, since **M** is square and idempotent $\mathbf{X}_i'\mathbf{M}\mathbf{X}_i = \mathbf{X}_i'\mathbf{M}'\mathbf{M}\mathbf{X}_i$ so that $\tilde{\beta}$ can be estimated by applying OLS to the transformed model $\mathbf{y}_{mi} = \mathbf{Z}_i\beta + u_i$, with $\mathbf{Z}_i = \mathbf{M}\mathbf{X}_i$. Now, if $\tilde{\beta}$ is estimated either by using the transformed model or by using (12) directly, the $\tilde{\alpha}_i$ can be calculated subsequently by (11).

16. In order to use this model for imputation of records with missing values, we need the record specific parameter estimates $\tilde{\alpha}_i$ for these records. These cannot be calculated by (11) directly because they depend on the values $\mathbf{y}_{mi}$ that are missing. However, $\tilde{\alpha}_i$ depends only on $\mathbf{y}_{mi}$ in the form $\mathbf{R}_m'\mathbf{y}_{mi}$ which, according to (3), equals $-\mathbf{R}_o'\mathbf{y}_{oi}$ and can thus be calculated using the observed values.

17. In this approach to consistent imputation the adjustments are part of the model and estimated simultaneously with the other regression coefficients. This is different from an approach where the regression predictions are calculated first and adjustments are calculated subsequently. In our case this last approach would be to first estimate $\hat{\beta}$ in model (5) and then $\tilde{\alpha}_i$ according to (11). To illustrate the differences, we have, for 6 records of a business survey, calculated predictions for variables (2), (3) and (4) of the illustration in section 2. The model was a regression with, for each variable, a single predictor: its value from the previous year. In table 1 we show the Residual Sum of Squares for three types of predictions: 1. Model (5) without adjustments (predictions inconsistent with the constraints), 2. Model (5) with adjustments calculated after estimating $\hat{\beta}$ with OLS and 3, predictions according to model (8) with coefficients estimated by (11) and (12).

Table 1. Residual Sum of Squares (RSS) for three types of predictions

Model	*RSS*
Model (5) without adjustments	324
Model (5) with adjustments afterwards	308
Model (8) incorporating the adjustments	214

18. The results clearly show that using the constraints to obtain adjusted consistent predictions result in more accurate prediction, both when the adjustments are calculated separately from the regression coefficients and when the adjustments are part of the model. The most accurate predictions are, however, obtained in the latter case.

IV. OTHER MODELS

A. Weighted least squares models

19. An often used imputation model for economic data is ratio imputation. This model can be seen as a heteroscedastic regression model with a single predictor **x** and variance proportional to **x**. In our case with a multivariate target variable we can write the ratio models for each variable as

$$\begin{bmatrix} y_{mi1} \\ \vdots \\ y_{miJ} \end{bmatrix} = \begin{bmatrix} x_{i1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & x_{iJ} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \\ \beta_J \end{bmatrix} + \begin{bmatrix} \varepsilon_{mi1} \\ \vdots \\ \varepsilon_{miJ} \end{bmatrix} \quad (13)$$

$$\mathbf{y}_{mi} = \mathbf{X}_i\beta + \varepsilon_{mi},$$

and $\mathrm{var}(\varepsilon_{mi}) = diag(\sigma_j^2 x_{ij}) = \mathbf{V}_{mi}$. If this model is estimated by Weighted Least Squares (WLS), it is well known that the estimated regression coefficients equal the ratio of the mean of y to the mean of x: $\bar{y}_{mj} / \bar{x}_j$, hence the name ratio-imputation.

20. Again we will need to use adjustments $\mathbf{a}_i$ to make the predictions consistent with the constraints. In this case it seems appropriate to require that the adjustments are as small as possible in a weighted least squares sense, i.e. to minimize $\mathbf{a}_i'\mathbf{V}_{mi}^{-1}\mathbf{a}_i$, since this will result in larger adjustments for variables with larger error variance. Analogous to (7) we can write the Lagrangian for this constrained minimization problem as:

$$L = \tfrac{1}{2}\tilde{\mathbf{a}}_i'\mathbf{V}_{mi}^{-1}\tilde{\mathbf{a}}_i - \tilde{\alpha}_i'(\mathbf{R}_m'(\mathbf{X}_i\tilde{\beta} + \tilde{\mathbf{a}}_i) - \mathbf{b}), \tag{14}$$

Differentiating w.r.t. $\tilde{\mathbf{a}}_i$ gives,

$\partial L / \partial \tilde{\mathbf{a}}_i = \mathbf{V}_{mi}^{-1}\tilde{\mathbf{a}}_i - \mathbf{R}_m\tilde{\alpha}_i = 0$ and so $\tilde{\mathbf{a}}_i = \mathbf{V}_{mi}\mathbf{R}_m\tilde{\alpha}_i$.

Thus, we have for the consistent adjusted predictions:

$$\tilde{\mathbf{y}}_{mi} = \mathbf{X}_i\tilde{\beta} + \mathbf{V}_{mi}\mathbf{R}_m\tilde{\alpha}_i, \tag{15}$$

21. Estimating this model by WLS, entails minimizing $\sum_{i\in O}(\mathbf{y}_{mi} - \mathbf{y}_{mi})'\mathbf{V}_{mi}^{-1}(\mathbf{y}_{mi} - \mathbf{y}_{mi})$, which yields the following normal equations:

$$\sum_{i\in O}\mathbf{X}_i'\mathbf{V}_{mi}^{-1}(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0}, \tag{16}$$

and

$$\mathbf{R}_m'\mathbf{V}_{mi}\mathbf{V}_{mi}^{-1}(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{R}_m'(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0}, \tag{17}$$

Following the same reasoning as below (10), (16) shows that the predictions will satisfy the constraints. Solving (17) for $\tilde{\alpha}_i$ yields

$$\tilde{\alpha}_i = (\mathbf{R}_m'\mathbf{V}_{mi}^{-1}\mathbf{R}_m)^{-1}\mathbf{R}_m'(\mathbf{y}_{mi} - \mathbf{X}_i\tilde{\beta}). \tag{18}$$

Substituting this result in (16) and solving for $\tilde{\beta}$ results in:

$$\tilde{\beta} = \left[\sum_{i\in O}\mathbf{X}_i'(\mathbf{V}_{mi}^{-1} - \mathbf{R}_m(\mathbf{R}_m'\mathbf{R}_m)\mathbf{R}_m')\mathbf{X}_i\right]^{-1}\left[\sum_{i\in O}\mathbf{X}_i'(\mathbf{V}_{mi}^{-1} - \mathbf{R}_m(\mathbf{R}_m'\mathbf{R}_m)\mathbf{R}_m')\mathbf{y}_{mi}\right] \tag{19}$$

B. Using a log-transform

22. Non-negative economic variables are often modelled on a log-scale. In this case also, we can use a model that incorporates the constraints to obtain consistent predictions. Consider the model

$$\mathrm{Log}\,\tilde{\mathbf{y}}_{mi} = \mathbf{X}_i\tilde{\beta} + \mathbf{R}_m\tilde{\alpha}_i, \tag{20}$$

and let $\mathbf{V}_{mi} = diag(\tilde{\mathbf{y}}_{mi})$. Thus, the variance is proportional to the mean. Estimating this model by WLS requires the minimization of

$$Q = \sum_i(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi})'\mathbf{V}_{mi}^{-1}(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) \tag{21}$$

Setting the derivatives of Q with respect to $\tilde{\beta}$ and $\tilde{\alpha}_i$ equal to zero yields the normal equations:

$$\sum_{i\in O}\mathbf{X}_i'\{diag(\partial\exp(\tilde{l}_{mi})/\partial\tilde{\mathbf{y}}_{mi}\}\mathbf{V}_{mi}^{-1}(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \sum_{i\in O}\mathbf{X}_i'(\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0}, \tag{22}$$

and

$$\mathbf{R}'_m \{diag(\partial \exp(\tilde{l}_{mi}) / \partial \tilde{\mathbf{y}}_{mi}\} \mathbf{V}_{mi}^{-1} (\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{R}'_\mathbf{m} (\mathbf{y}_{mi} - \tilde{\mathbf{y}}_{mi}) = \mathbf{0}, \quad (23)$$

with $\tilde{l}_{mi} = \mathbf{X}_i \tilde{\beta} + \mathbf{R}_m \tilde{\alpha}_i$ and using $\partial \exp(\tilde{l}_{mi}) / \partial \tilde{\mathbf{y}}_{mi} = diag(\tilde{\mathbf{y}}_{mi}) = \mathbf{V}_{mi}$.

23. By the same reasoning as below (10), equation (23) again shows the consistency of the predictions. Since (20) is a non-linear model, estimators for the coefficients cannot be expressed in closed form. However, the normal equations are the same as those for Loglinear models (or Poisson regression) and the iteratively reweighted least squares algorithm that is used for these models can be applied (See, e.g. McCullagh and Nelder, 1989).

References

De Waal, T. (2003). Processing Erroneous and Unsafe data, Ph.D. Thesis, Erasmus University, Rotterdam.

Little, R.J.A. and Rubin, D.B. (1987). Statistical Analysis with Missing Data. New York: Wiley.

McCullagh, P. and Nelder, J. A., (1989), Generalized Linear Models, Chapman and Hall, London.

Rao, C.R. (1973). Linear statistical inference and its applications, second edition, New York: Wiley.

Statistics Canada (2005). Banff: Functional description of the Banff system for edit and imputation. Generalized System Methods Section, Business Survey Methods Division, Statistics Canada, Ottawa.

Tempelman, C. (2004), Imputation of economic data under several linear restrictions, paper presented at the European Conference on Quality and Methodology in Official Statistics 2004 (Q2004), Mainz.
