Identification of Functional Forms and Predictor Variables in Generalized Variance Functions for Price Indices

MoonJung Cho
Cho.Moon@bls.gov *

John L. Eltinge †

Abstract

Generalized variance functions (GVFs) can provide useful approximations of error variances, especially for complex-survey cases in which (1) standard variance estimators have insufficient degrees of freedom for direct use, (2) confidentiality restrictions prevent the release of design information used in direct variance estimation, or (3) computational time requirements may be incompatible with short timelines for publication of estimates. Much of the GVF literature has considered variances of estimators for population proportions, and used population totals and sample sizes as predictors in the resulting variance-function models. Some surveys, however, have variance-estimation settings that meet criteria (1), (2) or (3) above, but involve point estimators that are complex nonlinear functions of the data. This paper derives the functional forms and predictors for a GVF in this setting; applies the results to a class of price-index estimators; and evaluates the properties of the resulting GVFs.

Key Words: Degrees of freedom, Design-based inference, International Price Program, Superpopulation model, Variance function model

1. Introduction

In analysis of data collected through complex sample designs, one often would like to use standard design-based or model-based variance estimators, $\hat{V}$, say. However, use of $\hat{V}$ may be problematic for one of several reasons, including the following.

1. Standard variance estimators have insufficient degrees of freedom for direct use.

2. Confidentiality restrictions prevent the release of design information used in direct variance estimation.

3. Computational time requirements may be incompatible with short timelines for publication of estimates.

Consequently, the sample survey literature has developed an alternative set of methods for estimation of "generalized variance functions" in which the true design or design-model variance is approximated as a function of known predictors Z. For some background on generalized variance functions for survey data, see Johnson and King (1987), Valliant (1987) and references cited therein.

Much of the GVF literature has focused on the variances of point estimators of population proportions or population totals related to a binary outcome variable. The current paper, however, considers the more complex setting in which the point

*U.S. Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Avenue NE, Washington, DC 20212

†U.S. Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Avenue NE, Washington, DC 20212

estimator of interest is a complex nonlinear function used in the estimation of price indices for the International Price Program (IPP) of the Bureau of Labor Statistics. Section 2 summarizes the sample design for the IPP Import program, and Section 3 describes the hierarchical structures used to produce nonlinear point estimators of the IPP price indices. Section 4 provides additional details on construction of weights and computation of price-index estimators. Section 5 presents several candidate variance function models for price index estimators. Section 6 discusses the use of simulation methods to fit the models introduced in Section 5. Section 7 presents numerical results from the application of these ideas to short-term price ratio estimators from the IPP. Section 8 reviews the main ideas from this paper and considers possible areas of future work.

2. Description of the IPP Import Sample Design

For a detailed description of the IPP sample design, see Bobbitt et al. (2007) and BLS Handbook of Methods (2007). For the current discussion, the following features of the design are of special interest.

Stratification: The population of all imported goods is partitioned into a total of approximately 120 strata, defined by the Harmonized Classification System. The import sampling frame is from the U.S. Customs and Border Protection (USCBP). This frame contains information about all import transactions that were filed with the USCBP during the reference year. For each transaction, the frame information includes a company identifier (usually the Employer Identification Number), the detailed product category (Harmonized Tariff number) of the goods that are being shipped and the corresponding dollar value of the shipped goods.

Primary Sample Units (PSUs): Within a given stratum h containing N_h companies, IPP selects n_h companies without replacement and with probabilities proportional to size. A given company may import products in several different strata. However, for purposes of the sample design, each company $\times$ stratum combination is treated as a distinct PSU, and selection decisions are independent across strata.

Secondary Sample Units (SSUs): Each selected PSU contains one or more classification groups (CGs) defined by a more detailed product classification. (The sampling strata generally are defined by the Harmonized Classification System at the two-digit level, while the CGs generally are defined at the six to ten digit level.) Within a given selected PSU, the IPP selects one or more CGs through systematic sampling with probability proportional to size. Here again, the size measure equals the trading dollar volume assigned to a given CG. Each given selected SSU receives a fixed number of price quotes. Note that within a selected company, a given CG may be selected in the sample more than once.

Tertiary Sample Units: Within a given selected SSU, field economists and the data provider (responding company) determine the listing of all applicable imported items. The field economist then selects a systematic sample of price quotes from this listing. Item-level selection probabilities are proportional to size and size is again equal to trading dollar value for the item in question. Thus certain items that have large trading dollar values may provide more than one price quote to a given SSU.

3. Description of the Hierarchical Estimation Structure

The IPP publishes three distinct sets of index estimates based on three different hierarchical classification structures know as, respectively, the Harmonized Classification System (HS), the North American Industrial Classification System (NAICS) and the Bureau of Economic Analysis (BEA) Classification System. We begin with a description of estimation under the Harmonized Classification System.

The IPP computes Classification Group level price indices through two steps. In the first step, IPP computes short-term ratio (STR) and long-term ratio (LTR) price indices at the establishment|CG (Weight Group) levels based on price quotes collected at the item level, using sample-based item weights. Specific formulas for STR and LTR indices are provided in Section 4 below.

Second, IPP computes STR and LTR values at the CG level, as well as the establishment|CG level sample weights (known commonly as weight group weights.)

Under the HS, the level of aggregation immediately above the CG level is commonly called the stratum-lower. The IPP computes STR and LTR indices for a given stratum-lower based on the CG level index estimators and fixed weights available at the CG level. The U.S. Census Bureau provides these CG level weights which are updated annually. The CG level fixed weights are intended to reflect the population values (trading dollar volumes) for each CG in the population.

Under the HS, at levels of aggregation above the stratum-lower, STR and LTR values are computed based on index values from the next-lowest levels of aggregation, again using fixed population weights.

Under both the NAICS and BEA systems, estimation of the price indices at the CG level and finer levels is identical to estimation at these levels under the HS. In this sense, it can be useful to focus special attention on properties of CG level index estimators. At coarser levels of aggregation, the HS, NAICS and BEA systems use different hierarchical structures and thus will use the CG-level index estimators in somewhat different ways.

4. Index and Weight Estimation

The IPP uses items that are initiated and re-priced every month to compute its indices of price change. These indices are calculated using a modified Laspeyres index formula. For each classification system, the IPP calculates the estimates of price change using an aggregation tree structure beginning with items, weight groups, classification groups, . . . , and finally overall. Weight groups are defined by the intersection of establishment and product classification group. Note that there could be many different levels, such as stratum-lower and stratum-upper which is above the stratum-lower. The formula is basically the same for all levels: each parent's index is computed from its children's indices. For example, a stratum index is computed from the stratum's children's indices. These children could be classification groups, stratum-lowers, stratum-uppers or any combination of them. Define $Child[h]$ to be the set of all stratum-lowers, stratum-uppers or classification groups directly below stratum h in an aggregation tree. In practice, θ_h^t, a short-term ratio for a stratum h at time t, is computed from the weighted long-term ratios, I_c^t and I_c^{t-1}, from its children's set.

$$\hat{\theta}_h^t = \sum_{c \in Child[h]} w_c I_c^t \left(\sum_{c \in Child[h]} w_c I_c^{t-1} \right)^{-1} \tag{1}$$

where w_c is the weight of an element c of $Child[h]$, and I_c^t the long-term price ratio of c at time t. $\hat{\theta}_h^t$ is then used in computing I_h^t, a long-term ratio for a stratum h at time t.

$$I_h^{t,0} = \prod_{u=0}^{t} \hat{\theta}_h^u \tag{2}$$

This general formula (1) is used until the desired aggregation level index is obtained. See Bobbitt et al. (2005, Section 4: Index Estimation) and Powers et al.(2006, Section 3: Point Estimation for the International Price Program) for more detailed explanation of the IPP index formula.

The weight group weight, w_g, is a trade dollar value divided by a selection probability of weight group g. Item weight is the weight group's weight divided by number of items in the weight group, and hence, the weight group weight, w_g can be expressed as the sum of item weights, w_{gi}, in the weight group, g: $w_g = \sum_i w_{gi}$. Starting in January 2004, the IPP changed to the fixed classification group weight, which is the total trade dollar value within each classification group. See Bobbitt et al. (2005, Section 5.1: Item and Weight Group Weight Formulas) for more detailed background on the IPP weight computation.

5. Methods for Estimation of the Coefficients of a Generalized Variance Function

Now consider approximations for the variance of a short-term ratio price index estimator $\hat{\theta}_{tg}$ for the time period $t = 1, \ldots, T$ and two-digit stratum $g = 1, \ldots, G$. In its most general form, we can write the GVF model

$$Y = X\gamma \tag{3}$$

where Y is a $TG \times 1$ vector with $[(t-1)\,G + g]\,th$ element equal to either V_{tg} or $ln(V_{tg})$; X is a $TG \times K$ - dimensional matrix of predictor variables with $[(t-1)\,G + g]\,th$ row providing predictors for V_{tg}; and γ is a $K \times 1$-dimensional vector of regression coefficients. Eight cases are of special interest:

Case 1: Constant variances
Let $K = 1$; $X = 1_{TG\times 1}$; and $\gamma = \gamma_0$ a constant scalar. Then model (3) reduces to

$$Y_{tg} = \gamma_0 \qquad t = 1, \ldots, T;\ g = 1, \ldots, G\,.$$

Case 2: Product effects with no time effects
Let $K = G$; $X = (X_1, \ldots, X_G)$ with $X_1 = 1_{TG\times 1}$; and X_g equal to a $TG \times 1$ vector with $[(t-1)\,G + (g-1)]\,th$ elements equal to 1 and all other elements equal to 0 for $g = 2, \ldots, G$ and $t = 1, \ldots, T$; and $\gamma = (\gamma_0, \gamma_{21}, \ldots \gamma_{2G-1})$. Then model (3) reduces to

$$Y_{tg} = \gamma_0 + \gamma_{2g} \qquad t = 1, \ldots, T;\ g = 1, \ldots, G\,. \tag{4}$$

Case 3: Time effects with no product effects
Let $K = T$; $X = (X_1, \ldots, X_T)$ with $X_0 = 1_{TG\times 1}$; and X_t equal to a $TG \times 1$ vector with $[(t-2)\,G + g]\,th$ elements equal to 1 and all other elements equal to 0 for $t = 2, \ldots, T$ and $g = 2, \ldots, G$; and $\gamma = (\gamma_0, \gamma_{11}, \ldots \gamma_{1T-1})$. Then (3) reduces to

$$Y_{tg} = \gamma_0 + \gamma_{1t}\,. \tag{5}$$

Case 4: Time and product effects with no interaction terms
Let $K = T + G - 1$; $X = (X_1, \ldots, X_{T+G-1})$ with $X_1 = 1_{TG\times1}$; and X_t equal to a $TG\times1$ vector with $[(t-2)\,G+g]\,th$ elements equal to 1 and all other elements equal to 0 for $t = 2, \ldots, T$; X_{T+g} equal to a $TG \times 1$ vector with $[(t-1)\,G+(g-1)]\,th$ elements equal to 1 and all other elements equal to 0 for $g = 2, \ldots, G$;
and $\gamma = (\gamma_0, \gamma_{11}, ...\gamma_{1T}, \gamma_{21}, \ldots, \gamma_{2G})$. Then (3) reduces to

$$Y_{tg} = \gamma_0 + \gamma_{1t} + \gamma_{2g} \qquad t = 1, \ldots, T;\; g = 1, \ldots, G \tag{6}$$

Case 5: Continuous predictors
For some of the cases considered in Sections 1 through 6, the variance function approximations led to the model

$$Y_{tg} = \gamma_1 + \sum_{k=2}^{K} \gamma_k X_{tgk} \tag{7}$$

for a specified set of K vectors of predictors, e.g., coefficients of variation for specified weights, expression (7) is a special case of (3) with $X_1 = 1_{TG\times1}$; and X_k equal to a $TG \times 1$ vector with $[(t-1)\,G+g]\,th$ elements equal to X_{tgk} and $\gamma = (\gamma_0, \gamma_2, ...\gamma_k)$. Using the notation from Case 5,

$$\begin{aligned} X_{tg2} &= n_{tg}^{-1} \sum_{i=1}^{n_{tg}} w_{tgi}^2 \\ &= (\bar{w}_{tg})^2 [C_{tg}^2 + 1] \end{aligned} \tag{8}$$

where $\bar{w}_{tg} = n_{tg}^{-1} \sum_{i=1}^{n_{tg}} w_{tgi}$, $C_{tg}^2 = (\bar{w}_{tg})^{-2} n_{tg}^{-1} \sum_{i=1}^{n_{tg}} (w_{tgi} - \bar{w}_{tg})^2$, and w_{tgi} is weight for child index i used in computation of the parent index estimator θ_{tg}, STR for time t and two-digit stratum g, and n_{tg} is number of child indices that contribute to the computation of the parent index θ_{tg}.
Case 6:
Fit (7) using an intercept and the predictor X_{tg2} only,

$$Y_{tg} = \gamma_1 + \gamma_2 + X_{tg2} + error\,. \tag{9}$$

Case 7:
Fit a more general form of (3) that includes an intercept, time and product effects (with no interactions) and the X_{tg2} effect,

$$Y_{tg} = \gamma_0 + \gamma_{1t} + \gamma_{2g} + \gamma_{2g}^{*} X_{tg2}\,. \tag{10}$$

This will use a total of $T + G$ parameters.
Case 8:

$$X_{tg3} = \bar{n}_{tg} \tag{11}$$

6. Use of Simulation Results in Fitting GVF Models

In general, one may consider two closely related variance-function models. The first model involves the true design variance of $\hat{\theta}_{tg}$, $V_{design}(\hat{\theta}_{tg})$. We approximate $V_{design}(\hat{\theta}_{tg})$ as a sum

$$\begin{aligned} V_{design}(\hat{\theta}_{tg}) &= V_{tg} \\ &= f(X_{tg}, \gamma) + q_{tg} \end{aligned} \tag{12}$$

where $f(\cdot, \cdot)$ is a function of specified form, X_{tg} is a vector of characteristics that may be associated with the design variance of $\hat{\theta}_{tg}$, and γ is a vector of unknown coefficients that we will need to estimate. In addition, q_{tg} reflects the lack of fit in approximating the true variances V_{tg} with functions $f(X_{tg}, \gamma)$. In the numerical work in subsequent sections, we will approximate V_{tg} with the simulation-based quantity

$$\hat{V}_{sim}(\hat{\theta}_{tg}) = (1000-1)^{-1} \sum_{s=1}^{1000} \left\{ \hat{\theta}_{tgs} - \hat{E}_{sim}(\hat{\theta}_{tg}) \right\}^2 \tag{13}$$

where

$$\hat{E}_{sim}(\hat{\theta}_{tg}) = 1000^{-1} \sum_{s=1}^{1000} \hat{\theta}_{tgs} \tag{14}$$

In addition, preliminary numerical work indicated that relatively good model fits were obtained after a logarithmic transformation, so we fit the models

$$ln\,\{\hat{V}_{sim}(\hat{\theta}_{tg})\} \quad = \quad X_{tg}\,\gamma + q^*_{tg} \tag{15}$$

with ordinary least squares regression. model(15) is equivalent to model(12) with $ln\,\{f(X, \gamma)\} \,=\, X_{tg}\,\gamma$ and

$$\begin{aligned} q^*_{tg} &= ln\left[1 + \{f(X_{tg}, \gamma))\}^{-1}\, q_{tg}\right] \\ &= ln\,[1 + exp(-X_{tg}\gamma)\, q_{tg}] \end{aligned} \tag{16}$$

The second model involves the variance estimators $\hat{V}_{tg}$. Consider the decomposition

$$\begin{aligned} \hat{V}_{tgs} &= V_{tg} \times (1 + e_{tgs}) \\ &= \{f(X_{tg}, \gamma) + q_{tg}\}(1 + e_{tgs}) \end{aligned} \tag{17}$$

The term e_{tgs} represents the relative error in $\hat{V}_{tg}$ as an estimator of V_{tg} , i.e., $e_{tgs} = (\hat{V}_{tgs}/V_{tg}) - 1$. Under standard approximations, we often treat $d(1+e_{tg})$ as following approximately a chi-square distribution on d degrees of freedom, where d is some fixed positive number. In that case, the errors e_{tg} will have a mean equal to zero and a variance equal to $2/d$. On the other hand, if $\hat{V}_{tg}$ is biased as an estimator of V_{tg}, then the error term e_{tg} will have a non-zero mean.

In subsequent numerical work, we will fit models for

$$\hat{E}_{sim}(\hat{V}_{tgs}) = 1000^{-1} \sum_{s=1}^{1000} \hat{V}_{tgs}\,. \tag{18}$$

Under the specified simulation design, $\hat{E}_{sim}(\hat{V}_{tgs})$ is an approximation to

$$V_{tg}(1 + \bar{e}_{tg}) \quad = \quad \{f(X_{tg}, \gamma) + q_{tg}\}(1 + \bar{e}_{tg}) \tag{19}$$

where $\bar{e}_{tg} \,=\, 1000^{-1} \sum_{s=1}^{1000} e_{tgs}$. Note especially that in our simulation work, we constructed a single finite population, and thus have a single finite population term q_{tg} which reflects the variability in generating a given finite population from a hypothetical superpopulation. On the other hand, we generated 1000 samples from this finite population, and $\bar{e}_{tg}$ represents the resulting average of the 1000 relative errors $(\hat{V}_{tg}/V_{tg}) - 1$.

Now apply the logarithmic transformation to model (19),

$$\begin{aligned} ln\{\hat{E}_{sim}(\hat{V}_{tg})\} &= ln\{f(X_{tg}, \gamma) + q_{tg}\} + ln(1 + \bar{e}_{tg}) \\ &= X_{tg}\gamma + q^*_{tg} + e^*_{tg} \end{aligned} \tag{20}$$

where $e^*_{tg} = ln(1 + e_{tg})$

Under the logarithmic transformation, if the terms e^*_{tg} have a mean of zero and are independent of X_{tg}, then models (15) and (20) will lead to the same coefficient vector γ, and will differ only in the magnitude of the variances of the error terms, $V(q^*_{tg})$ and $V(q^*_{tg} + e^*_{tg})$, respectively.

On the other hand, if the terms e^*_{tg} have a non-zero mean or are associated with the predictors X_{tg}, then the coefficients for models (15) and (20) generally will be different. Thus, comparison of results from (15) and (20) may help in exploration of cases in which the bootstrap variance estimator displays nontrivial bias.

7. Numerical results

The IPP created a large finite population U from the frame information (Chen et al., 2007; Cho and Eltinge, 2007, Appendix: Universe Creation Procedure). To simulate price ratios for the universe, the IPP used an historical database with 13 years (from September 1993 to June 2005) worth of price ratios which were stored in cells defined by classification group and month. There were about 4 million non-imputed item STRs in the historical data. The IPP then drew 1,000 samples based on an approximation to the complex IPP design.

Seven chapters (two-digit strata) were used in our simulation study based on the number of quotes in the historical database and preliminary results on their distributions (Cho and Eltinge, 2007, Section 3.5: Selecting strata). From each sample s, we computed short-term ratio point estimate $\hat{\theta}_{tgs}$ and bootstrap variance estimate $\hat{V}_{tgs}$ for $\hat{\theta}_{tgs}$ for $t = 1, \ldots, 36$ months, $g = 1, \ldots, 7$ chapters.

Consequently, the sample variance of the point estimates, computed from the 1000 samples, gives us an approximate value of the design variance of our STR estimator, $\hat{\theta_{tg}}$. Similarly, the sample mean of our bootstrap variance estimates, taken over the 1000 samples, gives us an approximate value of the design expectation of bootstrap variance estimator. For month t and chapter g, the following are the estimators of primary interest:

$$V_{sim} = \hat{V}_{sim}(\hat{\theta}_{tg})$$

in (13), and

$$\hat{E}_{sim}(\hat{V}_{Boot}) = 1000^{-1} \sum_{s=1}^{1000} \hat{V}_{Boot,s}$$

in (18).

Candidate predictors are:

n_{tgs}: the number of classification groups used in computation of chapter index estimator;

$\bar{n}_{tg} = 1000^{-1} \sum_{s=1}^{1000} n_{tgs}$: the average of n_{tgs} across the samples;

w_{tgis}: the weight of classification group i used in computation of chapter index estimator;

$\bar{w}_{tgs} = n_{tgs}^{-1} \sum_{i=1}^{n_{tgs}} w_{tgis}$: the average of w_{tgis} in a given sample;
$C_{tgs}^2 = \bar{w}_{tgs}^{-2} n_{tgs}^{-1} \sum_{i=1}^{n_{tgs}} (w_{tgis} - \bar{w}_{tgs})^2$: squared coefficient of variation of weights;
$X_{tgs} = \bar{w}_{tgs}^2 \left[C_{tgs}^2 + 1 \right]$ and $\bar{X}_{tg} = 1000^{-1} \sum_{s=1}^{1000} X_{tgs}$;
$Indicator\{g\}$: indicator for chapter membership.

The following are simple models that we considered:

$$ln\{\hat{E}_{sim}(\hat{V}_{tg})\} = \gamma_0 + \gamma_1 \bar{n}_{tg} \quad (21)$$

$$ln\{\hat{E}_{sim}(\hat{V}_{tg})\} = \gamma_0 + \gamma_2 \bar{X}_{tg} 10^{-15} \quad (22)$$

$$ln\{\hat{E}_{sim}(\hat{V}_{tg})\} = \gamma_0 + \gamma_1 \bar{n}_{tg} + \gamma_2 \bar{X}_{tg} 10^{-15} \quad (23)$$

$$\begin{aligned} ln\{\hat{E}_{sim}(\hat{V}_{tg})\} = \; & \gamma_0 + \gamma_1 \bar{n}_{tg} + \gamma_2 \bar{X}_{tg} 10^{-15} \\ & + \gamma_3 \bar{n}_{tg} \bar{X}_{tg} 10^{-15} \end{aligned} \quad (24)$$

$$ln\{\hat{E}_{sim}(\hat{V}_{tg})\} = \gamma_0 + \gamma_4 \times Indicator\{g\} \quad (25)$$

Note that $\hat{E}_{sim}(\hat{V}_{tg})$ was log-transformed and $\bar{X}_{tg}$ has been re-scaled for better fitting.

Table 1 displays model fitting results for Model (25) which was the best fit among models we considered; its R^2 was 0.95. Table 2 displays model fitting results for Model (23) which was the best model without using chapter membership; its R^2 was 0.80. Addition of interaction of $\bar{X}_{tg}$ and $\bar{n}_{tg}$ to Model(23) as in (24) gives only slight improvement to $R^2 = 0.82$.

Table 3 compares mean confidence interval widths of θ_{tg} based on the bootstrap variance estimator $\hat{V}_{Boot}$ and the GVF estimator V_{GVF}^*. Confidence interval widths from the bootstrap variance estimator were narrower than the ones from the GVF estimator except for P09 and P90 where the differences were rather marginal.

We also fit $ln\{V_{sim}\}$ for all models from (21) to (25). The resulting coefficient estimates and R^2 for $ln\{V_{sim}\}$ were similar with the ones from $ln\{\hat{E}_{sim}(\hat{V}_{tg})\}$.

Figures 1 and 2 display scatter plots of $ln\,\{\hat{E}_{sim}(\hat{V}_{tg})\}$ against $\bar{n}_{tg}$ and $\bar{X}_{tg}10^{-15}$ respectively. For both cases, the vertical axis corresponds to values of $ln\,\{\hat{E}_{sim}(\hat{V}_{tg})\}$. In Figure 1, note that each chapter has 252 values of $ln\,\{\hat{E}_{sim}(\hat{V}_{tg})\}$ which are grouped together at each chapter's $\bar{n}_{tg}$ value. In a similar way, $ln\,\{\hat{E}_{sim}(\hat{V}_{tg})\}$ are grouped at each chapter's $\bar{X}_{tg}10^{-15}$ value in Figure 2.

Quantile-Quantile (QQ) plots were plotted to assess standard χ^2 approximation for $Vratio = \hat{V}_{Boot,tg}/V_{sim}(\hat{\theta}_{tg})$. Figures 3 and 4 display quantile-quantile plots of $Vratio = \hat{V}_{Boot,tg}/\{V_{sim}(\hat{\theta}_{tg})\}$ against χ_d^2/d for P90 and P07 respectively. The red diamonds represent percentiles from $Vratio$ and the pth percentile is the value where p percent of the data lay below or equal to the value. Under regularity conditions, if $Vratio$ had a χ^2 distribution, then the percentiles from $Vratio$ would approximately follow the percentiles from χ^2. Then the QQ plot in Figures 3 and 4 should have its points arranged along a black line with a slope of 1 and an intercept of 0. We observed slight underestimation of $\hat{V}_{Boot,tg}$ in both P07 and P90.

Table 1: Model Fitting Results for Model (25): Best fit: $R^2 = 0.95$

Z	Coefficient	se	t-value	p-value
Intercept	-13.581	0.124	-109.16	< 0.0001
P07	7.961	0.176	45.25	< 0.0001
P08	7.081	0.176	40.25	< 0.0001
P09	5.215	0.176	29.64	< 0.0001
P22	1.214	0.176	6.90	< 0.0001
P61	-0.082	0.176	-0.46	0.6424
P74	4.575	0.176	26.00	< 0.0001

Table 2: Model Fitting Results for Model (23): Best fit without using chapter membership: $R^2 = 0.80$

Z	Coefficient	se	t-value	p-value
Intercept	-6.120	0.15084	-40.57	< 0.0001
$\bar{X}_{tg}$	-0.014	0.00068	-21.04	< 0.0001
$\bar{n}_{tg}$	-0.028	0.00116	-24.18	< 0.0001

Table 3: Comparison of Mean Confidence Interval Widths

Chapter	$\hat{V}_{Boot}$	V^*_{GVF}
P07	0.5281	0.5569
P08	9.7333	9.8833
P09	0.0973	0.0958
P22	0.0255	0.0280
P61	0.0371	0.0512
P74	0.0889	0.0934
P90	0.0072	0.0070

8. Discussion

This paper has considered generalized variance functions for point estimators of short-term ratio price indices from the Import component of the International Price Program of the U.S. Bureau of Labor Statistics. One could consider several extensions of this work. First, the numerical results presented here were based on seven chapters (two-digit product categories) that were selected to display a wide range of volatility across months, different patterns of seasonality, different numbers of observations in the historical database, and different distributions of price-change values within a given month. Other IPP chapters may have different characteristics, and thus may lead to different model-fitting results. In keeping with the GVF literature on fitting different GVF models for different groups of estimands (e.g., Wolter, 1985, Section 5.3), one could explore the extent to which different groups of IPP chapters may warrant different choices among the potential GVF models (21) - (25).

Second the numerical results in Section 7 suggest that confidence intervals for short-term ratio price indices have roughly comparable coverage rates and mean widths when these intervals are based on direct bootstrap variance estimators or the GVF model (25), respectively. In general, however, confidence interval performance will depend on a balance of several factors, including the stability of the direct bootstrap variance estimator; the magnitudes of the equation-error terms q_{tg} in equation (12); and the number of observations (t, g) that contribute to a given GVF model fit. Thus, it would be useful to develop tools to identify cases in which GVF-based confidence intervals may have better properties than bootstrap-based confidence intervals.

Third, P. Bobbitt, J. Himelein, and L. Lang have suggested the possible use of GVF models in the allocation of sample sizes at one or more levels in the sample-design hierarchy described in Section 2. It would be of interest to explore this in greater depth; to study efficiency gains that may follow from this method; and to evaluate the extent to which prospective efficiency gains depend on the goodness-of-fit for specific forms of model (3).

9. Acknowledgment

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the U.S. Bureau of Labor Statistics. The authors thank Jeff Blaha, Patrick Bobbitt, Jim Himelein and Larry Lang at the U.S. International Price Program Division for many helpful comments on the International Price Program; Zedenek Patek and Jean-Francois Beaumont at the Statistics Canada for helpful comments on the variance estimation; and Rick Valliant and Bert Balk for helpful comments on asymptotics for price indices.

REFERENCES

AKRITAS, M. and ARNOLD, S. (2000), "Asymptotics for Analysis of Variance When the Number of Levels is Large," *Journal of the American Statistical Association*, Vol. 95, No. 449 (Mar., 2000), 212-226.

BINDER, D. A. (1983), "On the Variances of Asymptotically Normal Estimators from Complex Surveys," *International Statistical Review*, 51, 279-292.

BOBBITT, P. A., CHO, M. J., and EDDY, R. M. (2005), "Comparing Weighting Methods in the International Price Program," *Proceedings of the American Statistical Association*, Government Statistics Section [CD-ROM], 1006-1014.

BOBBITT, P. A., PABEN, S. P., CHO, M. J., HIMELEIN, J. A., CHEN, T-C., and ERNST, L. R. (2007), "Application of the Bootstrap Method in the International Price Program," *Proceedings of the American Statistical Association*, Survey Research Methods [CD-ROM], 2910-2917.

BUREAU OF LABOR STATISTICS, *BLS Handbook of Methods*, U.S. Department of Labor. Available at http://www.bls.gov/opub/hom/homtoc.htm

CAI, T. T., and Wang, L. (2008), "Adaptive Variance Function Estimation in Heteroscedastic Nonparametric Regression," *Annals of Statistics*, 36, 2025-2054.

CASELLA, G., and BERGER, R. L. (1990), *Statistical Inference*, Duxbery.

CHEN, T-C., BOBBITT, P. A., HIMELEIN, J. A., PABEN, S. P., CHO, M. J., and ERNST, L. R.(2007), "Variance Estimation for International Price Program Indexes," *Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], 1427-1434.

CHO, M. J., CHEN, T-C, BOBBITT, P. A., HIMELEIN, J. A., PABEN, S. P., ERNST, L. R., and ELTINGE, J. L.(2007), "Comparison of Simulation Methods Using Historical Data in the U.S. International Price Program," *Proceedings of the American Statistical Association*, Third International Conference on Establishment Surveys [CD-ROM], 248-255.

COCHRAN, W. G. (1977), *Sampling Techniques*, John Wiley and Sons.

FISHER, I. K. (1922), *The Making of Index Numbers*, Boston: Houghlin Mifflin and Co.

JOHNSON, E. G., and KING, B. F. (1987), "Generalized Variance Functions for a Complex Sample Survey," *Journal of Official Statistics*, 3, 235-250.

KISH, L. (1995), "Methods for Design Effects," *Journal of Official Statistics*, 11, 55-77.

KORN, E. L., and GRAUBARD, B. I. (1990), "Simultaneous Testing of Regression Coefficients with Complex Survey Data: Use of Bonferroni t statistics," *The American Statistician*, 44, 270-276.

POWERS, R., ELTINGE, J. L., and CHO, M. J. (2006), "Evaluation of the Detectability and Inferential Impact of Nonresponse Bias in Establishment Surveys," *Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], 3577-3583.

RAO, J. N. K., WU, C. F. J., and YUE, K. (1992), "Some recent work on Resampling methods for complex surveys," *Survey Methodology*, Vol.18, No. 2, 209-217.

SHAO, J. (1996), "Resampling methods in Sample surveys," *Statistics*, 27, 203-254.

VALLIANT, R. (1987), "Generalized Variance Functions in Stratified Two-Stage Sampling," *Journal of American Statistical Association*, 82, 499-508.

VALLIANT, R. (1988), "Estimation of Laspeyres Price Indexes Using the Prediction Approach for Finite Population Sampling," *Journal of Business & Economic Statistics*, Vol. 6, No. 2 (Apr., 1988), 189-196.

VALLIANT, R. and MILLER, S. (1989), "A Class of Multiplicative Estimators of Laspeyres Price Indexes," *Journal of Business & Economic Statistics*, Vol. 7, No. 3 (Jul., 1989), 387-394.

VALLIANT, R. (1991), "Variance Estimation for Price Indexes from a Two-Stage Sample with Rotating Panels," *Journal of Business & Economic Statistics*, Vol. 9, No. 4 (Oct., 1991), 409-422.

WOLTER, K. M. (1985), *Introduction to Variance Estimation.* New York: Springer-Verlag.

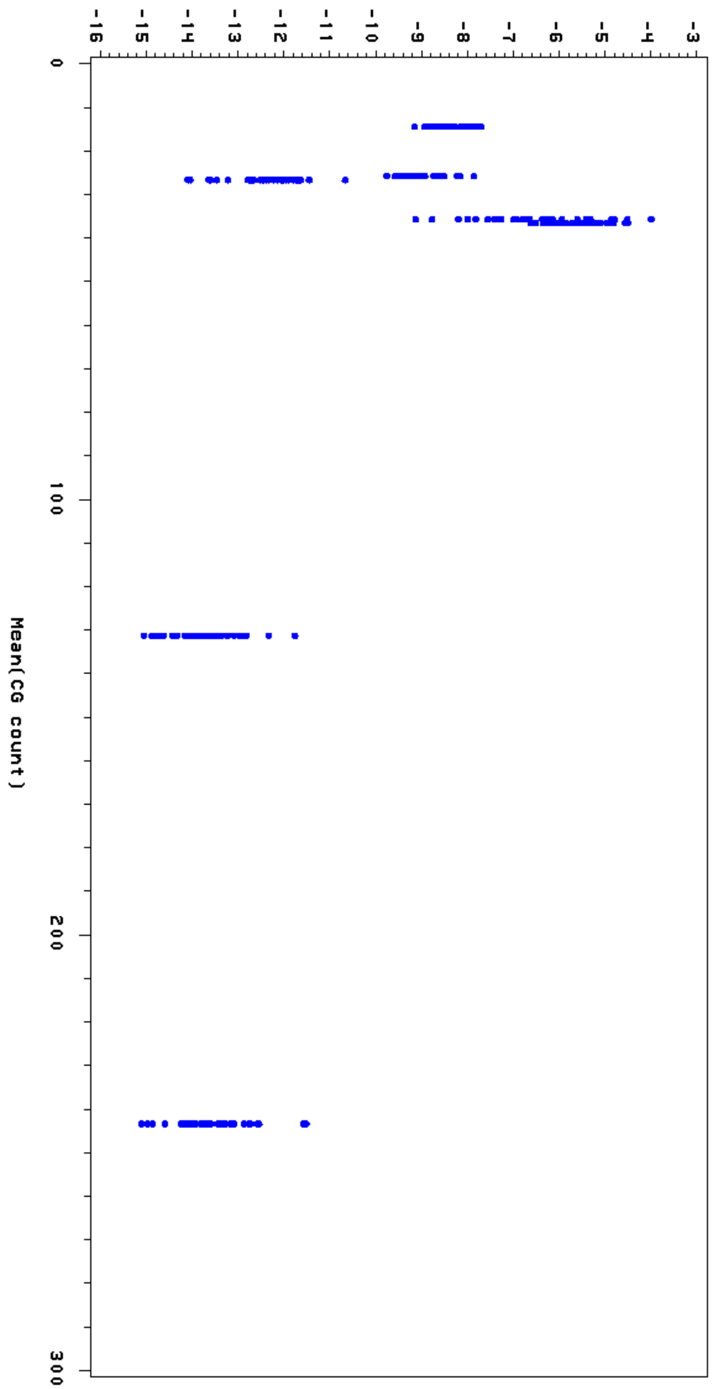

Figure 1: $ln\{\hat{E}_{sim}(\hat{V}_{tg})\}$ vs $\bar{n}_{tg}$

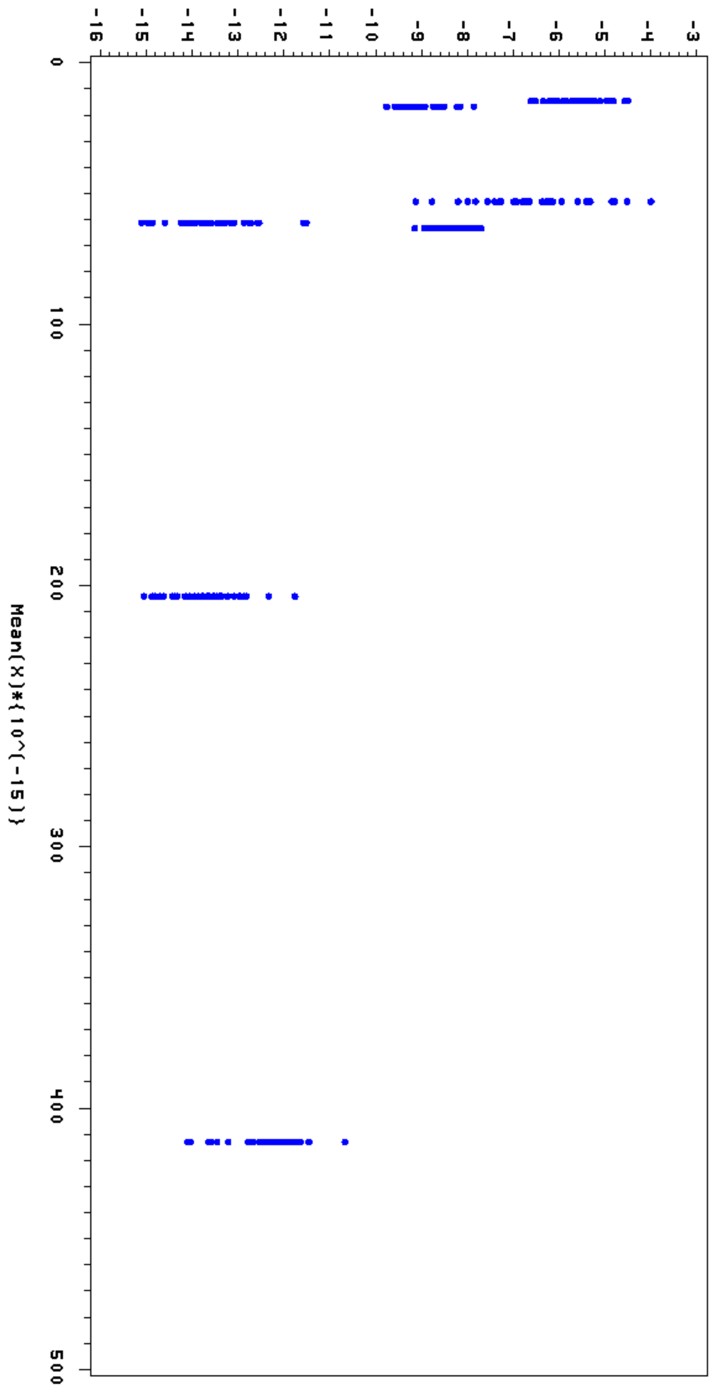

Figure 2: $ln\{\hat{E}_{sim}(\hat{V}_{tg})\}$ vs $\bar{X}_{tg}10^{-15}$

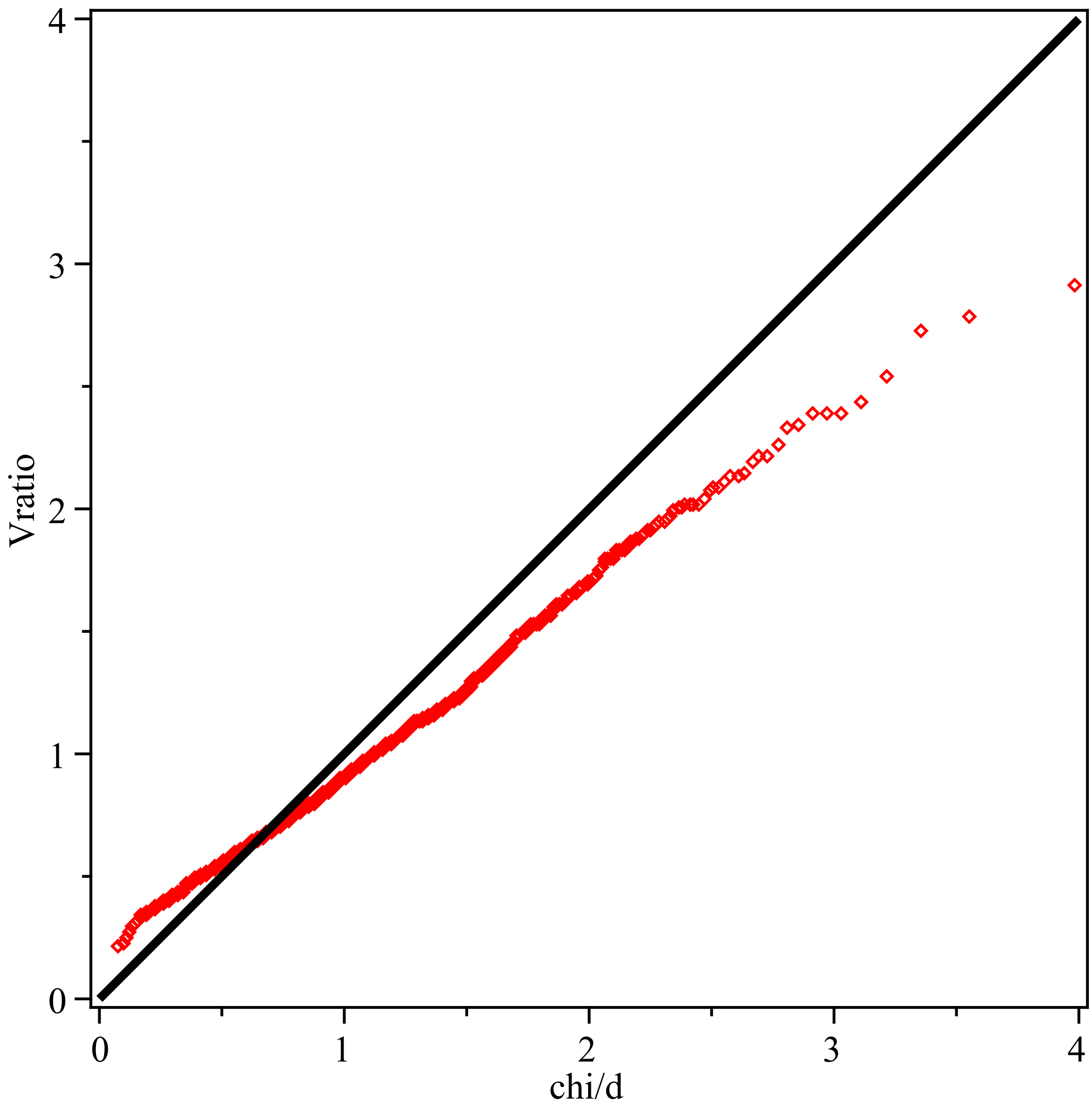

Figure 3: QQ plot of $Vratio$ vs χ_d^2/d (Chapter 90; Month=10; 1000 samples)

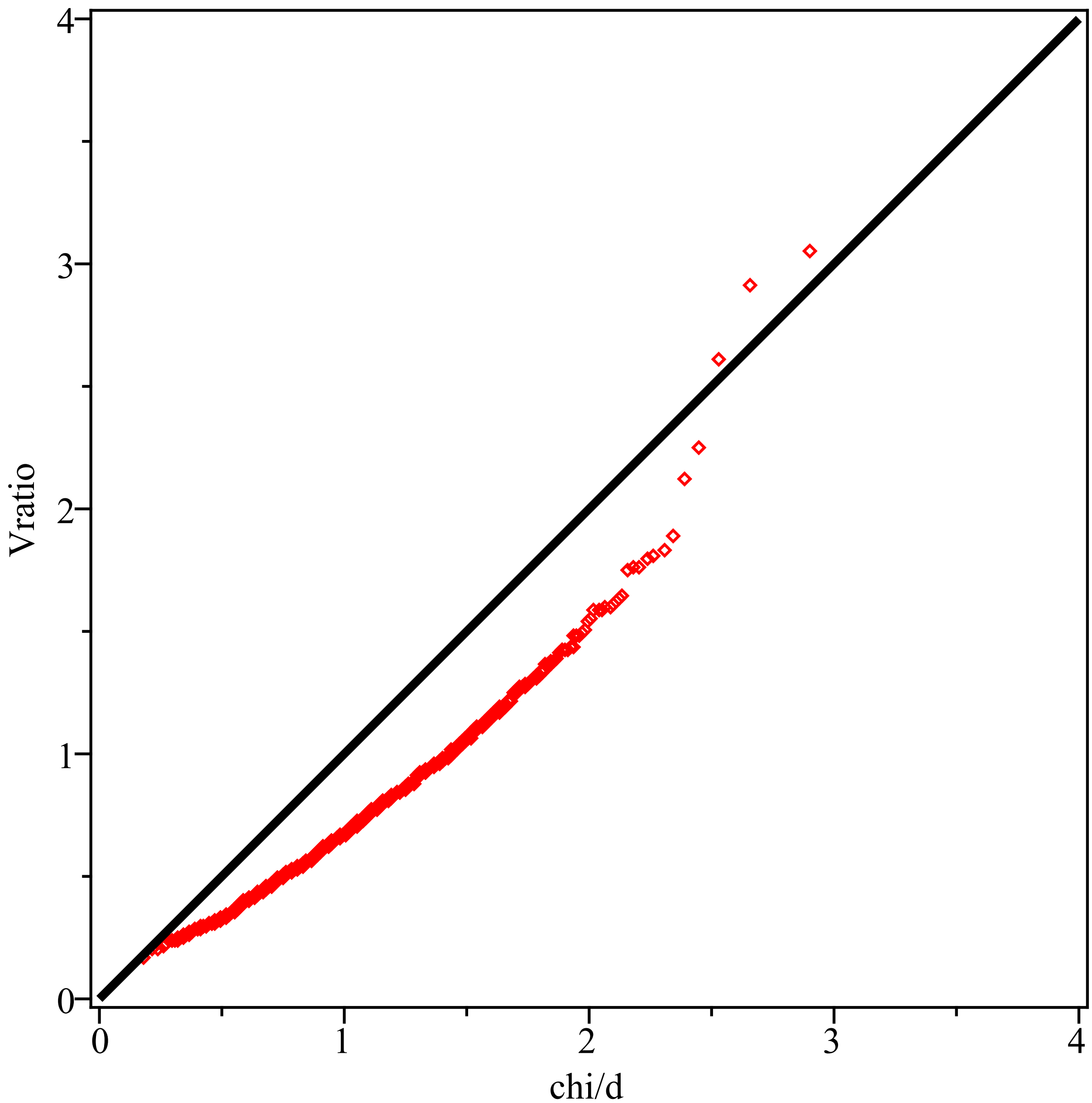

Figure 4: QQ plot of $Vratio$ vs χ_d^2/d (Chapter 07; Month=03; 1000 samples)