INVESTIGATING MODEL-BASED TIME SERIES METHODS TO IMPROVE ESTIMATES FROM MONTHLY VALUE OF CONSTRUCTION PUT-IN-PLACE SURVEYS

Thuy Trang T. Nguyen, William R. Bell, and James M. Gomish, U.S. Census Bureau
Thuy Trang T. Nguyen, U.S. Census Bureau, 4700 Silver Hill Road, Washington, DC 20233-6900

KEY WORDS: Signal extraction, sampling error autocorrelation, sampling error model

1. INTRODUCTION

The *Value of Construction Put in Place (VIP)* is a U.S. Census Bureau publication measuring the value of construction installed or erected at construction sites during a given month. The VIP estimates come from the monthly Construction Progress Reporting Survey (CPRS) augmented with estimates of a non-CPRS component based on regulatory filings, phasing of other Census data, administrative records, and trade association data. In July 2002 the Census Bureau began publishing the monthly VIP for new "types of construction" (TC) categories that reclassified and expanded the previous TC categories. (The latter can be found in U.S. Census Bureau (2002a).) The new TC categories contain many more series and levels of detail than do the old TC categories. This expansion to more levels of detail resulted in relatively small sample sizes and large sampling errors for the direct survey estimates for many categories.

In this paper we investigate the use of time series modeling and signal extraction methods to borrow information over time for improving the VIP estimates. Scott and Smith (1974) and Scott, Smith, and Jones (1977) proposed use of time series techniques to improve estimates in repeated surveys. More recent work in this area includes papers by Bell and Hillmer (1990), Binder and Dick (1989,1990), and Pfeffermann (1991). The approach requires the development of time series models for the sampling errors in the direct estimates as well as for the true underlying series being estimated. Here we develop such models for 70 VIP time series from a subset of the TC categories that refer to privately owned nonresidential construction. All these series start in January 1993 and end in December 2000, and are estimated entirely from the CPRS. These direct estimates have last-year (year 2000) average coefficient of variation (CV) ranging from 3% to 27%. Table 1 gives a complete list of the TCs and their last-month, last-year average, and last-four-year average CVs. We also perform signal extraction with the fitted models to examine the potential for variance reduction in the estimates by borrowing information over time through the models. Sections 2 through 4 develop these models for the VIP series, and Section 5 presents the signal extraction results.

We view the results presented here as preliminary due to some significant data limitations: a) the direct VIP estimates used in this study were adjusted with undercoverage and late selection factors, whereas the estimates of sampling variances and autocovariances that we had available did not reflect these adjustments, b) the length of the time series we analyzed is eight years, and we had estimates of sampling variances and autocovariances only for the last four years. The first of these limitations means that we are simply forced to assume that the sampling variances and autocovariances we had available are at least approximately valid for the time series of estimates we are analyzing. Further analysis might assess this assumption or partially address this limitation. The second limitation is perhaps more directly relevant to the results presented here. The limited amount of data available means that we had difficulty determining the appropriateness of the models we developed, and that even if we assume we chose models of appropriate form, we remain quite uncertain about the model parameters. This is important because the signal extraction results as computed are optimistic in that they assume the correct model is being used, including use of the true values of the model parameters.

2. MODELING OF THE SAMPLING VARIANCES AND AUTOCORRELATIONS

Given concerns about the high level of sampling error in the VIP point estimates there is reason for concern also about the level of sampling error in their corresponding estimates of sampling variances, autocovariances, and autocorrelations. To reduce the level of statistical uncertainty in these estimates we take the raw (direct) sampling variance and autocorrelation estimates and model them. Our philosophy is that, if the direct survey point estimates need to be improved via modeling, then so too do the direct survey variance and autocorrelation estimates.

Unfortunately, sampling variance and autocovariance estimates are not available prior to January 1997, i.e., we have no sampling variance and autocovariance estimates for the first four years of our observed time series (1993-1996). Because of the sample size and frame changes in January 1997 (Cartwright 1996; Mesenbourg 1997), this means we also lack information to relate pre-1997 sampling variances to post-1997 sampling variances, and similarly for autocovariances. Thus, here and in Section 3 we develop sampling error models based only on the post-1997 data. Use of these models for the full length of the observed time series in Section 4 necessitates some "heroic" assumptions as noted there.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

2.1 Modeling of the Sampling Variances

The sampling variance $Var_p(y_t)$ of the direct VIP survey estimate y_t is estimated using the stratified jackknife method. This is done using the VPLX program (Fay 1998). Sampling variances are expected to depend on sample size and possibly also on the VIP levels. Various alternative models for this dependence were thus fitted to the direct sampling variance estimates and compared empirically.

The first stage in modeling the variances was defining n_t, the sample size at time t, for each of the 70 TCs. The definition of n_t is not so obvious for the VIP estimates. We examined different definitions of sample size n_t. The alternative definitions differ in regard to the extent to which they count certainty cases (that is, having a sampling rate of 1-in-1) and how they treat the cases that were originally sampled as belonging to another TC (not the TC under consideration) but later discovered to belong to this TC. The latter cases were not part of the planned sample for the given TC, though they currently are part of the sample. One can argue that certainty cases do not contribute to sampling variability and thus should not be counted towards sample size. (This argument would be more compelling if the estimate were broken into pieces from the certainty and noncertainty portions of the sample, and the variance estimates were used for modeling the separate noncertainty portion.). It is possible that the choice among the alternative definitions of n_t will make little difference to the variance modeling for a given TC. To check this we computed 5×5 correlation matrices between the alternatively defined n_t's for each of the 70 TCs. The correlation coefficients exceeded .90 for almost all pairs of alternative n_t's for all TCs. This high correlation suggests that choice of a particular n_t is unlikely to appreciably affect the fit of the variance model. Our tentative choice was the definition that removed from the sample size count of the certainty cases. For those series where the correlation coefficients for the chosen n_t definition and the other n_t definitions were less than .90 we compared fits of variance models (discussed below) using these alternative definitions. The models of the chosen n_t definition had smaller AICs than models for the other n_t definitions, and so this definition remained the preferred choice.

A natural way to account for possible dependence of sampling variances on the level of the VIP estimates is to model relative sampling variances rather than directly model the sampling variances. The relative sampling variance of y_t is defined as $RelVar_p(y_t) = Var_p(y_t) / (Y_t)^2$, where Y_t is again the underlying population quantity estimated by y_t (true VIP for the TC). From a Taylor series linearization, $RelVar_p(y_t)$ is approximately $Var_p[\log(y_t)]$, a property that will be relevant to the time series modeling of Section 4. Since Y_t is unknown, we use $Rel\hat{V}ar_p(y_t) = \hat{V}ar_p(y_t) / (y_t)^2$ to estimate the relative sampling variances.

To investigate alternative possibilities for the dependence of sampling variances on level and sample size, we fitted the following generalized variance function (GVF) models by linear regression (Wolter (1985, ch. 5) discusses GVFs.): 1) $Rel\hat{V}ar_p(y_t) = b_1 / n_t + error$. 2) $Rel\hat{V}ar_p(y_t) = b_0 + b_1 / n_t + error$. 3) $Rel\hat{V}ar_p(y_t) = b_0 + error$. 4) $\hat{V}ar_p(y_t) = b_1 / n_t + error$. 5) $\hat{V}ar_p(y_t) = b_0 + b_1 / n_t + error$. 6) $\hat{V}ar_p(y_t) = b_0 + error$. 7) $\hat{V}ar_p(y_t) = b_1 n_t + error$. 8) $\hat{V}ar_p(y_t) = b_0 + b_1 n_t + error$. Models 1 to 3 allow for dependence of sampling variances on level through modeling of the relative variances, whereas models 4 to 8 imply no explicit dependence of sampling variances on level. Models 1 and 4 allow for sampling variability to be inversely proportional to sample size, and models 2 and 5 generalize this dependence with an intercept term. Models 3 and 6, however, allow no dependence of sampling variability on sample size. Models 7 and 8, which imply that sampling variances increase with increasing sample size (assuming $b_1 > 0$), require some explanation. Such dependence is possible because sample size increases with the level of construction activity (more active projects in sample), as does the level of the VIP series, and as would the variance of the estimates of VIP.

We examined scatter plots of $Rel\hat{V}ar_p(y_t)$ or $\hat{V}ar_p(y_t)$ versus n_t, with the fitted GVF curves superimposed. While some, not all, of the plots were quite noisy, those that were not suggested that models 4 to 6 are unreasonable, i.e., sampling variances are positively related to the level of the series. The plots also suggested dependence of sampling variances on sample size, eliminating model 3 from consideration. We thus kept four models for further analysis (models 1, 2, 7, and 8) and discarded the others. We computed AICs from these four models using results of the regression fits and assuming the error terms were normal and homoscedastic. ($AIC = m \times log(SSE / m) + 2 \times p$ where m is the number of data points in the fit, p is the number of parameters in the GVF, and SSE is the regression error sum of squares.) Model 2 had the smallest AIC for 62 out of 70 series.

Because the normality assumption for the error terms in models 1 to 8 is questionable given that the data are estimated variances and relative variances, we also tried fitting models for the logs of the variances and relative variances. These analogs to models 1, 2, 7, and 8 are: 1-log) $\log[Rel\hat{V}ar_p(y_t)] = \log(b_1 / n_t) + error$. 2-log) $\log[Rel\hat{V}ar_p(y_t)] = \log(b_0 + b_1 / n_t) + error$. 7-log) $\log[\hat{V}ar_p(y_t)] = \log(b_1 n_t) + error$. 8-log) $\log[\hat{V}ar_p(y_t)] = \log(b_0 + b_1 n_t) + error$. Note that models 1-log and 7-log reduce to linear models that can be fit by linear regression, while models 2-log and 8-log require fitting by nonlinear regression (done using PROC NLIN in SAS (1990)). We compared AICs for the four models, finding that of these models 2-log had the smallest AIC for 64 out of the 70 series. We thus discarded models 7-log and 8-log along with models 7 and 8. We then compared AICs for models 1, 1-log, 2, and 2-log. For the log models we added to the AICs – 2 times the log-Jacobian of the log transformation, which is $-2\Sigma\log|J_t|$ for $t = 1$ to m, where $J_t = \partial\log(\nu_t) / \partial\nu_t = 1 / \nu_t$ and ν_t is the t^{th} observation of the data being modeled: $\nu_t = Rel\hat{V}ar_p(y_t)$ for model 1-log and model 2-log. Model 2-log had the smallest AIC for 63 of the 70 series, and for three of the other TCs model 1-log, a special case of model 2-log, had the lowest AIC. Of the remaining four TCs for which model 2 was preferred by AIC, there was only one TC for which the difference was substantial. To avoid the complexity of using different variance models for a few different TCs, we adopted model 2-log for all 70 TCs.

2.2 Modeling of the Sampling Autocorrelations

For all 70 TCs in this study, we produced estimates of sampling autocovariances and autocorrelations for each pair of months from January 1997 through December 2000. Like the sampling variances, the sampling autocovariances between time t and $t - k$ $C\hat{o}v_p(y_t, y_{t-k})$ were also estimated using the VPLX program with the stratified jackknife method. Follows, the estimated sampling autocorrelations are computed from the estimated sampling autocovariances and variances as $C\hat{o}rr_p(y_t, y_{t-k}) = C\hat{o}v_p(y_t, y_{t-k}) / [V\hat{a}r_p(y_t) V\hat{a}r_p(y_{t-k})]^{.5}$.

Assuming stationarity of the autocorrelations, for each TC we averaged all the estimated autocorrelations for a given lag, that is, averaging 47 estimated lag-1 autocorrelations, 46 estimated lag-2 autocorrelations, etc. We then used the averaged autocorrelations to calculate corresponding partial autocorrelations by solving the Yule-Walker equations of successively higher order (Box and Jenkins 1976, pp. 64-65). Graphs of the resulting autocorrelation function (ACF) and partial autocorrelation function (PACF) were produced and examined for all 70 TCs.

The patterns of the ACF and PACF plots were quite similar across all 70 TC's. The ACF is dominated by an exponential decay (apart from some persistent positive, though small, autocorrelations at higher lags that were not characteristic of the 70 TCs in general). The PACF has a large spike at lag 1 and a much smaller spike at lag 2 (more so for some series than others). Candidate models for such patterns include the first order or second-order autoregressive (AR(1) or AR(2)) model and the mixed ARMA(1,1) model. Again, for simplicity, we wanted to use the same model for all the TCs. The AR(2) seemed to be a suitable choice for this purpose.

3. DEVELOPMENT OF THE SAMPLING ERROR MODEL

In Section 4 we develop models for the time series of the logarithms of the VIP estimates for the 70 TCs, denoting the time series for a given TC by $\log(y_t)$. Here we complete development of the models for the sampling error component e_t of $\log(y_t)$. In Section 2.1 we developed models for $Var(e_t) = Var_p[\log(y_t)]$, noting that from a Taylor series linearization $Var_p[\log(y_t)] \approx RelVar_p[(y_t)]$. In Section 2.2 we noted that the sampling error autocorrelations generally appeared to be well-modeled by an AR(2) model. Putting these two parts of the model together, we have the following general form of the sampling error model: $e_t = h_t\tilde{e}_t$ where h_t is the standard deviation of e_t, i.e., $h_t = [Var(e_t)]^{.5} \approx \{RelVar_p[(y_t)]\}^{.5}$ and $\tilde{e}_t$ has variance one and follows the AR(2) model $(1 - \phi_1 B - \phi_2 B^2)\, \tilde{e}_t = c_t$. B is the backward shift operator and c_t is white noise. So that $\tilde{e}_t$ has variance one we need to set $Var(c_t)$ so that the variance computed from the AR(2) model above is one. From Box and Jenkins (1976, p. 62) this implies that $\sigma_c^2 = Var(c_t) = [(1+\phi_2)/(1-\phi_2)]\times[(1-\phi_2)^2 - \phi_1^2]$. To estimate the parameters ϕ_1 and ϕ_2 we used the averaged sampling autocorrelations developed in Section 2.2 and applied the Yule-Walker equations for the AR(2) model. From Box and Jenkins (1976, p. 60) this gives $\hat{\phi}_1 = [r_1(1 - r_2)] / (1 - r_1^2)$ and $\hat{\phi}_2 = (r_2 - r_1^2) / (1 - r_1^2)$ where r_1 and r_2 are the averaged sampling autocorrelations at lags 1 and 2.

The estimates of h_t come from the fitted variance models developed in Section 2.1. The model chosen there (model 2-log) is fitted to the estimates, $\log[RelV\hat{a}r_p(y_t)]$, which are taken as estimates of $\gamma_t = \log[Var(e_t)]$. Denote the fitted values by $\hat{\gamma}_t = \log(\hat{b}_0 + \hat{b}_1 / n_t)$. We wish to convert these to estimates of $Var(e_t)$. Simple exponentiation is one obvious way to do this, i.e., we set $h_t^2 = \exp(\hat{\gamma}_t)$.

A more involved approach to converting fitted values from the sampling variance model 2-log to estimates of $Var(e_t)$ attempts to correct for bias from the log transformation. We assume that the direct relative variance estimates, $v_t = RelV\hat{a}r_p(y_t)$, are approximately unbiased for $Var(e_t)$. If we knew the parameters b_0 and b_1, and the variance of the error term (say ω^2) in model 2-log, then from properties of the lognormal distribution $E(v_t) = exp(\gamma_t + .5\omega^2) \equiv exp(b_0 + b_1 / n_t + .5\omega^2)$. Since we only have the fitted model we assume the fitted values $\hat{\gamma}_t = \log(\hat{b}_0 + \hat{b}_1 / n_t)$ are approximately normally distributed with means γ_t and variances $Var(\hat{\gamma}_t)$. This means that $\exp(\hat{\gamma}_t)$ is approximately lognormal with $E(e^{\hat{\gamma}_t}) = e^{\gamma + .5Var(\hat{\gamma}_t)}$. Assuming that the residual variance from the fit of model 2-log $\hat{\omega}^2 = \sum[\log(v_t) - \hat{\gamma}_t]^2 / (m-2)$ is a consistent estimate of the true residual variance ω^2, this implies that $E\left(e^{\hat{\gamma}+.5[\hat{\omega}^2 - Var(\hat{\gamma}_t)]}\right) \approx E\left(e^{\gamma+.5\omega^2}\right) = Var(e_t)$ so we set $h_t^2 = \exp\{\hat{\gamma}_t + .5[\hat{\omega}^2 - Var(\hat{\gamma}_t)]\}$.

We tried both of these approaches and found they sometimes gave different results. The differences from the two approaches were greater than 10% for 30 out of 70 time series, but the other 40 series had less than 10% difference. The series with large differences tended to be noisier than the series with less than 10% differences. The sampling error standard error, h_t, that will be used in section 4 is the result of the second approach.

Another problem is that we had sampling variance estimates only from January 1997 to December 2000, and so fitted the sampling variance models using data from this period. Because of the sample design changes in January 1997 (sampling rate changes and sampling frame change) we really have no information to relate sampling variances prior to 1997 to those from 1997 on. Therefore, from January 1993 to December 1996 we simply set the value of h_t to its value for January 1997. This is not a good solution to this problem, but the only other option is to restrict the time series modeling to start in January 1997, which would give us only four years of data. We intend to pursue the second option later, but for now take the first course of filling in the earlier h_t values, keeping in mind that this is a significant limitation to our results.

4. DEVELOPMENT OF MODELS FOR THE TIME SERIES OF THE DIRECT ESTIMATES

The direct log VIP estimate is equal to the true log VIP plus sampling error, $\log(y_t) = \log(Y_t) + e_t$. The model for the observed time series $\log(y_t)$ is determined by the models for the two components $\log(Y_t)$ and e_t; we call such a model a "component model." When the model for $\log(Y_t)$ includes regression terms, we call the model for $\log(y_t)$ a RegComponent model. Given the models for the sampling error components e_t developed in Section 3, and given a specified form for a time series model for $\log(Y_t)$, we can fit the resulting RegComponent model to the observed series $\log(y_t)$ to estimate the unknown parameters of the model for $\log(Y_t)$. In doing so the parameters of the model for the sampling error component e_t are held fixed. The REGCMPNT program developed by the Time Series Staff of the Census Bureau performs this type of model fitting.

As part of exploratory analysis to determine suitable forms for the models for the true time series $\log(Y_t)$, we used the X-12-ARIMA program (U.S. Census Bureau 2002b) to fit some RegARIMA models (regression models with error terms following ARIMA models), ignoring the sampling error components. This allowed us to check for trading-day effects and outliers in the series. Any outliers found were carried over for use in the RegComponent model since the REGCMPNT program does not perform outlier detection. For the ARIMA models we started in all cases with the airline model (Box and Jenkins 1976, ch. 9). In cases where the estimate of the seasonal moving average parameter was close to 1 we cancelled the seasonal difference and noninvertible seasonal MA operator and converted the model to an ARIMA(0,1,1) with fixed seasonal effects and a trend constant.

Having made a preliminary determination of the need for trading-day effects and outliers, the resulting RegComponent models were fitted by the REGCMPNT program for each of the 70 VIP series. We used airline models for $\log(Y_t)$ except when the ARIMA model fitting results suggested fixed seasonality. The REGCMPNT fitting results were examined and changes were made to the models when they exhibited any of the following properties: 1) if an outlier (included in the RegComponent model using the appropriate regression variable) had a t-statistic less than 3.8 (the critical value used in the X-12-ARIMA outlier detection), the outlier was dropped from the model. 2) If the model included trading-day effects but the chi- squared statistic testing the significance of the trading-day effects was insignificant at the .05 level then the trading-day effects were dropped from the model. 3) If the model included fixed seasonal effects and the fixed seasonal p-value was extremely large (p-value > .45; note that these p-values tended to be either very large or less than .05), then the fixed seasonal effects would be dropped from the model, leaving a nonseasonal model. 4) If the estimate of the nonseasonal MA parameter was near 1, the nonseasonal difference was cancelled with the nonseasonal MA operator and a trend constant was added to the model. 5) If the estimate of the seasonal MA parameter was now near 1, the seasonal difference was cancelled and a fixed seasonal and a trend constant were included.

We continued to modify models as needed until the results seemed reasonable or simply the best that we could do. We did not feel the need to stray from the airline model since in the cases where the RegComponent model seemed not to fit well the situation generally was not improved much by changing the airline model to some other model. Basically, some of the series were just quite noisy and difficult to model.

Given that we lacked sampling variance estimates for the first four years of our series our sampling error models are questionable for this period. Motivated by this, we shortened the VIP series to the four years starting in January 1997 for which we did have sampling variance estimates, and tried fitting the RegComponent models to these shortened series. Unfortunately, we were generally unsuccessful in modeling these extremely short series.

5. APPLICATION OF SIGNAL EXTRACTION RESULTS TO INVESTIGATE POTENTIAL FOR IMPROVING ESTIMATES OF THE TRUE VIP SERIES

The REGCMPNT program produces finite sample signal extraction estimates of the component series along with signal extraction error variances for these estimates. We denote the signal extraction estimates for the log VIP series by $\log(\hat{Y}_t)$. Our interest here is primarily in the signal extraction error variances, denoted $\mathrm{Var}[\log(Y_t) - \log(\hat{Y}_t)]$. The square roots of these error variances can be interpreted in percentage terms, analogous to CVs. When compared to the original sampling error CVs for the direct VIP estimates, this provides a measure of the improvement from signal extraction.

The signal extraction results from REGCMPNT assume that the correct model is used. In particular, no allowance is made in the signal extraction error variances to account for uncertainty due to using estimated model parameters. With reasonably long time series the consequence of this is generally some amount of understatement of the signal extraction error variance. With the limitations of our modeling (very short time series, some series have high levels of sampling error, no sampling variances prior to 1997) we are quite uncertain about the true values of our model parameters. This raises the possibility that the signal extraction variances we examine here significantly understate the true error variances. However, overstatement of variances could also occur if the innovations variance in the model for $\log(Y_t)$ is overestimated. The bottom line here is that, due to the significant amount of uncertainty about our model parameters, the signal extraction variances generally provide at best rough indications of potential for improvement from signal extraction. Results for single series should not be taken too seriously, particularly for those series with high levels of sampling error. Results considered over all 70 VIP series probably provide better general indications of potential for improvement. The results should not be taken as precise quantifications of the potential improvement.

Table 1. Sampling Coefficient of Variation (CV) and Percentage Improvements in CV from Signal Extraction for the Last-month Average, Last-year Average, and Last-four-year Average

	Last Month		Last Year Average		Last Four-Year Average	
Types of Construction	Sampling CV (%)	(%) CV Improve-ment	Sampling CV (%)	(%) CV Improve-ment	Sampling CV (%)	(%) CV Improve-ment
Lodging	5	3	5	4	5	5
Office	4	13	4	16	5	27
Commercial	3	30	3	30	4	43
Health Care	5	32	4	33	5	46
Educational	5	14	5	18	6	24
Religious	5	34	5	34	7	50
Public Safety	9	1	8	2	8	3
Amuse & Rec	5	8	5	12	6	17
Transportation	5	6	6	9	8	11
Sewer & WstDisp	23	14	27	17	22	18
Water Sup Sys	28	31	26	40	26	46
Manufacturing	3	4	3	7	4	8
Food/Bev/Tobac	10	10	10	14	12	17
Textile/App//Leath	14	6	12	8	17	10
Wood	10	3	15	6	16	8
Furniture	12	10	24	14	24	18
Paper Products	16	27	17	30	24	37
Print/Publishing	17	5	15	7	15	9
Chemical/Allied	7	2	8	4	8	5
Petroleum/Coal	7	6	17	10	12	12
Rubber/Plastics	9	2	10	2	8	3
Stone/Clay/Glass	6	3	8	3	13	9
Primary Metal	8	2	5	2	7	3
Fabricated Metal	9	3	11	4	11	5
Machinery/Non-el	17	5	19	9	17	12
Computer/Elect/El	7	2	7	3	7	3
Transportation	8	15	9	18	13	24
Miscellaneous	8	3	8	5	12	7
Financial	12	52	11	46	14	53
Automotive	8	27	9	30	12	40
Food/Beverage	9	38	9	34	12	49
Multi-Retail	6	9	5	13	6	18
Other Commercial	8	25	8	24	13	45
Warehouse	4	13	5	17	6	22
Hospital	7	28	5	31	7	45
Medical Building	8	13	9	18	9	21
Special Care	9	49	9	45	11	49
Preschool	22	28	24	31	33	50
Primary & Second	10	17	11	21	13	27
Higher Education	6	28	15	27	9	38
Other Educational	18	9	7	15	16	19
House of Worship	6	35	6	36	8	50
Other Religious	9	45	9	43	12	50
Theme/Amusemnt	10	3	13	5	16	6
Sports	15	6	10	10	12	11
Fitness	15	30	14	31	18	43
Perform/mtCnter	10	4	12	5	14	7
Social Places	13	30	12	32	16	42
MovieTheatr & Stud	9	3	10	5	9	7
Air	4	2	4	3	7	4
Land	17	4	16	6	16	8
General Offices	4	12	4	15	6	26
Auto Sales	16	19	17	24	20	36
Auto Service/Parts	13	24	13	26	18	38
Parking	8	6	14	10	18	12
Food	11	36	12	31	15	47
Dining/Drinking	18	39	16	39	18	50
Fast Food	20	35	26	41	29	51
General Merchand	5	16	9	20	10	28
Shopping Center	9	6	7	11	8	16
Shopping Mall	21	12	12	16	10	19
Other Stores	13	24	12	26	19	43
Drug Stores	18	23	9	25	22	40
Building Supplies	9	6	16	10	14	20
General Warehses	4	10	5	18	6	28
Instructional	8	29	9	30	11	40
Dormitory	14	45	13	41	21	47
Sport/Rec Facility	19	12	21	17	22	20
Gallery/Museum	22	10	16	17	17	22
Auxiliary Buildings	13	12	13	17	16	21

Keeping these limitations in mind, Table 1 presents results derived from the signal extraction variances produced by REGCMPNT. The table shows the estimated improvements from signal extraction over the sampling CVs of the direct estimates. To summarize the results, the percent improvements are shown in the CVs of the last month estimate, along with the average percent improvements over the last year of estimates and the last four years of estimates. The percent improvements shown are multiplicative percent improvements on the sampling CVs expressed as percents. Thus, if the sampling CV was 20% and the improvement was 25%, then the signal extraction CV was $(1 - .25)\times 20\% = 15\%$.

Table 1 indicates that there was a wide range of estimated improvements. Note the average improvement in the CVs over the last year ranged from 2% (public safety TC) to 46% (financial TC). For those TCs whose last-year average sampling error CV is less than 10%, the average percent improvement from signal extraction ranged between 2% and 45%. For TCs whose last-year average sampling error CV is greater than 10%, the average percent improvement ranged between 4% and 46%. This showed potential for improvements both for TCs with relatively small sampling CVs as well as for TCs with relatively large sampling CVs. However, improvements in the accuracy of estimates for TCs whose sampling CVs are already quite low (say < 5%) may not be of much interest.

6. CONCLUSIONS

The results presented provide rough indications of potential for improvement of the VIP estimates through time series modeling and signal extraction. Assessment of the actual improvements that can be realized, however, is made difficult by the significant data limitations (short series with no sampling variance estimates for the first four years). These limitations leave us with considerable uncertainty about the parameters of our models, and this affects the validity of the signal extraction results. (The signal extraction variances can be thought of as estimates, here fairly imprecise estimates, of the true variances of the errors in the signal extraction estimates.) The high level of sampling error in some of the series is another limitation on the modeling results as it too contributes to uncertainty about model parameters. Series with high levels of sampling error are the most interesting in regard to potentially improving on the accuracy of the direct survey estimates. Unfortunately, high levels of sampling error make series more difficult to model.

In the future we hope to do additional work to at least partially address some of the limitations of this study. First, we will soon have available one additional year of VIP estimates to extend our time series, and can also generate corresponding sampling variance estimates. The resulting series will still be rather short, but not quite so short as before. Second, we intend to pursue a Bayesian approach to inference with our models to recognize uncertainty about the model parameters. (This will most likely recognize uncertainty only about the parameters in the models for $\log(Y_t)$, taking the fitted sampling error models as given, but it is the uncertainty about the parameters in the models for $\log(Y_t)$ that is of most concern.) The goal here is not really to reduce the uncertainty, but simply to account for it in the signal extraction results. Finally, if we are able to achieve satisfactory results with the Bayesian approach, we intend to use the models to investigate model-based seasonal adjustment of the VIP series, and the potential for the use of the models to improve seasonal adjustment results. Again, this will probably need to be done with a Bayesian approach.

ACKNOWLEDEMENTS

We would like to thank Brian Monsell of the Statistical Research Division for his assistance in writing routines to compute the partial autocorrelations with the autocorrelations as input. We thank Bob Fay of the Director's Office for his assistance with VPLX. We also thank Masato Asanuma of the Manufacturing and Construction Division for providing the information regarding the survey design.

REFERENCES

Bell, William R. and S. C. Hillmer (1990), "The Time Series Approach to Estimation for Repeated Surveys," *Survey Methodology*, **16**, 195-215.

Binder, David A. and Peter J. Dick (1989), "Modeling and Estimation for Repeated Surveys," *Survey Methodology*, **15**, 29-45.

---------- (1990), "A Method for the Analysis of Seasonal ARIMA Models," *Survey Methodology*, **16**, 239-253.

Box, George E. and Gwilym M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, California: Holden-Day Inc.

Cartwright, David C. (1996), "Construction Progress Reporting Surveys (CPRS) System Redesign – Revised Sampling Rates", MCD Memorandum, April 29, 1996.

Fay, Robert E. (1998), VPLX Program Documentation, Washington DC: U.S. Census Bureau.

Mesenbourg, Thomas L. (1997), "Construction Progress Reporting Surveys System Redesign – Revised Sampling Rates", MCD Memorandum, July 16, 1997.

Pfeffermann, Danny (1991), "Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys," *Journal of Business and Economic Statistics*, **9**, 163-175.

SAS Institute, Inc. (1990), *SAS® Language: Reference, Version 6, First Edition*, Cary, NC: SAS Institute, Inc.

Scott, Alastair J. and T.M.F. Smith (1974), "Analysis of Repeated Surveys Using Time Series Methods," *Journal of the American Statistical Association*, **69**, 674-678.

Scott, Alastair J., T.M.F. Smith, and Roger G. Jones (1977), "The Application of Time Series Methods to the Analysis of Repeated Surveys," *International Statistical Review*, **45**, 13-28.

U.S. Census Bureau (2002a), "Value of Construction Put in Place", (Report No. C30/02-5 (Annual)), Washington, DC: U.S. Department of Commerce.

---------- (2002b), *X-12-ARIMA Reference Manual, Final Version 0.2*, Washington, DC: U.S. Census Bureau.

Wolter, Kirk M. (1985), *Introduction to Variance Estimation*, New York: Springer-Verlag Inc.