Imputation Methods for the Current Employment Statistics Survey

Kennon R. Copeland[1] and Lan Pham[2]

[1]NORC at the University of Chicago, [2]U.S. Bureau of Labor Statistics
[1]4350 East-West Hwy, Ste 800, Bethesda, MD 20814, (Copeland-Kennon@norc.org)
[2]2 Massachusetts Avenue, N.E. Washington, DC 20212 (Pham.Lan@bls.gov)

I. Introduction

The Bureau of Labor Statistics' (BLS) Current Employment Statistics (CES) survey collects employment, hours, and earnings data monthly from a sample of over 300,000 U.S. establishments. To provide timely information, initial estimates are generated three to four weeks after the survey reference period. Final estimates, incorporating late reports received after production of the preliminary estimates, are released two months later. Benchmark estimates, incorporating administrative population data from the BLS' ES-202 program for March of the prior year, are released annually with the data for May.

Nonresponse potentially introduces bias into survey estimates, if respondents differ from nonrespondents relative to the variables of interest, and also reduces the effective sample size of a survey, thereby increasing variances for survey estimates. Estimation methods are developed so as to account for nonresponse and lessen its impact on bias and variance. These methods, however, assume nonresponse is ignorable within defined estimation cells and, hence, do not distinguish among various patterns of nonresponse. Nonresponse is used here to encompass both nonreporting and late reporting. Late reporting is temporal nonresponse, as the data become available at a subsequent point in time.

The objective of this research is to identify imputation methods that yield decreased revisions in month-to-month change and total employment estimates. Alternative methods are proposed, and findings are presented relative to criteria under which the alternative methods should be applied and performance of the resulting employment estimates relative to the current method.

This paper builds upon previous research into CES late reporting and nonresponse imputation reported in Copeland (2003a) Copeland (2003b), Copeland (2004a), Copeland (2004b), and Copeland and Valliant, 2007). The CES imputation research conducted in this project was divided into two phases – model definition (Phase I) followed by model performance assessment (Phase II). Section II presents a brief overview of the CES survey design and estimation methodology, Section III defines and profiles the issue associated with CES late reporting and nonresponse, Section IV discusses model definition, Section V discusses model refinement and assessment, and Section VI provides a summary and next steps.

II. CES Survey Design and Estimation

The BLS completed a major redesign of the CES survey in 2003 (Werking, 1997; Bureau of Labor Statistics, 2003), moving the survey from its historical quota sample design to a probability sampling basis. The new sample design is a stratified, simple random sample of establishments from the BLS's Longitudinal Data Base, with strata defined by state, industry, and employment size. Sampling rates for each stratum are determined through optimum allocation to minimize variance of total employment level.

Data must be reported within a two to three week period for inclusion in the initial published estimates (referred to as first closing estimates) for the month. As additional responses are received after this first closing of the collection period, the estimates for a given month are revised twice (referred to as second and third closing estimates) to incorporate data from late reporters. The first closing estimate of month-to-month change is derived by subtracting the prior month's second closing estimate from the current month's first closing estimate.

Estimates are generated through use of a weighted link-relative estimator, which uses a weighted sample trend within an estimation cell, based upon common reporters between the prior and current months, to move forward the prior month's estimate for that cell. Units determined to be atypical are treated separately. The current estimator for total employment (Bureau of Labor Statistics, 2001) takes the following form for month t and closing $k\,(=1,2,3)$

$$\hat{Y}_{t|k} = \sum_{c=1}^{C}\left[\frac{\sum_{i\in s_{ct,(t-1)|k},R} w_i Y_{ti}}{\sum_{i\in s_{ct,(t-1)|k},R} w_i Y_{(t-1)i}}\left(\hat{Y}_{c(t-1)|(k+1)} - \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{(t-1)i}\right)\right] + \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{ti} =$$

$$\sum_{c=1}^{C}\left[LR_{ct,(t-1)|k}\left(\hat{Y}_{c(t-1)|(k+1)} - \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{(t-1)i}\right)\right] + \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{ti}$$

where

$c\,(=1,\ldots,C)$ refers to estimation cell (defined by industry and, for selected industries, region)

$s_{ct,(t-1)|k}$ represents the set of sample establishments in estimation cell c that, as of closing k, reported data for both months t and $t-1$

$Atyp$ represents the set of sample establishments in $s_{ct,(t-1)|k}$ determined to be Atypical

R represents the set of reporting sample establishments in $s_{ct,(t-1)|k}$ not classified as Atypical

w_i is the sampling weight for sample establishment i

Y_{ti} is the total employment reported for month t by sample establishment i

$\hat{Y}_{c(t-1)|(k+1)}$ represents the prior month, $t-1$, link-relative estimate for estimation cell c based upon data reported as of closing $k+1$ (with a maximum value of three) for month $t-1$ (which corresponds to closing k for month t)

The link-relative, $LR_{ct,(t-1)}$, is thus a growth rate estimate for the period $t-1$ to t. Differences in estimates between closings will be due solely to the inclusion of late responses, while differences between estimated and benchmark values will be due to the combined effects of sampling, nonresponse, late reporting (if comparing first or second closing estimates), and measurement error. An implicit assumption of the estimator is that, within an estimation cell and closing, establishments not reporting data (nonsampled, nonreporting, late reporting) for both months t and $t-1$ are assumed to have the same growth rate as for those establishments reporting data, i.e., the sampling, response, and reporting timeliness mechanisms are ignorable (Rubin, 1976).

III. CES Late Reporting/Nonresponse

Survey nonresponse is frequently classified on the basis of reason for nonresponse. Panel surveys add another dimension to the response mechanism, that being response status at different points in time. Little and David (1983) distinguished three types of panel survey nonresponse – attrition (sample unit stops reporting), late entry (sample unit does not report initially), and reentry (sample unit has a gap in reporting).

Copeland (2003b) developed the following alternative classification, appropriate for the CES survey, reflecting the current month's reporting status, and timing of reporting for current month reporters:

- Reporters
  - Early Reporters
    - First Closing Reporters – sample establishments reporting data for the month prior to the cutoff date for processing first closing estimates
  - Late Reporters
    - Second Closing Reporters – sample establishments reporting data for the month after the cutoff date for processing first closing estimates, but prior to the cutoff date for processing second closing estimates
    - Third Closing Reporters – sample establishments reporting data for the month after the cutoff date for processing second closing estimates, but prior to the cutoff date for processing third closing estimates
- Nonreporters – sample establishments not reporting data for the month
  - Complete Nonreporters – sample establishments not reporting for any month

Attritors – sample establishments reporting data for at least one month, but which no longer report data

Episodic Nonreporters – sample establishments not reporting data for the month, but which do report for a subsequent month

All three nonreporter types impact the overall accuracy of the CES estimates, regardless of closing. Late reporters (second closing reporters, third closing reporters) affect the accuracy of preliminary estimates only. The extent of the impact late reporters have on the preliminary estimates can be assessed by examining the direction and magnitude of revisions between first and third closing estimates.

IV. Imputation Research Phase I: Model Definition

Phase I of the research was intended to identify characteristics for consideration in imputation and to develop alternative imputation models for broader assessment in Phase II.

A. Consideration of Alternative Models

Consideration of alternatives was restricted to several simple models that would be operationally efficient to implement. All assumed prior history for a particular establishment provided useful ancillary information for estimating current month employment. The following models were considered:

Model 1: Year-lagged-establishment: Establishment growth rate for current month is consistent with that seen for the same month in the prior year

$$LR_{ti} = LR_{(t-12)i} \Rightarrow \hat{Y}_{ti} = Y_{(t-1)i} \times \frac{Y_{(t-12)i}}{Y_{(t-13)i}}$$

Model 2: Month-lagged establishment: Establishment growth rate for the current month is consistent with that seen for the immediately prior month

$$LR_{ti} = LR_{(t-1)i} \Rightarrow \hat{Y}_{ti} = Y_{(t-1)i} \times \frac{Y_{(t-1)i}}{Y_{(t-2)i}}$$

Model 3: Year-lagged deviation: Deviation of establishment growth rate for the establishment from the corresponding cell growth rate for the current month is consistent with that seen for the same month in the prior year

$$LR_{tci} - LR_{tc} = LR_{(t-12)ci} - LR_{(t-12)c} \Rightarrow \hat{Y}_{tci} = Y_{(t-1)ci} \times \left[LR_{tc} + \left(\frac{Y_{(t-12)ci}}{Y_{(t-13)ci}} - LR_{(t-12)c} \right) \right]$$

B. Model Performance Assessment

The performance of each of these models was compared to the performance of the current model (the Link Relative, or Model 0)

$$LR_{tci} = LR_{tc} \Rightarrow \hat{Y}_{tci} = Y_{(t-1)ci} \times LR_{tc}$$

First, a subset of the CES sample establishments was identified for use in assessing model performance. This subset (referred to here as "Influencers") consisted of establishments with a weighted employment greater than or equal to 0.1% (or >[5/{# of reporting establishments}], if this latter value is greater than 0.1%) of the total weighted employment for the supersector in at least one month between Jan '04 and May '08[1]. All data reported for the Influencer establishments was used in the analysis, resulting in 42,596 establishment-months, distributed by supersector and size as indicated in Table 4.

[1] This definition provided for not only designation of large establishments (e.g., most establishments with 500+ reported employment in at least one month), but also relatively larger establishments in smaller cells. Use of a criteria taking into account number of reporting establishments was used to avoid automatically designating Influencer establishments in cases where there were a small number of reporting establishments, in which case it would be expected that at least one establishment would account for >0.1% of the total weighted employment for a supersector.

Table 4

CES Sample Influencers

Supersector	# of Establishment-Months								
		Size (Based on prior month employement)							
	Total	10,000+	5,000-9,999	2,500-4,999	1,000-2,499	500-999	250-499	100-249	<100
Total	42,596	7,086	4,882	5,412	7,642	4,970	4,102	4,011	4,491
Business	1,832	292	353	130	226	215	99	243	274
Construction	701		1	8	63	112	134	202	181
Education	1,849	1,068	333	124	83	82	79	26	54
Finance	3,805	195	733	1,036	374	401	400	357	309
Government	3,791	3,642	102	10	27		10		
Information	7,715	62	178	724	2,786	1,024	909	1,039	993
Leisure	1,301	69	550	198	210	112	28	77	57
Durable Goods Mfg	2,820	504	1,226	694	88	179	47	28	54
Mining	1,870	9	25	68	579	272	185	313	419
Non-durable Goods Mfg	2,957	97	290	731	254	474	557	401	153
Other Services	2,375	3		242	395	556	341	300	538
Retail Trade	1,559	625	444	18	153	91	109	32	87
Transportation	5,074	444	497	1,058	1,025	785	465	280	520
Utilities	2,418	67	72	258	1,270	254	194	163	140
Wholesale Trade	2,529	9	78	113	109	413	545	550	712

Next, performance of the current and three alternative models was determined using both an establishment level assessment and an employment estimate level assessment.

The establishment level assessment was derived by measuring the error in the predicted value for the establishment determined by the model:

$$Err\left(\hat{Y}_{ti}^{M}\right) = \hat{Y}_{ti}^{M} - Y_{ti}$$

where

$\hat{Y}_{ti}^{M}$ = predicted value for establishment *i* in month *t*, as determined by model *M*

Y_{ti} = reported value for establishment *i* in month *t*

Several summary statistics were examined at the supersector, size, and supersector x size level:

Median Absolute Error

$$MedAbsErr(M) = \underset{t,i}{median}\left|Err\left(\hat{Y}_{ti}^{M}\right)\right|$$

Error IQR

$$ErrIQR(M) = \underset{t,i}{75thPercentile}\left(Err\left(\hat{Y}_{ti}^{M}\right)\right) - \underset{t,i}{25thPercentile}\left(Err\left(\hat{Y}_{ti}^{M}\right)\right)$$

Error 90/10

$$Err90/10(M) = \underset{t,i}{90thPercentile}\left(Err\left(\hat{Y}_{ti}^{M}\right)\right) - \underset{t,i}{10thPercentile}\left(Err\left(\hat{Y}_{ti}^{M}\right)\right)$$

The employment estimate level assessment was derived by measuring the error in the 1st closing employment estimate at the supersector level determined by the model, relative to the 3rd closing employment estimate:

$$Err\left(\hat{Y}_{SS,t}^{M}\right) = \hat{Y}_{SS,t}^{M(1)} - \hat{Y}_{SS,t}^{(3)}$$

where

$\hat{Y}_{SS,t}^{M(1)}$ = 1st closing estimate for supersector *SS* in month *t*, using imputed employment as determined by model *M* for late reporting Influencer establishments (note, current model uses only 1st closing reporters, not imputed values for late reporting Influencer establishments)

$\hat{Y}^{(3)}_{SS,t}$ = 3rd closing estimate for supersector *SS* in month *t*

C. Results

Results were examined from both an establishment level and an establishment estimation level perspective. NOTE: All referenced tables in the Establishment Level Assessment section are provided at the end of this document, rather than with the text, due to the size of the tables.

Establishment Level Assessment. Looking at the summary statistics at the supersector level (Table 5), it can be seen that the current model yields the best performance measures (highlighted in yellow in the table) for all three summary statistics for twelve of the fifteen supersectors. For Education and Government, however, both Models 1 and 3 yield better performance measures for all three summary statistics than does the current model. For Retail Trade, Model 1 yields better Median Absolute Error and Error IQR performance measures than does the current model, while the current model yields the best Error 90/10 performance measure.

Looking at the summary statistics at the size level (Table 6), it can be seen that the current model yields the best performance measures (highlighted in yellow in the table) for all three summary statistics for all but the largest size category. For the 10,000+ size category, however, both Models 1 and 3 yield better performance measures for all three summary statistics than does the current model.

It is by examining summary statistics at the size x supersector level that the relationship between size and supersector can be seen. Detailed information analogous to that provided in Tables 5 and 6 were prepared in a Table 7 not presented her due to size, and frequencies of the best performance measure were tallied separately by size and supersector and are provided in Tables 8, 8a, 9, and 9a. Tables 8 and 9 provide frequencies based on all Models, whereas Tables 8a and 9a provide frequencies based solely on Model 1, which is the model which has the best performance.

As can be seen in Tables 8 and 8a, Model 1 yields better performance measures than the current model in roughly half the supersectors for establishments of size 5,000+. For establishments of size less than 2,500, the current model tends to yield best performance across supersectors, with the exception of Construction (where Model 2 yields better performance measures for establishments of size 500-2,499) and Retail Trade (where both Models 1 and 3 yield better performance measures for establishments of size 500-2,499. This latter finding can be ascertained by examining the detailed information (not presented here).

As can be seen in Tables 9 and 9a, Model 1 yields better performance measures than the current model for Government (which only has sufficient sample for establishments of size 5,000+) and yields better performance measures than the current model in roughly half the size categories for Education and Retail Trade. For the remaining supersectors, the current model tends to yield the best performance measures across size categories.

Further examination as provided in Tables 10 and 10a shows that Model 1 offers performance improvement opportunities for establishments of size 5,000+ for many supersectors (with the exception of Finance, Durable Goods Mfg, and Wholesale Trade), while performance improvement opportunities for establishments of size 500-4,999 appears limited to Retail Trade.

Establishment Estimate Level Assessment. Based upon findings from the establishment level assessments, employment level assessment was carried out for Model 1. In addition, the set of Influencer establishments used in the employment estimate level assessment was restricted to large (5,000+) establishments, as this is the set for which Model 1 appears to offer opportunities for improvement across most supersectors. Note that not all months were included in the employment estimate level for all supersectors, as there could be some months in which no large Influencer establishments were late reporters.

Model 1 yielded the lower absolute error for just over 60% of the cases. When looking at extreme errors, both the current model and Model 1 yielded roughly the same number of errors greater than 0.5 percentage points; however, when looking at errors greater than 0.2 percentage points, Model 1 did yield fewer errors, primarily in Government, Education, and Transportation. While positive results for the first two supersectors were expected given the findings from the establishment level assessment, the finding for Transportation was unexpected. Comparing errors of smaller sizes show virtually no difference between the current model and Model 1 (results not shown).

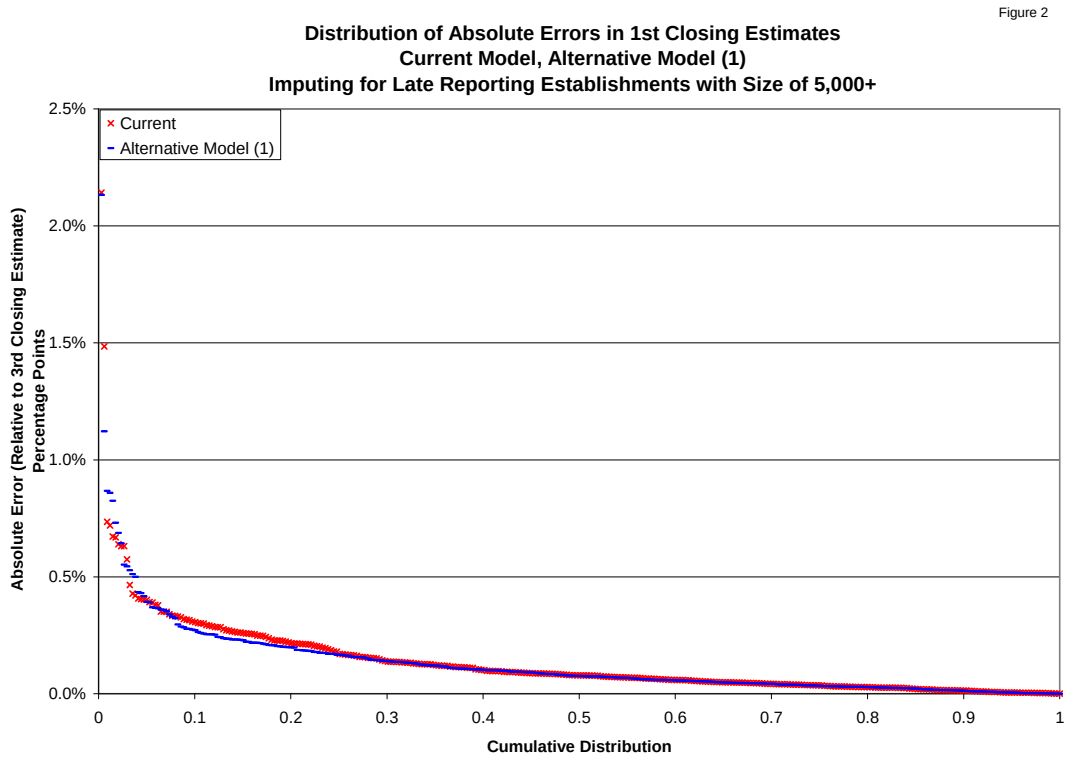
Table 11

Model Performance for Employment Estimates
Jan '04 to May '08

Supersector	# Lower Abs(Err)		# Abs(Err) >0.5%		# Abs(Err) >0.2%	
	Model		Model		Model	
	(0) Current	(1) Year-lagged Establishment	(0) Current	(1) Year-lagged Establishment	(0) Current	(1) Year-lagged Establishment
Business	25	40	0	0	6	5
Construction	0	1	0	0	0	0
Education	27	41	1	0	11	8
Finance	21	39	0	0	2	2
Government	25	41	6	5	18	11
Information	10	17	1	2	2	4
Leisure	19	41	0	0	6	5
Durable Goods Mfg	22	41	0	1	5	6
Mining	1	2	0	0	1	1
Non-durable Goods Mfg	20	38	0	0	5	5
Other Services	2	2	0	0	0	0
Retail Trade	17	39	1	1	12	11
Transportation	27	41	2	4	19	15
Utilities	15	24	0	0	9	8
Wholesale Trade	11	15	0	0	1	1
Total	242	422	11	13	97	82

To show the impact of Model 1 on errors greater than 0.2 percentage points, the distribution of the absolute errors were graphed (Figure 2). The graph shows that between roughly the ninth and twenty-fifth percentile of the distribution (which correspond to absolute errors for the current model from 0.18 to 0.32 percentage points), Model 1 yields smaller absolute errors than does the current model (generally on the order of 0.03 to 0.04 percentage points smaller). For the remainder of the distribution, the two models yield essentially the same absolute errors.

Figure 2



D. Conclusions

Results suggest use Model 1 in imputing current month employment for selected late reporters offers opportunity for improvement in first closing employment estimates. Further analysis is required to refine criteria in terms of size and supersector for application.

In addition, there are other characteristics which should be considered, which were not part of this analysis. Work by Butani, et al (2005), suggested employment growth may differ by single vs. multi-establishments employers. While this work utilized reported employment size, further analysis should make use of design size for establishments to provide more stability in the definition criteria. Consideration of performance by month and relative to overall employment trends (e.g., use of prior year employment growth may not be a good predictor when employment trends have reversed – in these situations, Model 3 may be useful) should also be part of the analysis.

V. Phase II: Model Refinement and Assessment

During Phase I several alternative methods were explored, and one method (Model 1) recommended for assessment during Phase II. The work in Phase II is intended to refine the criteria under which Model 1 should be applied, and to assess the performance of the resulting employment estimates relative to the current model.

A. Criteria Refinement

Results from Phase I of this research suggested the Model 1 may offer improved performance when applied to large establishments and in selected sectors. The first activity in the Phase II research was to conduct further analysis to identify criteria defining the set of establishments for which improved performance under Model 1 can be expected.

Accuracy by Establishment Characteristics: The first analysis entailed assessing the accuracy of imputed employment at the establishment level by selected characteristics: establishment size; establishment type; state; calendar month; difference in monthly change between current month and same month prior year; direction of monthly change for current month and same month prior year.

Change and direction characteristics were determined as follows:

For each estimation cell, the difference between the current month 1st closing link relative and the 1st closing link relative from 12 months prior was calculated

For each estimation cell, the direction of monthly change was defined as

0 if link relatives were either both greater than 1.0 or both less than 1.0

-1 if current month link relative was greater than 1.0 and 12 months prior link relative was less than 1.0

$+1$ if current month link relative was less than 1.0 and 12 months prior link relative was greater than 1.0

Within each supersector, sort in decreasing order the observed differences across all months and estimation cells and segment into quintiles to create a "growth change" variable (valued as 1, 2, 3, 4, or 5, based on the quintile)

The accuracy of imputation based on the current method and Model 1 was calculated for each month/establishment in employment size class 7 (500-999 employees) or 8 (1,000+ employees), based upon establishments reporting employment by 3rd closing in a month. The relative error for imputation was defined as

$$Err\left(\hat{Y}_{ti}^{M}\right) = \frac{\hat{Y}_{ti}^{M} - Y_{ti}}{Y_{ti}}$$

where

$\hat{Y}_{ti}^{M}$ represents the imputed employment for establishment in month

Y_{ti} represents the actual reported employment for establishment in month

The summary statistic examined was the median absolute error for month/establishments within a given

characteristic set.

$$MedAbsErr(M) = \underset{t,i\in A}{median}\left|Err\left(\hat{Y}_{ti}^{M}\right)\right|$$

Table 12 provides summary information from the imputation assessment. As can be seen Method 1 yields lower median absolute errors for many characteristics, for Government, Information, and Utilities. It can also be seen that for establishments in Quintile 1 and Quintile 5 relative to growth change, Method 1 yields lower median absolute errors for one-third or more of the supersectors.

Table 12

Characteristics for Which Imputation Model 1 Yields Lower Median Absolute Error than Current Method
Jan '05 - Jun '08

Supersector	Establishment Size		Establishment Type		Growth Change					Direction Change			Month	State
	Size Class 7	Size Class 8	Single	Multi	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	-1	0	1	#	#
Business													0	0
Construction					X				X				0	6
Durable Goods Mfg													0	3
Education													0	1
Finance													0	0
Government	X	X			X	X	X	X	X		X		6	43
Information		X		X		X	X	X		X	X	X	12	25
Leisure			X		X				X				0	5
Mining	X										X		7	18
Non-durable Goods Mfg					X				X				0	9
Other Services													0	6
Retail Trade													1	0
Transportation					X								0	3
Utilities	X	X		X	X	X	X	X	X	X	X	X	11	20
Wholesale Trade													0	13

Impact on Estimates: The second analysis entailed comparing monthly levels and month-over-month growth rates under the current model and Model 1 to actual growth rates at the establishment level, summarized by various establishment characteristics. The time period for the criteria refinement analysis was January 2005 1st closing through June 2008 3rd closing.

Three methods applying Model 1 were included in this analysis. These methods represent different forms in which imputed employment could be applied for implementation. At the estimation cell level, these models are:

Method 1.1: Imputed values are used in deriving the Link-Relative, with the form of the estimator being the same as that currently used.

$$\hat{Y}_{ct|k}^{(1.1)} = LR_{ct,(t-1)|k}^{(1.1)}\left(\hat{Y}_{c(t-1)|(k+1)}^{(1.1)} - \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{(t-1)i}\right) + \sum_{i\in s_{ct,(t-1)|k},Atyp} Y_{ti}$$

where

$$LR_{ct,(t-1)|k}^{(1.1)} = \frac{\sum_{i\in s_{ct,(t-1)|k},R} w_i Y_{ti} + \sum_{i\in s_{c(t-1)|k}} w_i \hat{Y}_{ti}}{\sum_{i\in s_{ct,(t-1)|k},R} w_i Y_{(t-1)i} + \sum_{i\in s_{c(t-1)|k}} w_i Y_{(t-1)i}}$$

$s_{c(t-1)|k}$ represents the set of sample establishments in estimation cell c that, as of closing k, reported data for both month $t-1$, but not for month t

Method 1.2: Imputed values are not use in deriving the Link-Relative; rather the weighted sums of values for the

establishments with imputed values are handled separately as are those for unweighted sums for Atypical establishments.

$$\hat{Y}_{ct|k}^{(1.2)} = LR_{ct,(t-1)|k}^{(0)} \left(\hat{Y}_{c(t-1)}^{(1.2)} - \sum_{i \in s_{ct,(t-1)|k},Atyp} Y_{(t-1)i} - \sum_{i \in s_{c(t-1)|k}} w_i Y_{(t-1)i} \right) + \sum_{i \in s_{ct,(t-1)|k},Atyp} Y_{ti} + \sum_{i \in s_{c(t-1)|k}} w_i \hat{Y}_{ti}$$

where

$LR_{ct,(t-1)|k}^{(0)} =$ current link-relative

Method 1.3: This is analogous to Method 1.2, the difference being unweighted sums of values for the establishments with imputed values are handled separately as are those for Atypical establishments.

$$\hat{Y}_{ct|k}^{(1.2)} = LR_{ct,(t-1)|k}^{(0)} \left(\hat{Y}_{c(t-1)}^{(1.2)} - \sum_{i \in s_{ct,(t-1)|k},Atyp} Y_{(t-1)i} - \sum_{i \in s_{c(t-1)|k}} Y_{(t-1)i} \right) + \sum_{i \in s_{ct,(t-1)|k},Atyp} Y_{ti} + \sum_{i \in s_{c(t-1)|k}} \hat{Y}_{ti}$$

Performance was assessed on the basis of relative mean square error to the current 3rd closing estimates, and was assessed by calendar month within supersector, s:

$$\mathrm{Re}lMSE_s^{(M)} = \frac{1}{LC_s} \sum_{c_s=1}^{C_s} \sum_{l=1}^{L} \left(\frac{\hat{Y}_{c_s,l,m}^{(M)} - \hat{Y}_{c_s,l,m}^{(3)}}{\hat{Y}_{c_s,l,m}^{(3)}} \right)^2$$

where

M refers to an imputation method (Current, 1.1, 1.2, 1.3)

s refers to a calendar month (January, February, etc.)

l refers to a calendar year (2005, 2006, 2007, 2008)

c_s refers to an estimation cell within supersector s

Table 13 provides a summary of the comparison of RelMSE to that for the current method for the three alternative methods. First, Methods 1.2 and 1.3 generally yield lower RelMSE for more months than Method 1. Second, the Method 3 yields lower RelMSE than the current method in at least eight months for Education, Government, Leisure, Other Services, and Retail Trade.

Table 13
Number of Calendar Months for Which Alternative Method has Lower RelMSE than Current Method
Jan '05 - Jun '08

Supersector	Method 1.1	Method 1.2	Method 1.3
Business	6	7	6
Construction	4	5	6
Durable Goods Mfg	4	3	3
Education	6	7	9
Finance	4	7	7
Government	9	9	8
Information	4	6	5
Leisure	9	9	9
Mining	5	5	5
Non-durable Goods Mfg	4	4	7
Other Services	5	8	9
Retail Trade	7	8	8
Transportation	6	6	6
Utilities	7	7	6
Wholesale Trade	5	5	6

VI. Summary and Next Steps

In summary, it appears that use of a simple model making use of the prior year's month-over-month growth for first closing nonreporters meeting some minimum size criteria within selected supersectors could offer improvements to the stability of employment revisions. This approach would derive imputed employment for designated first closing nonreporters as

$$\hat{Y}_{ti} = Y_{(t-1)i} \times \frac{Y_{(t-12)i}}{Y_{(t-13)i}}$$

The set of first closing nonreporters for which imputation would be carried out would be restricted on the basis of size (larger establishments) and possibly supersector. An advantage of the proposed imputation model is that imputed employment can be pre-processed once the prior month reporting is complete, with imputations used based upon the results of current month reporting outcomes.

Further refinement of imputation criteria is desirable. In addition, further analysis of performance when employment trends are changing is needed to determine thresholds which identify when the imputation model will not yield accurate results.

References

Bureau of Labor Statistics. (2001), Chapter 7, "Estimation," *Current Employment Statistics Manual*: U.S. Bureau of Labor Statistics, Washington, D.C.

Bureau of Labor Statistics. (2003), "BLS Establishment Estimates Revised to Incorporate March 2002 Benchmarks," *Employment and Earnings*: U.S. Bureau of Labor Statistics, Washington, D.C.

Butani, S., Werking, G., and Kapani, V. (2005), "Alternative Imputation Methods for Employment Data," *Monthly Labor Review*: U.S. Bureau of Labor Statistics, Washington, D.C., 3-15

Copeland, K.R. and Valliant, R. 2007. "Imputing for Late Reporting in the Current Employment Statistics Survey." *Journal of Official Statistics*, 23:69-90.

Copeland, K.R. 2004a. "An Approach to Late Reporting and Nonresponse Imputation for the Current Employment Statistics Survey," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Copeland, K.R. 2004b. “Panel Survey Estimation in the Presence of Late Reporting and Nonresponse,” PhD Thesis, University of Maryland, College Park, Maryland.

Copeland, K.R. 2003a. “Nonresponse Adjustment in the Current Employment Statistics Survey.” *Proceedings of the Federal Committee on Statistical Methodology Conference*.

Copeland, K.R. 2003b. “Reporting Patterns in the Current Employment Statistics Survey,” *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Little, R.J.A. and David, M.H. (1983), “Weighting Adjustments for Non-response in Panel Surveys,” Working paper: U.S. Bureau of the Census, Washington, D.C.

Rubin, D.B. (1976), “Inference and Missing Data,” *Biometrika*, **63**, 581-592.

Werking, G.S. (1997), “Overview of the CES Redesign,” *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 512-516.

Table 5

Establishment Level Performance Measures
by Supersector

Supersector	Model (0) Current	(1) Year-lagged Establishment	(2) Month-lagged Establishment	(3) Year-lagged Deviation
	Median Absolute Error			
Business	1.52%	2.05%	2.15%	2.02%
Construction	4.13%	6.70%	5.68%	8.13%
Education	1.18%	0.74%	1.21%	0.77%
Finance	0.81%	1.17%	1.07%	1.31%
Government	1.49%	0.63%	1.34%	0.70%
Information	1.16%	1.60%	1.59%	1.84%
Leisure	2.21%	2.48%	2.86%	2.51%
Durable Goods Mfg	0.75%	0.91%	0.87%	1.08%
Mining	1.60%	2.11%	2.17%	2.36%
Non-durable Goods Mfg	1.08%	1.49%	1.44%	1.75%
Other Services	2.33%	3.13%	3.98%	3.16%
Retail Trade	1.54%	1.54%	1.85%	1.57%
Transportation	1.42%	1.89%	1.96%	1.98%
Utilities	0.65%	0.78%	0.82%	1.00%
Wholesale Trade	1.30%	2.24%	2.13%	2.76%
	Error IQR			
Business	3.05%	3.96%	4.34%	4.09%
Construction	8.31%	13.49%	11.58%	17.01%
Education	2.35%	1.48%	2.48%	1.57%
Finance	1.61%	2.34%	2.12%	2.63%
Government	3.01%	1.26%	2.69%	1.39%
Information	2.32%	3.20%	3.18%	3.71%
Leisure	4.43%	4.96%	5.62%	5.06%
Durable Goods Mfg	1.48%	1.83%	1.74%	2.21%
Mining	3.27%	4.20%	4.31%	4.85%
Non-durable Goods Mfg	2.12%	3.00%	2.86%	3.52%
Other Services	4.59%	6.35%	8.04%	6.39%
Retail Trade	3.16%	3.06%	3.81%	3.14%
Transportation	2.84%	3.81%	3.90%	3.95%
Utilities	1.30%	1.54%	1.64%	2.01%
Wholesale Trade	2.62%	4.46%	4.25%	5.54%
	Error 90/10			
Business	10.53%	14.80%	15.91%	14.86%
Construction	25.81%	44.51%	34.74%	62.38%
Education	6.74%	5.86%	12.56%	5.95%
Finance	4.60%	6.79%	6.96%	7.81%
Government	8.06%	4.31%	11.64%	4.24%
Information	7.22%	11.19%	11.25%	11.23%
Leisure	10.93%	15.48%	17.34%	15.34%
Durable Goods Mfg	4.01%	5.83%	5.98%	9.14%
Mining	8.57%	13.80%	13.27%	21.59%
Non-durable Goods Mfg	6.50%	9.72%	9.57%	11.48%
Other Services	15.13%	19.44%	28.50%	19.50%
Retail Trade	7.84%	7.96%	11.03%	7.96%
Transportation	8.61%	12.13%	13.35%	12.08%
Utilities	3.16%	4.20%	4.76%	4.81%
Wholesale Trade	8.68%	14.28%	13.39%	20.76%

Table 6

Establishment Level Performance Measures
by Size

Size	Model			
	(0) Current	(1) Year-lagged Establishment	(2) Month-lagged Establishment	(3) Year-lagged Deviation
	Median Absolute Error			
10,000+	1.13%	0.68%	1.11%	0.76%
5,000-9,999	0.91%	0.92%	0.97%	1.07%
2,500-4,999	0.85%	1.09%	1.03%	1.24%
1,000-2,499	1.03%	1.27%	1.30%	1.47%
500-999	1.42%	1.87%	1.95%	2.09%
250-499	1.56%	2.26%	2.23%	2.46%
100-249	2.01%	3.04%	3.13%	3.20%
<100	2.77%	5.22%	5.09%	5.50%
	Error IQR			
10,000+	2.27%	1.37%	2.26%	1.50%
5,000-9,999	1.82%	1.87%	1.94%	2.13%
2,500-4,999	1.70%	2.19%	2.06%	2.50%
1,000-2,499	2.05%	2.53%	2.61%	2.93%
500-999	2.80%	3.77%	3.92%	4.22%
250-499	3.10%	4.50%	4.47%	4.94%
100-249	4.02%	6.09%	6.27%	6.37%
<100	5.71%	10.42%	10.22%	10.98%
	Error 90/10			
10,000+	6.45%	4.36%	9.56%	4.54%
5,000-9,999	5.03%	5.87%	6.67%	6.53%
2,500-4,999	4.68%	6.30%	7.01%	7.23%
1,000-2,499	5.56%	7.38%	8.04%	7.97%
500-999	7.45%	10.58%	11.20%	11.73%
250-499	8.70%	13.52%	13.63%	14.73%
100-249	11.25%	17.44%	16.98%	19.01%
<100	20.88%	35.20%	35.96%	37.08%

Table 8

Frequency with which Model has Best Performance Measure
by Size (within Supersector)

Size	Model			
	(0) Current	(1) Year-lagged Establishment	(2) Month-lagged Establishment	(3) Year-lagged Deviation
	Median Absolute Error			
10,000+	4	5	2	0
5,000-9,999	5	4	2	1
2,500-4,999	10	2	1	0
1,000-2,499	12	0	1	1
500-999	12	1	1	0
250-499	11	0	1	0
100-249	11	0	0	0
<100	13	0	1	0
	Error IQR			
10,000+	4	4	2	1
5,000-9,999	6	4	2	0
2,500-4,999	12	1	0	0
1,000-2,499	12	1	1	0
500-999	13	0	0	1
250-499	11	1	0	0
100-249	11	0	0	0
<100	14	0	0	0
	Error 90/10			
10,000+	4	5	0	2
5,000-9,999	9	2	0	1
2,500-4,999	13	0	0	0
1,000-2,499	13	0	0	1
500-999	13	1	0	0
250-499	11	0	0	1
100-249	11	0	0	0
<100	13	0	0	1

Table 9

Frequency with which Model has Best Performance Measure
by Supersector (within Size)

Supersector	Model			
	(0) Current	(1) Year-lagged Establishment	(2) Month-lagged Establishment	(3) Year-lagged Deviation
	Median Absolute Error			
Business	7	1	0	0
Construction	3	1	2	0
Education	3	3	1	0
Finance	8	0	0	0
Government	0	1	0	1
Information	6	1	1	0
Leisure	5	1	1	0
Durable Goods Mfg	5	0	1	0
Mining	5	0	1	0
Non-durable Goods Mfg	7	1	0	0
Other Services	6	0	0	0
Retail Trade	2	2	1	1
Transportation	7	1	0	0
Utilities	7	0	1	0
Wholesale Trade	7	0	0	0
	Error IQR			
Business	7	1	0	0
Construction	5	0	1	0
Education	4	3	0	0
Finance	8	0	0	0
Government	0	2	0	0
Information	7	0	1	0
Leisure	5	1	1	0
Durable Goods Mfg	5	0	1	0
Mining	6	0	0	0
Non-durable Goods Mfg	7	1	0	0
Other Services	6	0	0	0
Retail Trade	2	2	0	2
Transportation	7	1	0	0
Utilities	7	0	1	0
Wholesale Trade	7	0	0	0
	Error 90/10			
Business	7	0	0	1
Construction	6	0	0	0
Education	5	2	0	0
Finance	8	0	0	0
Government	1	1	0	0
Information	6	2	0	0
Leisure	5	0	0	2
Durable Goods Mfg	6	0	0	0
Mining	6	0	0	0
Non-durable Goods Mfg	7	1	0	0
Other Services	6	0	0	0
Retail Trade	2	1	0	3
Transportation	7	1	0	0
Utilities	8	0	0	0
Wholesale Trade	7	0	0	0

Table 8a

Frequency with which Model has Best Performance Measure
by Size (within Supersector)

Size	Model (0) Current	Model (1) Year-lagged Establishment
	Median Absolute Error	
10,000+	5	6
5,000-9,999	6	6
2,500-4,999	10	3
1,000-2,499	13	1
500-999	13	1
250-499	12	0
100-249	11	0
<100	14	0
	Error IQR	
10,000+	5	6
5,000-9,999	6	6
2,500-4,999	12	1
1,000-2,499	12	2
500-999	13	1
250-499	11	1
100-249	11	0
<100	14	0
	Error 90/10	
10,000+	4	7
5,000-9,999	9	3
2,500-4,999	13	0
1,000-2,499	13	1
500-999	13	1
250-499	12	0
100-249	11	0
<100	13	1

Table 9a

Frequency with which Model has Best Performance Measure
by Supersector (within Size)

Supersector	Model (0) Current	(1) Year-lagged Establishment
	Median Absolute Error	
Business	7	1
Construction	5	1
Education	4	3
Finance	8	0
Government	0	2
Information	6	2
Leisure	6	1
Durable Goods Mfg	6	0
Mining	5	1
Non-durable Goods Mfg	7	1
Other Services	6	0
Retail Trade	3	3
Transportation	7	1
Utilities	7	1
Wholesale Trade	7	0
	Error IQR	
Business	7	1
Construction	5	1
Education	4	3
Finance	8	0
Government	0	2
Information	7	1
Leisure	5	2
Durable Goods Mfg	6	0
Mining	6	0
Non-durable Goods Mfg	7	1
Other Services	6	0
Retail Trade	2	4
Transportation	7	1
Utilities	7	1
Wholesale Trade	7	0
	Error 90/10	
Business	7	1
Construction	6	0
Education	5	2
Finance	8	0
Government	1	1
Information	6	2
Leisure	5	2
Durable Goods Mfg	6	0
Mining	6	0
Non-durable Goods Mfg	7	1
Other Services	6	0
Retail Trade	3	3
Transportation	7	1
Utilities	8	0
Wholesale Trade	7	0

Table 10

Frequency with which Model has Best Performance Measure
by Supersector (within Size 5,000+)

Supersector	Model (0) Current	Model (1) Year-lagged Establishment
	Median Absolute Error	
Business	1	1
Construction	0	0
Education	0	2
Finance	2	0
Government	0	2
Information	0	2
Leisure	1	1
Durable Goods Mfg	2	0
Mining	0	0
Non-durable Goods Mfg	1	1
Other Services	0	0
Retail Trade	1	1
Transportation	1	1
Utilities	1	1
Wholesale Trade	1	0
	Error IQR	
Business	1	1
Construction	0	0
Education	0	2
Finance	2	0
Government	0	2
Information	1	1
Leisure	0	2
Durable Goods Mfg	2	0
Mining	0	0
Non-durable Goods Mfg	1	1
Other Services	0	0
Retail Trade	1	1
Transportation	1	1
Utilities	1	1
Wholesale Trade	1	0
	Error 90/10	
Business	1	1
Construction	0	0
Education	0	2
Finance	2	0
Government	1	1
Information	0	2
Leisure	1	1
Durable Goods Mfg	2	0
Mining	0	0
Non-durable Goods Mfg	1	1
Other Services	0	0
Retail Trade	1	1
Transportation	1	1
Utilities	2	0
Wholesale Trade	1	0

Table 10a

Frequency with which Model has Best Performance Measure
by Supersector (within Size 500 - 4,999)

Supersector	Model (0) Current	(1) Year-lagged Establishment
	Median Absolute Error	
Business	3	0
Construction	2	1
Education	2	1
Finance	3	0
Government	0	0
Information	3	0
Leisure	3	0
Durable Goods Mfg	3	0
Mining	2	1
Non-durable Goods Mfg	3	0
Other Services	3	0
Retail Trade	0	2
Transportation	3	0
Utilities	3	0
Wholesale Trade	3	0
	Error IQR	
Business	3	0
Construction	2	1
Education	2	1
Finance	3	0
Government	0	0
Information	3	0
Leisure	3	0
Durable Goods Mfg	3	0
Mining	3	0
Non-durable Goods Mfg	3	0
Other Services	3	0
Retail Trade	0	2
Transportation	3	0
Utilities	3	0
Wholesale Trade	3	0
	Error 90/10	
Business	3	0
Construction	3	0
Education	3	0
Finance	3	0
Government	0	0
Information	3	0
Leisure	3	0
Durable Goods Mfg	3	0
Mining	3	0
Non-durable Goods Mfg	3	0
Other Services	3	0
Retail Trade	0	2
Transportation	3	0
Utilities	3	0
Wholesale Trade	3	0