

CPS BRIDGE TEAM TECHNICAL REPORT 1

What Would the Unemployment Rate Have Been Had the Redesigned Current Population
Survey Been in Place from September 1992 to December 1993 ? :

A Measurement Error Analysis

Prepared by

Stephen M. Miller

For the CPS Bridge Team

February 4, 1994

Revised March 3, 1994

Executive Summary of

CPS BRIDGE TEAM TECHNICAL REPORT 1

What Would the Unemployment Rate Have Been Had the Redesigned Current Population Survey Been in Place from September 1992 to December 1993 ? :
A Measurement Error Analysis

February 1994

The official monthly labor force estimates from January 1994 onwards will be based on data from a comprehensively redesigned Current Population Survey (CPS). The redesign incorporates changes in the basic questionnaire and collection methodology. In addition, these estimates will be constructed using 1990-Census-based population controls rather than 1980-Census-based population controls.

To better understand the effect of the CPS redesign on published estimates, a Parallel Survey, intended to provide annual average labor force estimates, was conducted from July 1992 to December 1993 using a separate monthly sample of households approximately one-fifth the size of the regular CPS sample. Because of the small sample size, monthly estimates from the Parallel Survey are highly variable.

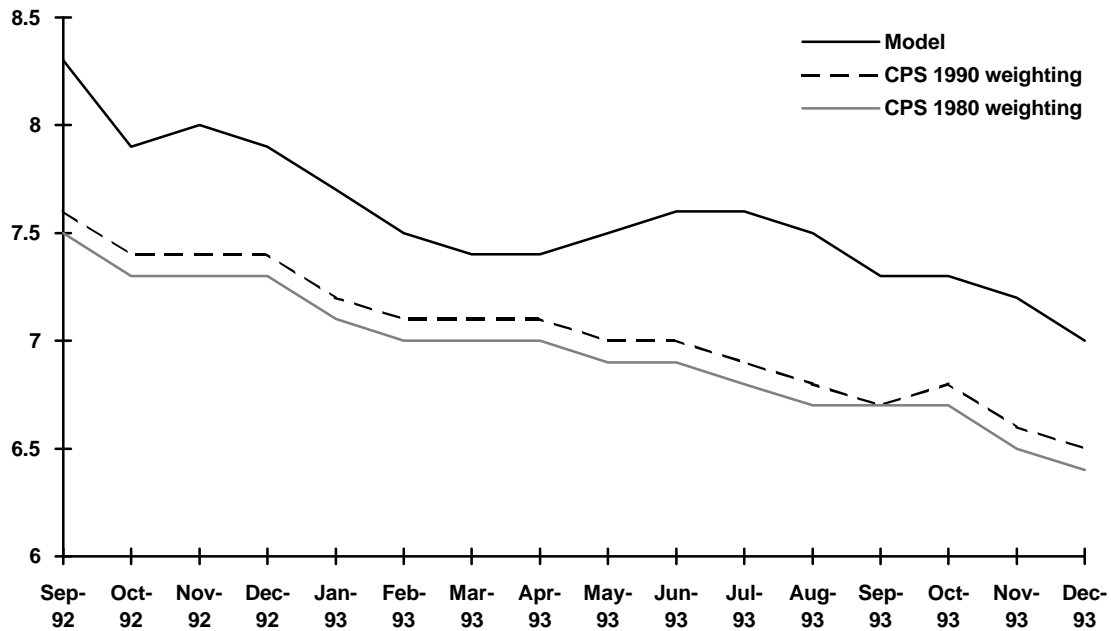
To help analysts interpret the 1994 monthly CPS estimates, the CPS Bridge team has modeled the Parallel Survey estimates from September 1992 through December 1993 for the purposes of producing more accurate estimates of what the CPS might have looked like during that period had the redesigned CPS been in place. The modeling has used only data from the Parallel Survey and the CPS and yields monthly estimates with approximately one-half the sampling variability of the monthly Parallel Survey estimates, but still about twice the variance of the regular CPS estimates. While the model estimates are less variable than the monthly estimates from the Parallel Survey, it should be noted that, because not all design differences could be taken into account, the model estimates may not provide a completely accurate picture of what the estimates from the redesigned CPS would have looked like during that period. The model estimates were seasonally adjusted with factors computed under the old CPS, so not all seasonality may be accounted for.

Several seasonally adjusted national unemployment rate series are shown below. The CPS *1980 weighted* estimates are the official seasonally adjusted rates as originally published. The CPS *1990 weighted* estimates show the effect on the original CPS estimates of switching to the 1990 Census adjusted population controls. The change to *1990 weighting* results in a 0.1 increase in the unemployment rate over the *1980 weighted* estimates. The model estimates are our estimates of what the unemployment rate would have been had the redesigned survey been in place. They are uniformly higher than both of the CPS estimates, averaging 0.6 higher than the *1980 weighted* estimates over the period. All three series show a decline in the unemployment rate of more than a percentage point between September 1992 and December 1993.

Seasonally Adjusted Unemployment Rates

Date	CPS 1980 weighting	CPS 1990 weighting	Model
1992: Sep.	7.5	7.6	8.3
Oct.	7.3	7.4	7.9
Nov.	7.3	7.4	8.0
Dec.	7.3	7.4	7.9
1993: Jan.	7.1	7.2	7.7
Feb.	7.0	7.1	7.5
Mar.	7.0	7.1	7.4
Apr.	7.0	7.1	7.4
May.	6.9	7.0	7.5
Jun.	6.9	7.0	7.6
Jul.	6.8	6.9	7.6
Aug.	6.7	6.8	7.5
Sep.	6.7	6.7	7.3
Oct.	6.7	6.8	7.3
Nov.	6.5	6.6	7.2
Dec.	6.4	6.5	7.0

Seasonally Adjusted Unemployment Rates



Contents

Twelve labor force series are modeled and used to construct unemployment rates for the period September 1992 to December 1993 which are comparable to the rates that will be produced from the redesigned CPS starting in January 1994. This report, which is the product of joint work between the Bureau of Labor Statistics and the Bureau of the Census, is laid out as follows. The *Background* section provides an introduction to the parallel survey used to evaluate the effect on annual average labor force estimates using the new questionnaire and collection procedures, as well as a discussion of the problems inherent in trying to develop an historical time series comparable to the estimates from the redesigned CPS. The *Summary of Statistical Methodology* section provides an overview of the methods used to produce the comparable series. The *Presentation of Selected Labor Force Series* section provides a summary of the *comparable data* to the official CPS data published during the same period. Finally, two *Appendices* provide additional details about the statistical modeling used.

Acknowledgments

The CPS Bridge Team consists of Chester E. Bowie, Shail J. Butani, Lawrence S. Cahoon, Cathryn S. Dippo, Robert E. Fay, Donna L. Kostanich, Elizabeth A. Martin, Stephen M. Miller, Anne E. Polivka, and N. Clyde Tucker. Other participants in the team's work included Katharine Abraham, Jack Bregger, Mick Couper, Thomas Plewes, and Jennifer Rothgeb. Outside consultants included Wayne Fuller, and Robert Groves. This report was prepared by Stephen Miller, who also was responsible for the modeling. Robert Fay computed variances for the Parallel Survey, and also provided CPS estimates for September 1992. Additional assistance was provided by Ken Robertson, Athar Malik, Ed Robison, and Janice Lent.

Background

Current Population Survey (CPS)

The Current Population Survey (CPS) is a national survey of approximately 60,000 households which provides a monthly picture of the nation's labor force characteristics. The CPS began in 1940, and the questionnaire has remained essentially unchanged since the last redesign in 1967. The current redesign takes into account changes which have taken place in the workforce and structure of American society, and takes advantage of improvements in survey methodology, such as computer-assisted data collection and questionnaire methods based on the theories of cognitive psychology. Further discussion of the changes in CPS can be found in the September 1993 issue of the *Monthly Labor Review*.

In addition to the questionnaire and mode of collection changes, new weights will be used to derive estimates from the sample. Starting in January 1994, the 1990-undercount-adjusted census population counts will be used. We will refer to this as *90 weighting*. The data used in this report are all based upon *90 weighting* unless otherwise stated. (The old weights were based on the 1980 Census of Population.)

Parallel Survey

In order to better understand the effects of the CPS redesign changes upon published estimates, BLS sponsored a Parallel Survey conducted from July 1992 to December 1993 by the Census Bureau. Using a monthly sample of approximately 12,000 households, the design of the Parallel Survey sample was a national stratified multistage probability sample which yields unbiased estimates for the nation as a whole. In contrast, the CPS sample has a State-based design which yields unbiased estimates for States and the Nation. In this report, we will utilize data for September 1992 through December 1993. The data collected during the initial two months, July and August of 1992, were not used due to startup effects in the Parallel Survey.

Polivka (1994) provides annual average estimates from the Parallel Survey. Annual averages give our best overall understanding about the differences between the redesigned CPS and the CPS prior to January 1994. These have been used as the basis of comparison because their sampling errors are much lower than those of the monthly Parallel Survey

estimates, and their use eliminates the need for seasonally adjusting the data. However, to interpret 1994 monthly estimates in comparison to 1993, analysts need comparable monthly estimates, both on a seasonally adjusted and unadjusted basis. Providing these estimates is a formidable task because of the many design differences between the CPS and the Parallel Survey. In this report, we present a set of estimates which attempt to account for some, but not all, of these differences. The major design differences are summarized in table 1.

Table 1. Design Features of CPS and the Parallel Survey (PS) from Sep. 92 - Dec. 93.

	CPS	PS
<i>SAMPLING DESIGN</i>		
Type of Design	State-Based	National
Number of sample primary sampling units (PSU)	729	283
Number of geographic areas in sample PSU's	1,973	579
Monthly sample size (occupied housing units)	58,900	12,000
Post-stratification	600 cells Plus States	559 cells NO States
Composite estimation	Yes	No
Standard Error of Monthly Unemployment Rate	0.11	0.25
Standard Error of Annual Avg. Unemployment Rate	0.07	0.16
<i>COLLECTION PROCEDURES</i>		
Mode of collection	PAPI, CATI	CAPI, CATI
% Centralized CATI	9%	18%
% True Month-in-Sample in September 1992	100%	50%
% True Month-in-Sample in October 1992	100%	50%
% True Month-in-Sample in November 1992	100%	50%
% True Month-in-Sample in December 1992	100%	50%
% True Month-in-Sample in January 1993	100%	54%
% True Month-in-Sample in February 1993	100%	58%
% True Month-in-Sample in March 1993	100%	64%
% True Month-in-Sample in April 1993	100%	76%
% True Month-in-Sample in May 1993	100%	78%
% True Month-in-Sample in June 1993	100%	86%
% True Month-in-Sample in July 1993	100%	93%
% True Month-in-Sample in August 1993	100%	97%
% True Month-in-Sample in September 1993	100%	100%
% True Month-in-Sample in October 1993	100%	100%
% True Month-in-Sample in November 1993	100%	100%
% True Month-in-Sample in December 1993	100%	100%

The model-based estimates presented in this report take into account differences between the CPS and Parallel Survey in sampling error and population controls but do not account for differences in post-stratification procedures, sample rotation, rotation group bias patterns, composite estimation, percent CATI, and seasonal patterns in monthly estimates. A brief discussion of the implications for the model-based estimates of these latter differences follows. For an in-depth discussion of the differences in sampling design between CPS and the Parallel Survey, see Kostanich and Cahoon (1994).

Poststratification

Poststratification is a statistical procedure used to ensure, among other things, that estimates aggregate to known population totals. In the CPS, estimates are controlled to match State totals. Estimates from the Parallel Survey were not controlled to match State totals. In addition, because of the small sample size of the Parallel Survey, more collapsing of race and ethnicity cells had to be done than is the case under regular CPS.

Sample rotation

The CPS uses a 4-8-4 rotation sampling scheme which means that households are interviewed for 4 consecutive months, not interviewed for 8 months, and then interviewed for 4 additional months before being permanently dropped from the sample. Each month, a new panel or rotation group is introduced, replacing the panel which received its 8th interview the previous month. Thus, each month 1/8 of the total sample is being interviewed for the 1st, 2nd, ..., and 8th time. This particular sampling scheme attempts to balance respondent burden with improved estimates of month-to-month and year-to-year change. (In any given month, 75 percent of the interviewed households were interviewed in the previous month and 50 percent were interviewed in the same month in the previous year.) Historically, it has been shown that, for a given month, the expected value of the estimated unemployment rate from each of the 8 rotation groups is different. For example, the estimated unemployment rate for the month-in-sample 1 (MIS1) panel is, on average, 7.4 percent higher than the estimated rate based on the month-in-sample 2 (MIS2) panel and 8.8 percent higher than the estimated rate based on the month-in-sample 8 panel. The estimated unemployment rate from the CPS for MIS1 through MIS4 combined is, on average, 4.3 percent higher than for MIS5 through MIS8. Although this basic 4-8-4 rotation scheme was used in the Parallel Survey, it takes 16 months to completely phase in this type of design, and the first month in which 1/8 of the sample

received its 8th interview was September 1993. Thus, if the rotation group bias pattern under the old methodology were to hold true under the new methodology, we might expect estimates from the Parallel Survey to be higher than they would have been if the sample had been fully phased in before September 1992.

Rotation Group Bias Patterns

A major change in the redesigned CPS questionnaire is the inclusion of questions each month to determine if persons not in the labor force are discouraged workers. In the "old" CPS, from January 1970 through December 1993, questions for determining discouraged-worker status were asked only of MIS4 and MIS8 households. Between January 1967 and December 1969, these questions had been asked of MIS1 and MIS5 households. When the discouraged worker questions were moved from MIS1 and MIS5 to MIS4 and MIS8, the rotation group bias patterns changed dramatically. Thus, we expect the rotation group bias pattern for the new CPS to be different from the old. Because of the small sample sizes in the Parallel Survey and the incomplete rotation phase-in, we cannot estimate what the new pattern is. On the other hand, for the parallel testing period, monthly estimates based only on true MIS1 households were 4 percent higher than for true MIS2 to MIS4 combined, as compared to a 3 percent difference in CPS, providing some evidence that there was a change.

Composite Estimation

The CPS estimates are computed each month using a composite estimator to increase the accuracy of both the monthly estimates and the estimates of month-to-month changes. The composite estimator is a weighted linear combination of two estimates for the current month. One is based on data from the entire sample for the current month; the other derives from the previous month's composite estimate plus an estimate of over-the-month change from the 75% of the sample interviewed in both months. The particular form of the composite estimator was chosen based upon the rotation group bias patterns found under the old CPS methodology. Since, as mentioned earlier, we do not have accurate estimates of the rotation group bias pattern for the new methodology, we cannot modify the form of the composite estimator. To avoid mixing January 1994 estimates based on the new methodology by using December 1993 estimates based on the old methodology, the estimate for January will not be composited. The composite estimator will be restarted in early 1994. There will be an effect due to starting up the composite estimator, but we

are unable to predict the exact effect. Because of the problems inherent in the sample rotation pattern for the Parallel Survey mentioned above, the model estimates reported in the present paper were developed based upon noncomposited Parallel Survey data.

Centralized CATI Percentage

The percentage of centralized computer-assisted telephone interviewing (CATI) was 18 percent in the Parallel Survey and 9 percent in the CPS. The redesigned CPS will have approximately 13 percent centralized CATI in January 1994, and this will increase to 18 percent sometime in 1994. For information on the implications of these differences on estimates see Kostanich and Cahoon (1994).

Seasonal Adjustment

The method used by BLS for seasonal adjustment is X-11 ARIMA, with seasonal factors reestimated every 6 months. The most recent set of factors was estimated in December 1993. For the seasonally adjusted estimates in this report, we used these most recent seasonal factors computed using data from the CPS through December 1993 based upon 1980 weighting. It was not possible to compute accurate seasonal factors from the Parallel Survey, since it usually requires at least 5 years of data to produce good monthly factors. The factors obtained from CPS data through December 1993 are the factors that will be used starting in January 1994 to adjust the redesigned CPS. BLS will monitor the seasonally adjusted estimates from the redesigned CPS to look for possible changes in seasonal patterns.

Summary of Statistical Methodology

The modeling effort reported here looked at the 12 basic series that, when aggregated, comprise the civilian labor force. These are listed in table 2.

Table 2. The 12 modeled series

1	Agricultural employment, women 16-19
2	Agricultural employment, men 16-19
3	Agricultural employment, women 20+
4	Agricultural employment, men 20+
5	Nonagricultural employment, women 16-19
6	Nonagricultural employment, men 16-19
7	Nonagricultural employment, women 20+
8	Nonagricultural employment, men 20+
9	Unemployment, women 16-19
10	Unemployment, men 16-19
11	Unemployment, women 20+
12	Unemployment, men 20+

Since the civilian noninstitutional population is known, having estimates of these 12 series makes it possible to compute estimates of those not in the labor force by subtraction. A large number of rate series then can be computed. Our approach was to produce modeled monthly unadjusted estimates for each of the 12 series and then use the standard seasonal adjustment procedure for each of the 12 series individually. Once the model estimates were computed, estimates for those not in the labor force were obtained by subtraction. The remaining series were then used to produce seasonally adjusted unemployment rates, employment-to-population ratios, and labor force participation rates.

The importance of decreasing the variability of the monthly estimates from the Parallel Survey can be seen from table 3. The first column represents the average over the 16-month period, September 1992 to December 1993, of the ratios of the variance of the monthly Parallel Survey estimates to the variance of the monthly CPS estimates. The column labeled Model gives the average, over the same 16 months, of the ratios of the variance of the monthly model estimates to the variance of the monthly CPS estimates.

Table 3. Ratio of Parallel Survey and Model variance to CPS variance

	PS	Model
Agricultural Employment		
Women 16-19	5.1	1.4
Men 16-19	5.8	1.7
Women 20+	6.4	3.8
Men 20+	3.8	2.3
Nonagricultural Employment		
Women 16-19	3.5	1.5
Men 16-19	3.7	1.3
Women 20+	5.1	2.5
Men 20+	4.4	2.2
Unemployment		
Women 16-19	4.3	1.2
Men 16-19	4.1	2.0
Women 20+	4.6	2.5
Men 20+	4.3	1.5

The average of the ratios in the first column is about 4.5, which is close to the ratio of the sample size of CPS to the Parallel Survey. In order to improve the monthly estimates of each of these 12 series, modeling was done for each series separately.

There are many different methods that could have been used to improve the monthly Parallel Survey estimates. Our approach, which may be improved upon at a later time, is described below. All of the modeling was done on non-seasonally adjusted data. Let

$$(1) \quad \begin{aligned} Y_t &= \text{Observed monthly CPS estimate under the old method} \\ y_t &= \text{True monthly CPS value under the old method} \end{aligned}$$

and let

$$(2) \quad Y_t = y_t + e_t$$

where e_t is the CPS sampling error, and $t = 1, \dots, T$, where $T = 16$. Also, let

- (3) X_t = Observed monthly Parallel Survey estimate under the new method
 x_t = True monthly Parallel Survey value under the new method

and let

(4) $X_t = x_t + u_t$

where u_t is the Parallel Survey sampling error. We will make the assumption that

(5) $y_t = b_0 + b_1 x_t$

which says that the true values have a linear structural relationship. Our goal is to estimate the true value x_t for each of the 16 months. The estimates of the x_t are what we refer to as the *model estimates*. The model parameters b_0 and b_1 allow the CPS observation to contribute to the monthly estimate. In effect, the model uses both the CPS and Parallel Survey estimate to give an improved monthly estimate. By using information from both the CPS and the Parallel survey, the model is able to reduce the variance of the monthly estimate as defined under the Parallel Survey. The reduction in variance can be seen in the column labeled Model in table 3. On average, over all 12 characteristics, the models produced a 57 percent decrease in the variance relative to the variance of the monthly Parallel Survey estimates. It should be noted, however, that the variance of the model is still greater than the variance of the monthly CPS estimate by about a factor of two.

In addition to estimating the model parameters and the monthly estimates x_t for each of the 16 months, we computed two test statistics. The first was a statistic C_G^2 tests the hypothesis that the model, as given in equation C_G^2 suggest that the model is deficient,

C_G^2 measures how much unexplained variability remains in the data apart from that

The construction and properties of C_G^2 are discussed more fully in the appendix, but under the null hypothesis that the model specified in equation (5) is adequate, C_G^2 approximately distributed as a chi-square random variable on 14 degrees of freedom.

The second test statistic was a *structural relation test* statistic C_5^2 . This statistic tests the null hypothesis $H_0: b_0 = 0, b_1 = 1$, which implies that there is no difference between the true values as measured under CPS and the Parallel Survey. Large values of the test statistic C_5^2

Survey are measuring different true values. There are other structural relationship tests one might also consider, but that selected seems a natural one. The construction and C_5^2 are discussed more fully in the appendix, but under the null hypothesis C_5^2 is approximately distributed as a chi-square random variable on 2 degrees of freedom.

4 in terms of the two tests just discussed. Table 4 lists the actual test statistics along with p underlying hypotheses. For example, a small p hypothesis, while a large p hypothesis. Values of 0.05 and 0.10 are often used as references when judging the sizes of p -values; when a p to be significant at the 0.05 level of significance. Small p of-fit statistics G^2 indicate that the model is inadequate, i.e., that the relationship in equation (5) does not adequately explain the observed variation in the data apart from sampling variability. Small p -values for the structural relationship test indicate that the model cannot be further simplified to one where CPS and the Parallel Survey are estimating the same underlying true value.

Table 4. Summary of model fitting for the 12 labor force characteristics.

	Model		Structural	
	goodness-of-fit		relationship test	
	C_G^2	p -value	C_S^2	-value
Agricultural Employment				
	10.60	0.72		0.27
Men 16-19		0.27	2.21	
Women 20+	7.72		18.22	0.00
	23.28	0.06		0.10
Nonagricultural Employment				
	19.32	0.15		0.08
Men 16-19		0.83	1.32	
Women 20+	19.48		1.95	0.38
	12.88	0.54		0.26
Unemployment				
	8.96	0.83		0.00
Men 16-19		0.64	8.17	
Women 20+	10.31		64.99	0.00
	11.95	0.61		0.19

None of the goodness-of-fit statistics in table 4 indicate substantial model failure. On the the underlying true values between CPS and the Parallel Survey, in particular those for the unemployed adult women, unemployed teenage women and adult women employed in groups, except adult men, are significantly different.

Appendix 1 presents detailed graphs of the model estimates for each of the 12 series, in data and the model estimates over time. The modeled estimates produce a smoother monthly series than the Parallel Survey, but one aspect of the model estimates needs to be average of the Parallel Survey estimates over the same 16 months. The model estimates seem to underestimate employment in agriculture, in particular, and slightly underestimate

total employment. Because total unemployment also was slightly underestimated, this did not have a noticeable effect on the reported unemployment rate estimate, but may have had a slight effect on the employment-to-population ratios and labor force participation rates presented in the next section. These last two rates may be slightly underestimated, on average. Since our main concern in this report was with producing unemployment rate estimates, estimation procedures that would have produced higher levels of employment were not investigated. Other estimation procedures may be studied in the future.

Table 5. Ratio of 16 Month Averages of
Model Estimates to Parallel Survey Estimates

	Ratio of Averages
Agricultural Employment	
Women 16-19	0.794
Men 16-19	0.889
Women 20+	0.979
Men 20+	0.932
Nonagricultural Employment	
Women 16-19	0.993
Men 16-19	1.003
Women 20+	0.999
Men 20+	1.000
Unemployment	
Women 16-19	0.996
Men 16-19	1.004
Women 20+	0.999
Men 20+	0.997

In this section, we present graphs and tables portraying seasonally adjusted and unadjusted model estimates, along with estimates from the CPS. The first two graphs show total

from the model and from the CPS under both *90 weighting* and *80 weighting*. The next

are followed by six tables, each formatted similarly to table A-33 from the monthly *Employment and Earnings*

seasonally adjusted model estimates, and table 6, the unadjusted estimates. Table 7 gives the seasonally adjusted composited CPS estimates based upon _____, with compositing beginning in October 1992, and table 8 gives the unadjusted estimates. Table _____ *80 weighted data*,

The following abbreviations are used in all of the tables:

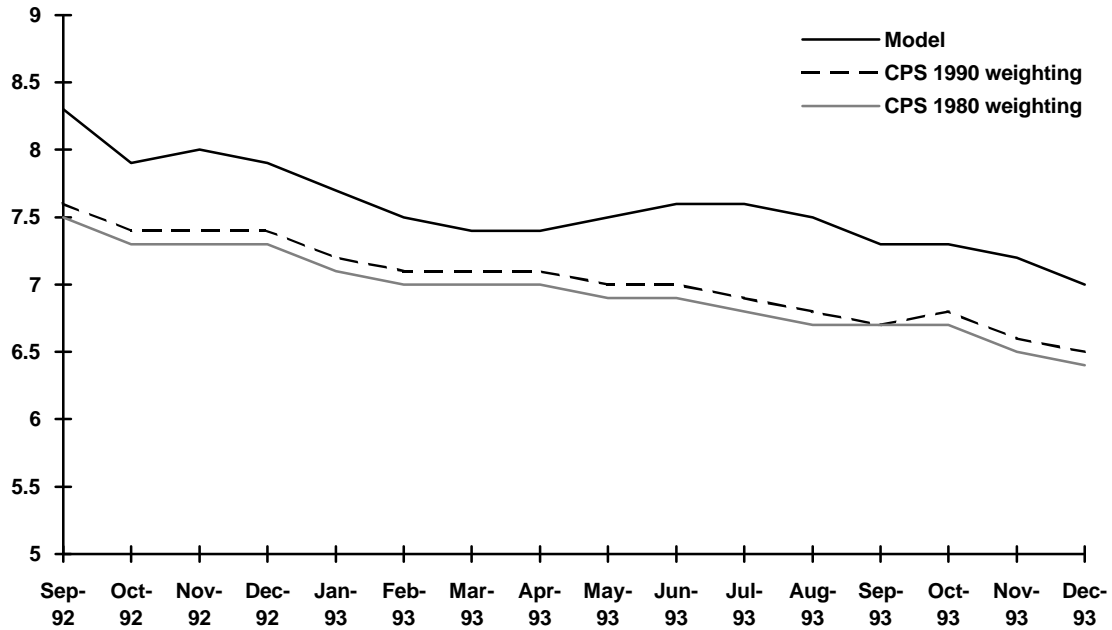
LFPR

EPR

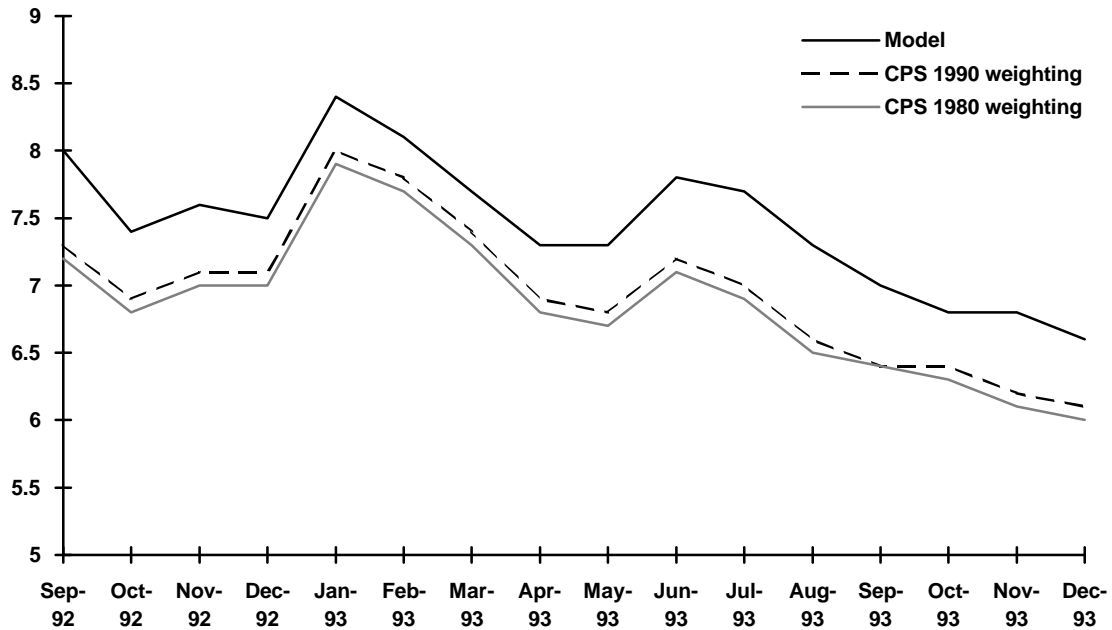
Employment to population ratio

Unemployment rate

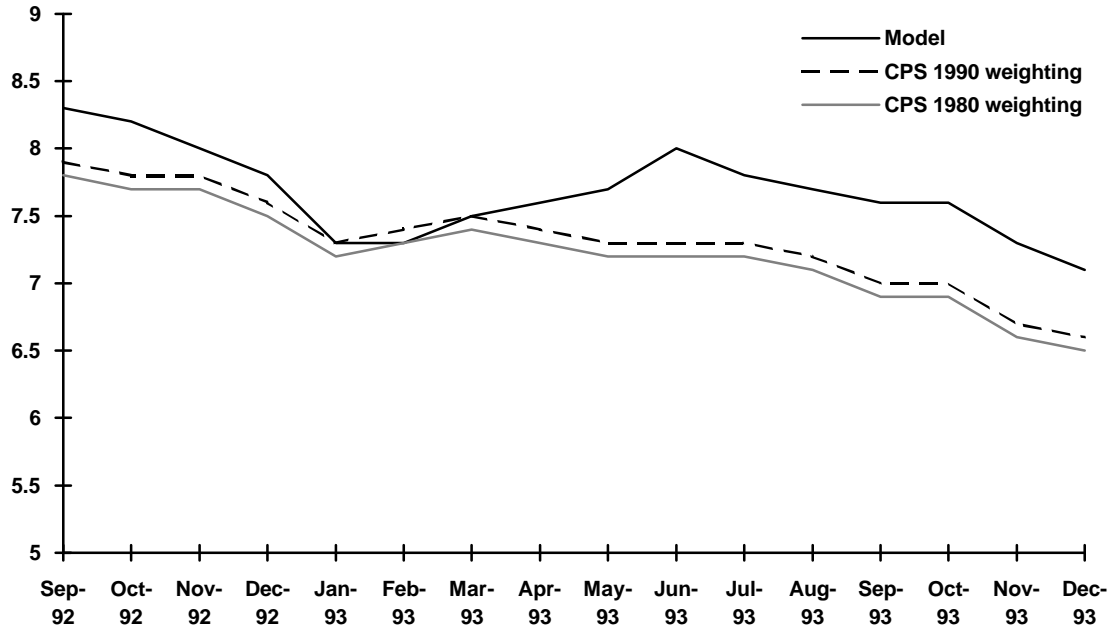
Seasonally Adjusted Unemployment Rates



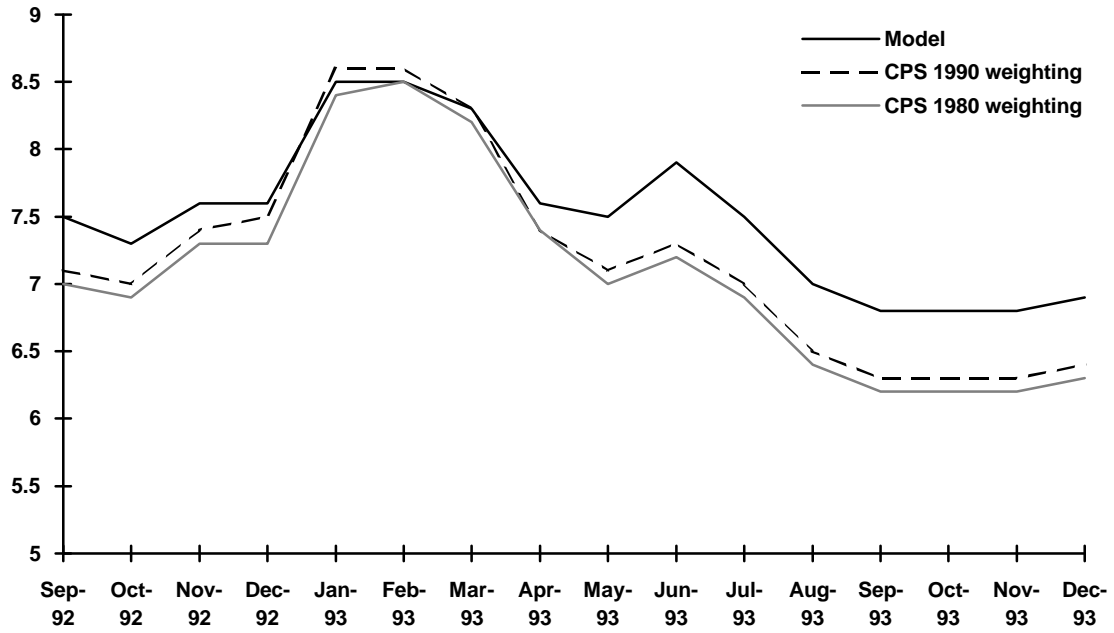
Unadjusted Unemployment Rates



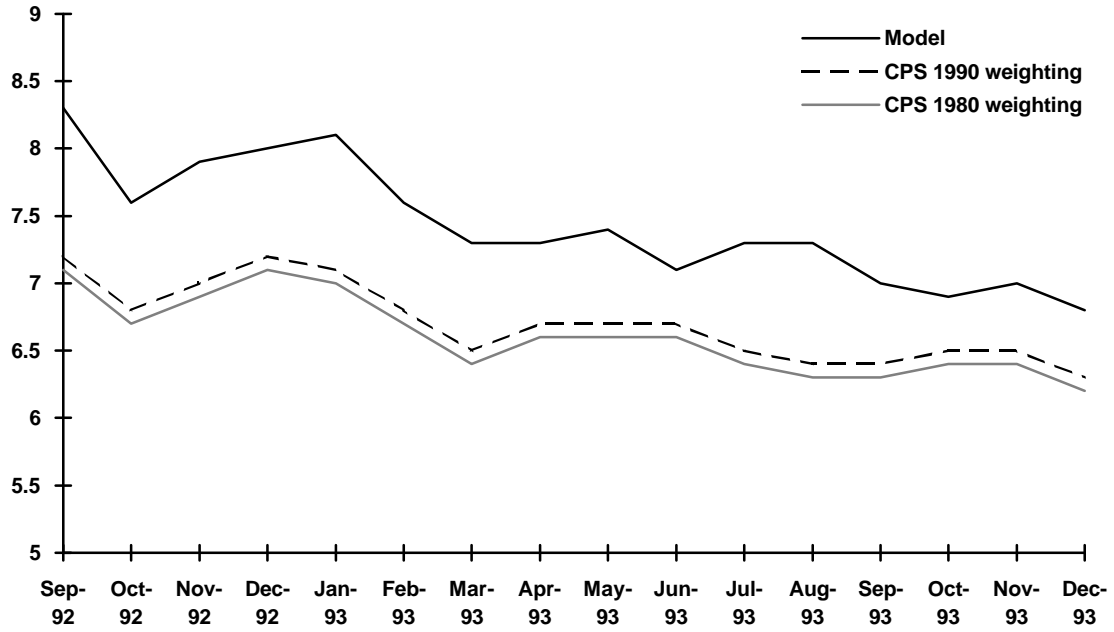
Seasonally Adjusted Unemployment Rates for Men



Unadjusted Unemployment Rates for Men



Seasonally Adjusted Unemployment Rates for Women



Unadjusted Unemployment Rates for Women

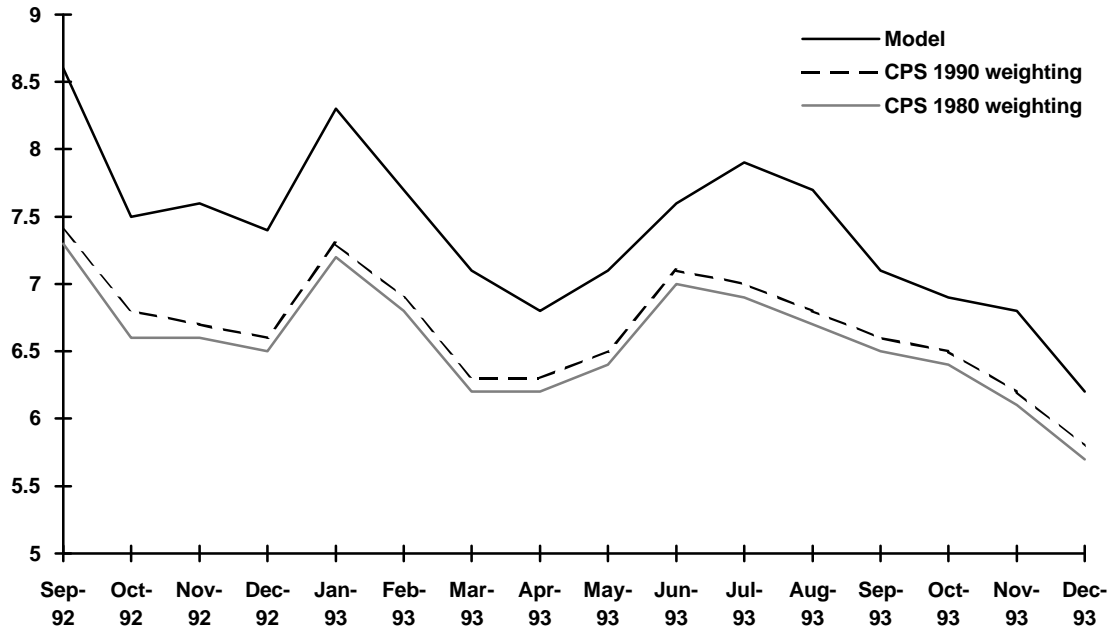


Table 5. Seasonally Adjusted Model Estimates (Based on 90 weighting)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	67.2	66.7	66.8	66.7	66.7	66.6	66.5	66.5	66.7	66.8	66.8	66.8	66.5	66.7	66.7	66.5
EPR	61.6	61.4	61.5	61.4	61.6	61.6	61.6	61.6	61.7	61.7	61.7	61.8	61.7	61.8	61.9	61.9
UER	8.3	7.9	8.0	7.9	7.7	7.5	7.4	7.4	7.5	7.6	7.6	7.5	7.3	7.3	7.2	7.0
Men 16+																
LFPR	75.7	75.4	75.2	75.0	74.7	74.8	74.9	74.9	75.1	75.1	75.0	75.0	74.8	75.0	74.9	74.9
EPR	69.4	69.2	69.2	69.2	69.3	69.3	69.3	69.2	69.4	69.2	69.2	69.2	69.2	69.3	69.5	69.5
UER	8.3	8.2	8.0	7.8	7.3	7.3	7.5	7.6	7.7	8.0	7.8	7.7	7.6	7.6	7.3	7.1
Men 20+																
LFPR	77.3	77.2	77.0	76.7	76.3	76.3	76.4	76.5	76.7	76.7	76.7	76.8	76.7	76.9	76.7	76.6
EPR	71.5	71.4	71.3	71.3	71.4	71.3	71.3	71.3	71.4	71.4	71.4	71.5	71.3	71.5	71.7	71.7
UER	7.5	7.5	7.3	7.1	6.5	6.5	6.7	6.8	6.9	6.9	6.9	6.9	6.9	6.9	6.6	6.4
Women 16+																
LFPR	59.4	58.7	59.0	59.0	59.3	59.0	58.8	58.9	58.9	59.0	59.2	59.3	58.8	58.9	59.0	58.9
EPR	54.5	54.3	54.4	54.3	54.5	54.5	54.5	54.6	54.6	54.8	54.9	54.9	54.7	54.9	54.9	54.9
UER	8.3	7.6	7.9	8.0	8.1	7.6	7.3	7.3	7.4	7.1	7.3	7.3	7.0	6.9	7.0	6.8
Women 20+																
LFPR	59.9	59.3	59.4	59.5	59.8	59.5	59.3	59.3	59.3	59.5	59.6	59.7	59.2	59.3	59.4	59.3
EPR	55.4	55.2	55.3	55.2	55.4	55.4	55.5	55.5	55.5	55.6	55.6	55.8	55.6	55.6	55.7	55.7
UER	7.6	6.9	7.0	7.2	7.3	6.8	6.4	6.4	6.5	6.4	6.7	6.5	6.2	6.1	6.2	5.9
Teens 16-19																
LFPR	54.7	52.7	53.5	53.8	54.3	54.4	54.1	53.9	54.2	54.8	54.3	53.7	53.0	53.6	53.6	53.4
EPR	42.8	42.4	42.5	42.9	43.1	43.1	42.6	42.6	42.9	42.6	43.3	43.0	43.0	43.2	43.3	43.0
UER	21.6	19.6	20.6	20.3	20.7	20.8	21.1	21.0	20.8	22.1	20.2	19.9	18.9	19.4	19.3	19.4

Table 6. Unadjusted Model Estimates (Based on 90 weighting)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	67.1	66.6	66.6	66.4	66.1	66.1	66.1	66.0	66.6	67.7	67.9	67.4	66.4	66.6	66.6	66.3
EPR	61.7	61.7	61.6	61.4	60.5	60.7	61.0	61.2	61.7	62.4	62.7	62.5	61.7	62.0	62.0	61.9
UER	8.0	7.4	7.6	7.5	8.4	8.1	7.7	7.3	7.3	7.8	7.7	7.3	7.0	6.8	6.8	6.6
Men 16+																
LFPR	75.5	75.1	74.8	74.5	74.0	74.2	74.4	74.3	75.0	76.4	76.6	75.9	74.7	74.8	74.6	74.4
EPR	69.8	69.6	69.1	68.8	67.7	67.9	68.2	68.6	69.4	70.3	70.8	70.6	69.6	69.7	69.5	69.2
UER	7.5	7.3	7.6	7.6	8.5	8.5	8.3	7.6	7.5	7.9	7.5	7.0	6.8	6.8	6.8	6.9
Men 20+																
LFPR	77.3	77.0	76.7	76.5	76.0	76.1	76.2	76.2	76.7	77.2	77.3	77.1	76.7	76.8	76.5	76.4
EPR	72.1	71.9	71.4	71.1	70.1	70.2	70.5	70.9	71.6	72.2	72.2	72.2	71.9	72.0	71.8	71.5
UER	6.7	6.7	6.9	7.0	7.8	7.7	7.5	6.9	6.7	6.6	6.5	6.3	6.2	6.2	6.2	6.3
Women 16+																
LFPR	59.3	58.9	59.1	59.0	58.8	58.7	58.5	58.4	58.8	59.7	60.0	59.6	58.7	59.1	59.2	58.9
EPR	54.2	54.5	54.7	54.6	53.9	54.1	54.3	54.4	54.6	55.2	55.3	55.0	54.5	55.0	55.2	55.2
UER	8.6	7.5	7.6	7.4	8.3	7.7	7.1	6.8	7.1	7.6	7.9	7.7	7.1	6.9	6.8	6.2
Women 20+																
LFPR	60.0	59.6	59.7	59.6	59.6	59.3	59.2	59.1	59.3	59.6	59.6	59.5	59.3	59.5	59.7	59.4
EPR	55.3	55.5	55.7	55.6	55.1	55.2	55.5	55.5	55.7	55.7	55.3	55.3	55.5	55.9	56.1	56.1
UER	7.8	6.8	6.7	6.7	7.6	6.9	6.2	6.0	6.2	6.6	7.1	7.1	6.4	6.0	6.0	5.6
Teens 16-19																
LFPR	52.2	50.6	51.3	50.8	49.2	50.0	49.8	49.5	52.4	63.1	67.1	60.9	50.7	51.5	51.5	50.5
EPR	40.9	40.7	40.7	41.3	38.5	39.0	38.7	39.3	41.1	47.5	54.0	50.7	41.2	41.5	41.6	41.6
UER	21.7	19.5	20.6	18.8	21.7	22.1	22.3	20.6	21.7	24.6	19.5	16.7	18.8	19.3	19.2	17.7

Table 7. Seasonally Adjusted CPS 90 weighted Estimates (Composited beginning with October 1992)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	66.5	66.2	66.4	66.3	66.2	66.2	66.2	66.2	66.4	66.4	66.4	66.4	66.2	66.4	66.4	66.4
EPR	61.5	61.4	61.4	61.4	61.4	61.5	61.5	61.5	61.7	61.7	61.8	61.9	61.7	61.9	62.0	62.1
UER	7.6	7.4	7.4	7.4	7.2	7.1	7.1	7.1	7.0	7.0	6.9	6.8	6.7	6.8	6.6	6.5
Men 16+																
LFPR	76.0	75.7	75.7	75.6	75.4	75.5	75.5	75.5	75.6	75.6	75.6	75.6	75.2	75.4	75.2	75.2
EPR	70.0	69.8	69.8	69.8	69.8	69.9	69.9	69.9	70.1	70.0	70.1	70.2	69.9	70.1	70.2	70.2
UER	7.9	7.8	7.8	7.6	7.3	7.4	7.5	7.4	7.3	7.3	7.3	7.2	7.0	7.0	6.7	6.6
Men 20+																
LFPR	77.7	77.6	77.5	77.4	77.2	77.2	77.3	77.2	77.4	77.3	77.4	77.4	77.1	77.3	77.1	77.1
EPR	72.3	72.0	72.1	72.1	72.1	72.1	72.0	72.2	72.3	72.3	72.3	72.4	72.2	72.4	72.5	72.5
UER	7.0	7.1	7.0	6.9	6.6	6.6	6.8	6.6	6.5	6.5	6.5	6.4	6.3	6.3	5.9	5.9
Women 16+																
LFPR	57.8	57.5	57.8	57.8	57.7	57.7	57.6	57.6	57.9	57.9	57.9	57.9	57.9	58.1	58.2	58.3
EPR	53.6	53.6	53.7	53.7	53.6	53.8	53.8	53.8	54.0	54.1	54.1	54.2	54.2	54.3	54.4	54.6
UER	7.2	6.8	7.0	7.2	7.1	6.8	6.5	6.7	6.7	6.7	6.5	6.4	6.4	6.5	6.5	6.3
Women 20+																
LFPR	58.4	58.2	58.4	58.5	58.3	58.3	58.2	58.3	58.4	58.5	58.4	58.5	58.5	58.7	58.8	58.9
EPR	54.6	54.6	54.8	54.7	54.6	54.7	54.9	54.7	55.0	55.0	55.0	55.2	55.1	55.2	55.4	55.5
UER	6.5	6.2	6.3	6.5	6.4	6.1	5.8	6.1	6.0	6.0	5.9	5.8	5.8	5.9	5.8	5.7
Teens 16-19																
LFPR	52.3	50.4	51.3	51.6	51.2	51.8	51.4	51.5	52.3	51.7	52.2	51.8	51.1	51.0	51.1	50.8
EPR	41.4	41.1	40.9	41.5	41.1	41.6	41.3	41.1	41.9	41.6	42.6	42.4	41.9	41.4	41.8	41.8
UER	21.0	18.4	20.4	19.5	19.7	19.7	19.6	20.3	19.9	19.6	18.4	18.3	17.9	18.8	18.3	17.7

Table 8. Unadjusted CPS 90 weighted Estimates (Composited beginning with October 1992)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	66.4	66.2	66.2	66.1	65.6	65.8	65.8	65.6	66.3	67.3	67.5	67.0	66.1	66.4	66.3	66.2
EPR	61.6	61.6	61.5	61.4	60.3	60.6	60.9	61.1	61.7	62.4	62.8	62.6	61.8	62.1	62.1	62.1
UER	7.3	6.9	7.1	7.1	8.0	7.8	7.4	6.9	6.8	7.2	7.0	6.6	6.4	6.4	6.2	6.1
Men 16+																
LFPR	75.9	75.4	75.3	75.1	74.6	74.9	75.1	74.8	75.5	76.8	77.2	76.5	75.1	75.2	74.9	74.7
EPR	70.5	70.2	69.8	69.5	68.2	68.4	68.8	69.3	70.1	71.2	71.8	71.6	70.4	70.4	70.2	69.9
UER	7.1	7.0	7.4	7.5	8.6	8.6	8.3	7.4	7.1	7.3	7.0	6.5	6.3	6.3	6.3	6.4
Men 20+																
LFPR	77.8	77.5	77.3	77.1	76.8	77.0	77.1	76.9	77.4	77.9	78.0	77.7	77.1	77.3	76.9	76.8
EPR	72.9	72.6	72.2	71.9	70.8	71.0	71.3	71.8	72.5	73.1	73.2	73.2	72.8	73.0	72.7	72.4
UER	6.3	6.4	6.6	6.8	7.9	7.9	7.6	6.7	6.3	6.2	6.1	5.9	5.6	5.6	5.6	5.8
Women 16+																
LFPR	57.7	57.7	57.9	57.8	57.3	57.4	57.3	57.2	57.8	58.5	58.6	58.2	57.7	58.2	58.4	58.3
EPR	53.4	53.8	54.0	54.0	53.1	53.4	53.7	53.6	54.0	54.4	54.5	54.3	53.9	54.4	54.7	54.9
UER	7.4	6.8	6.7	6.6	7.3	6.9	6.3	6.3	6.5	7.1	7.0	6.8	6.6	6.5	6.2	5.8
Women 20+																
LFPR	58.5	58.5	58.7	58.6	58.2	58.2	58.2	58.1	58.4	58.7	58.3	58.4	58.5	58.9	59.1	59.1
EPR	54.5	55.0	55.2	55.0	54.3	54.5	54.9	54.8	55.1	55.1	54.7	54.7	55.0	55.5	55.8	55.9
UER	6.7	6.1	6.0	6.1	6.6	6.2	5.7	5.7	5.7	6.1	6.3	6.3	6.0	5.8	5.6	5.3
Teens 16-19																
LFPR	49.9	48.3	49.1	48.6	46.2	47.5	47.1	47.0	50.5	60.2	65.1	59.1	48.8	49.0	49.1	48.0
EPR	39.4	39.4	39.1	39.9	36.6	37.5	37.4	37.8	40.1	46.6	53.3	50.2	40.1	39.8	40.2	40.4
UER	20.9	18.3	20.5	17.9	20.7	21.0	20.6	19.7	20.8	22.6	18.1	15.1	17.7	18.7	18.2	15.8

Table 9. Seasonally Adjusted CPS 80 weighted Estimates (Composited beginning with October 1992)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	66.4	66.1	66.2	66.2	66.0	66.1	66.0	66.1	66.2	66.2	66.2	66.2	66.0	66.2	66.2	66.3
EPR	61.4	61.3	61.4	61.4	61.3	61.4	61.4	61.4	61.6	61.6	61.7	61.7	61.6	61.8	61.9	62.0
UER	7.5	7.3	7.3	7.3	7.1	7.0	7.0	7.0	6.9	6.9	6.8	6.7	6.7	6.7	6.5	6.4
Men 16+																
LFPR	75.8	75.4	75.5	75.3	75.1	75.3	75.3	75.3	75.4	75.3	75.3	75.3	74.9	75.1	75.0	75.0
EPR	69.9	69.6	69.7	69.7	69.7	69.8	69.7	69.7	69.9	69.8	69.9	70.0	69.7	69.9	70.0	70.1
UER	7.8	7.7	7.7	7.5	7.2	7.3	7.4	7.3	7.2	7.2	7.2	7.1	6.9	6.9	6.6	6.5
Men 20+																
LFPR	77.4	77.2	77.2	77.0	76.8	76.9	76.9	76.9	77.0	77.0	77.0	77.0	76.7	77.0	76.8	76.8
EPR	72.0	71.8	71.8	71.8	71.9	71.8	71.8	71.9	72.1	72.0	72.0	72.1	71.9	72.2	72.3	72.3
UER	7.0	7.1	6.9	6.8	6.5	6.6	6.7	6.5	6.5	6.5	6.5	6.4	6.3	6.2	5.9	5.8
Women 16+																
LFPR	57.8	57.5	57.7	57.8	57.7	57.7	57.6	57.6	57.9	57.8	57.8	57.9	57.9	58.1	58.2	58.3
EPR	53.7	53.6	53.7	53.7	53.7	53.8	53.9	53.8	54.0	54.1	54.1	54.2	54.2	54.3	54.5	54.7
UER	7.1	6.7	6.9	7.1	7.0	6.7	6.4	6.6	6.6	6.6	6.4	6.3	6.3	6.4	6.4	6.2
Women 20+																
LFPR	58.3	58.2	58.3	58.4	58.3	58.2	58.2	58.2	58.3	58.4	58.3	58.5	58.4	58.6	58.7	58.9
EPR	54.6	54.6	54.7	54.6	54.6	54.7	54.8	54.7	54.9	55.0	55.0	55.1	55.0	55.2	55.4	55.5
UER	6.4	6.1	6.2	6.4	6.3	6.0	5.7	6.0	5.9	5.9	5.8	5.7	5.8	5.8	5.7	5.7
Teens 16-19																
LFPR	52.5	50.5	51.4	51.7	51.4	52.0	51.5	51.7	52.4	51.5	51.8	51.6	51.2	51.1	51.2	50.9
EPR	41.5	41.2	41.0	41.6	41.3	41.8	41.5	41.2	42.1	41.5	42.2	42.1	42.0	41.4	41.8	41.9
UER	20.9	18.4	20.3	19.5	19.6	19.6	19.5	20.3	19.8	19.5	18.4	18.4	17.9	18.9	18.3	17.8

Table 10. Unadjusted CPS 80 weighted (Composited beginning with October 1992)

	SEP92	OCT92	NOV92	DEC92	JAN93	FEB93	MAR93	APR93	MAY93	JUN93	JUL93	AUG93	SEP93	OCT93	NOV93	DEC93
Total																
LFPR	66.2	66.0	66.1	65.9	65.4	65.6	65.7	65.5	66.1	67.1	67.3	66.8	65.9	66.2	66.1	66.0
EPR	61.5	61.6	61.5	61.3	60.3	60.6	60.8	61.0	61.7	62.3	62.7	62.4	61.7	62.0	62.1	62.0
UER	7.2	6.8	7.0	7.0	7.9	7.7	7.3	6.8	6.7	7.1	6.9	6.5	6.4	6.3	6.1	6.0
Men 16+																
LFPR	75.6	75.2	75.1	74.8	74.4	74.7	74.8	74.6	75.2	76.5	76.9	76.3	74.8	74.9	74.6	74.4
EPR	70.3	70.0	69.6	69.3	68.1	68.3	68.7	69.1	70.0	71.0	71.6	71.3	70.2	70.3	70.0	69.7
UER	7.0	6.9	7.3	7.3	8.5	8.5	8.2	7.4	7.0	7.2	6.9	6.4	6.2	6.2	6.2	6.3
Men 20+																
LFPR	77.4	77.1	76.9	76.7	76.5	76.7	76.8	76.6	77.1	77.6	77.6	77.4	76.8	77.0	76.6	76.5
EPR	72.6	72.3	71.9	71.6	70.5	70.7	71.0	71.5	72.3	72.8	72.9	72.9	72.5	72.7	72.4	72.1
UER	6.2	6.3	6.5	6.7	7.8	7.8	7.5	6.6	6.2	6.2	6.1	5.8	5.6	5.6	5.5	5.8
Women 16+																
LFPR	57.6	57.6	57.9	57.8	57.2	57.3	57.3	57.2	57.8	58.5	58.5	58.1	57.7	58.2	58.3	58.3
EPR	53.4	53.8	54.0	54.0	53.1	53.4	53.7	53.6	54.1	54.4	54.5	54.3	54.0	54.5	54.8	55.0
UER	7.3	6.6	6.6	6.5	7.2	6.8	6.2	6.2	6.4	7.0	6.9	6.7	6.5	6.4	6.1	5.7
Women 20+																
LFPR	58.4	58.5	58.6	58.5	58.1	58.1	58.1	58.0	58.3	58.6	58.2	58.3	58.4	58.9	59.0	59.0
EPR	54.5	54.9	55.2	55.0	54.3	54.5	54.9	54.7	55.1	55.0	54.6	54.7	55.0	55.5	55.8	55.9
UER	6.6	6.0	5.9	6.0	6.6	6.2	5.6	5.6	5.6	6.0	6.2	6.2	5.9	5.7	5.5	5.3
Teens 16-19																
LFPR	49.9	48.3	49.1	48.6	46.2	47.4	47.1	47.1	50.6	60.3	65.3	59.2	48.7	48.9	49.1	48.0
EPR	39.5	39.5	39.1	40.0	36.7	37.5	37.4	37.8	40.1	46.7	53.5	50.2	40.1	39.8	40.2	40.4
UER	20.9	18.2	20.4	17.8	20.6	20.9	20.5	19.7	20.7	22.5	18.1	15.1	17.7	18.7	18.1	15.7

Appendix 1 Modeled Series Graphs

The following 6 pages present further modeling results. Each page contains four plots. The two plots on the left are for teenagers, and the two plots on the right are for adults. In both cases, the top plot is a scatter plot of the monthly data for both the Parallel Survey and CPS for the particular characteristic being represented. The gray line is a 45 degree reference line, and the solid black line is the estimated structural relationship as represented in equation (5). The bottom graph is a set of time series for the 16 months from September 1992 to December 1993. The solid black line is the CPS, the dashed line is the Parallel Survey, and the gray line is the model. Dots represent monthly estimates in all these series. The upper and lower gray lines (without dots) are the upper and lower bounds of an approximate 90% prediction confidence bound around the monthly model estimates. The prediction intervals form a range where we would have expected the *observed CPS* to have fallen (approximately 90% of the time) had the new methodology and weighting been used during that period. The prediction intervals account for the variance in estimating the true monthly values, as well as the variance in the monthly CPS sampling error. This is discussed more fully in Appendix 2. In all cases, the data are reported in thousands.

Figure 1. Agricultural Employment, Women.

Figure 2. Agricultural Employment, Men.

Figure 3. Non Agricultural Employment, Women.

Figure 4. Non Agricultural Employment, Men.

Figure 5. Unemployment, Women.

Figure 6. Unemployment, Men.

Appendix 2 Details of Statistical Methods

As described earlier, we modeled each of the twelve labor force series separately. We will describe the common procedure used for each below.

$$(A1) \quad \begin{aligned} Y_t &= \text{Observed monthly CPS estimate under the old method} \\ y_t &= \text{True monthly CPS value under the old method} \end{aligned}$$

and let

$$(A2) \quad Y_t = y_t + e_t$$

where e_t is the CPS sampling error, and $t = 1, \dots, T$, where $T = 16$. Also, let

$$(A3) \quad \begin{aligned} X_t &= \text{Observed monthly Parallel Survey estimate under the new method} \\ x_t &= \text{True monthly Parallel Survey value under the new method} \end{aligned}$$

and let

$$(A4) \quad X_t = x_t + u_t$$

where u_t is the Parallel Survey sampling error. We will make the assumption that

$$(A5) \quad y_t = b_0 + b_1 x_t$$

which says that the true values are linearly structurally related. If we write

$$(A6) \quad \begin{aligned} \mathbf{e}' &= (e_1, \dots, e_{16}) \\ \mathbf{u}' &= (u_1, \dots, u_{16}) \end{aligned}$$

then we can write our next assumption as

$$(A7) \quad \begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} + \begin{bmatrix} \mathbf{e} \\ \mathbf{u} \end{bmatrix}$$

We have estimates of both Σ_{ee} and Σ_{uu} , which we will consider as known and fixed for the purposes of estimation. The estimate of Σ_{ee} was based upon generalized variance function estimates and correlations obtained from previous research on the CPS. The estimate of Σ_{uu} was computed directly by replication using Robert Fay's program VPLX.

We treated the unknown parameters of the problem as

$$(A8) \quad \eta = \eta(b_1, x_1, \dots, x_{16})$$

where we treated the x_t

η by weighted nonlinear least squares. Second, use those estimates in a one-step Gauss Newton estimation to get the final estimates, as well as estimates of variance. We next describe the actual estimation of the model parameters and estimated true values in more detail.

STEP 1 Initial estimates.

Let $Z_t = \eta(X_t)$, and $\varepsilon_t = u_t$ for $t = 1, \dots, 16$. Also let

$$(A9) \quad \Sigma_{\varepsilon\varepsilon t} \equiv \begin{bmatrix} \sigma_{\varepsilon_t}^2 & 0 \\ 0 & S_{uutt} \end{bmatrix}$$

Let

$$(A10) \quad \mathbf{m}_{ZZ} = \sum_{t=1}^{16} Z_t Z_t'$$

$$\Sigma_{\varepsilon\varepsilon \bullet\bullet} = 16 \sum_{t=1}^{16} \Sigma_{\varepsilon\varepsilon t} = \begin{bmatrix} \sigma_{\varepsilon\bullet}^2 & 0 \\ 0 & S_{uu\bullet\bullet} \end{bmatrix}$$

where $\bar{Z} = \frac{1}{16} \sum_{t=1}^{16} Z_t$. We estimated the parameters b_0 and β_1 by methods described in Fuller (1987). Let $\hat{\lambda}$ be the smallest root of the determinantal equation

$$(A11) \quad |\mathbf{m}_{ZZ} - \lambda \Sigma_{\varepsilon\varepsilon \bullet\bullet}| = 0$$

and let $\hat{\alpha} = \hat{\alpha}_2 \mathbf{1}$ be the generalized eigenvector associated with $\hat{\Gamma}$ which satisfies

$$(A12) \quad \mathbb{M}_{\text{uv}} - \hat{\Gamma} \Sigma_{\text{ee}} \hat{\alpha} = \mathbf{0}.$$

Define

$$(A13) \quad \hat{b}_1 = -\hat{a}_2 \hat{a}_1^{-1}$$

$$\hat{b}_0 = \bar{Y} - \bar{X} \hat{b}_1$$

$$\hat{x}_t = \mathbb{M}_{\text{uv}} + \hat{b}_1^2 s_{\text{uutt}} \left[s_{\text{uutt}} \hat{b}_1 \mathbb{M}_{\text{uv}} - \hat{b}_0 \right] \text{et}_{\text{ett}} X_t \quad \text{for } t = 1, \dots, 16.$$

Call the initial estimates

$$(A14) \quad \hat{\theta}^{\Psi} = \mathbb{M}_{\text{uv}} \hat{b}_0, \hat{x}_1, \dots, \hat{x}_{16} \hat{\epsilon}$$

STEP 2 Weighted Nonlinear Estimation.

Define the function

$$(A15) \quad G = \sum_{t=1}^{16} \left[s_{\text{ett}}^{-1} \left(\mathbb{M}_{\text{uv}} - b_0 - b_1 x_t \right) \text{et}_{\text{uutt}}^{-1} \left(\mathbb{M}_{\text{uv}} - x_t \right) \right]$$

Let $\hat{\theta}^{\Psi}$ be the solution to the minimization of G , obtained by using $\hat{\theta}^{\Psi}$ as an initial estimates. There is an alternative formulation of this step. Define

$$(A16) \quad H = \sum_{t=1}^{16} \left[s_{\text{ett}} + b_1^2 s_{\text{uutt}} \left(\mathbb{M}_{\text{uv}} - b_0 - b_1 X_t \right) \right]$$

and let $\tilde{\beta}$ be the minimizer of H , and let

$$(A17) \quad \tilde{x}_t = \mathbb{M}_{\text{uv}} + \tilde{b}_1^2 s_{\text{uutt}} \left[s_{\text{uutt}} \tilde{b}_1 \mathbb{M}_{\text{uv}} - \tilde{b}_0 \right] \text{et}_{\text{ett}} X_t \quad \text{for } t = 1, \dots, 16.$$

Then, $\hat{\theta}^u$ is equal to $\mathbb{M}_{\tilde{\theta}_0, \tilde{x}_1, \dots, \tilde{x}_{16}} \hat{\epsilon}$. It turns out, in general, that if $\Sigma_{\epsilon\epsilon t}$ is constant over t , then the estimates in Step 1 and Step 2 give the same result.

STEP 3 One-Step Gauss Newton Estimation

Define

$$(A18) \quad \begin{aligned} \mathbf{Z} &= \begin{bmatrix} Y_1 \\ \vdots \\ Y_{16} \\ X_1 \\ \vdots \\ X_{16} \end{bmatrix} \\ \mathbf{Y} &= (Y_1, \dots, Y_{16})' \\ \mathbf{X} &= (X_1, \dots, X_{16})' \\ \mathbf{f} &= \begin{bmatrix} b_0 + x_1 b_1 \\ \vdots \\ b_0 + x_{16} b_1 \end{bmatrix} \\ \mathbf{F} &= \frac{\partial \mathbf{f}}{\partial \theta} \end{aligned}$$

and let

$$(A19) \quad \begin{aligned} \mathbf{w} &= \mathbf{Z} - \mathbf{f} \\ \mathbf{V} &= \begin{bmatrix} \Sigma_{\epsilon\epsilon} & \mathbf{0} \\ \mathbf{0} & \Sigma_{uu} \end{bmatrix} \\ \tilde{\delta} &= [\mathbf{F}' \mathbf{V}^{-1} \mathbf{F} \quad \mathbf{F}' \mathbf{V}^{-1} \mathbf{w}]' \\ \hat{\theta}^u &= \hat{\theta} + \tilde{\delta} \end{aligned}$$

where $\hat{\theta}^u$ is the final estimate. The approximate distribution of $\hat{\theta}^u$ is given by

$$(A20) \quad \hat{\theta} \stackrel{asympt}{\sim} N(\theta, [F' \Omega^{-1} F]^{-1})$$

and we estimated the variance of $\hat{\theta}$ by

$$(A21) \quad \hat{V} = [F' \hat{\Omega}^{-1} F]^{-1}$$

When we speak of the approximate distribution, we refer to an approximation as variances of the survey errors get small. This definition of asymptotics implies that the estimators obtained in each of the first two steps are consistent estimators of θ , and that the final estimator in Step 3 is efficient in the sense that they have the same limiting distribution as the maximum likelihood estimator of θ .

From now on, we drop the superscript and refer to the final estimates as $\hat{\theta}$, as well as the components of $\hat{\theta}$. Note that the resulting estimates $\hat{x} = (\hat{x}_1, \dots, \hat{x}_{16})'$ are the estimates of the true (but unobservable) value under the *New CPS* definition, and \hat{V} is an estimate of the covariance matrix of the estimates, where \hat{V} is to lower 16×16 sub matrix of \hat{V} given above. The estimates $\hat{x} = (\hat{x}_1, \dots, \hat{x}_{16})'$ can also be thought of as a predictor of what the observed CPS might have been under the *New CPS* definition, and we can use

$$(A22) \quad \hat{V}^{-1} \Sigma_{ee}$$

as an estimator of the covariance matrix of the prediction errors, where we used the covariance matrix Σ_{ee} , the covariance matrix of the CPS sampling errors, in (A7). These prediction error variances were used to construct the prediction confidence intervals in Figures 1-6.

In addition, we computed the goodness-of-fit statistic

$$(A23) \quad C_G^2 = \frac{1}{n} \sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{\hat{V}_{ii}}$$

which is approximately distributed as a chi-square random variable with 14 degrees-of-freedom

In addition, we computed a structural relationship test, which tested the hypothesis $H_0: b_0 = 0, b_1 = 1$, which says that there is no difference in the true values as defined under the Parallel Survey and CPS. The test statistic used is

$$(A24) \quad C_S^2 = \mathbf{M}' \mathbf{b} \mathbf{e}' \mathbf{e}^{-1} \mathbf{e} \mathbf{M} \mathbf{b}$$

which is approximately distributed as a chi-squared random variable with 2 degrees-of-freedom under the null hypothesis, where $\hat{\mathbf{V}}_{\beta}$ is the upper 2×2 sub matrix of $\hat{\mathbf{V}}_{\beta}$ given above, and $\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$.

References

- Fuller, Wayne A. (1987). *Measurement Error Models*. John Wiley.
- Kostanich, Donna L. and Cohoon, Lawrence S. (1994). Effect of Design Differences Between the Parallel Survey and New CPS. *CPS Bridge Team Technical Report 3*. Bureau of Labor Statistics, Washington DC.
- Polivka, Anne E. (1994) Comparisons of Labor Force Estimates for the Parallel Survey and the Current CPS. *Overlap Analysis Team Technical Report 1*. Bureau of Labor Statistics, Washington DC.