BY THE COMPTROLLER GENERAL

Report To The Congress

OF THE UNITED STATES

Funds Needed To Develop CPI Quality Control System

The Consumer Price Index (CPI) is one of the most important statistics produced by the Government. A 1-percent increase in the CPI, for example, can increase Federal outlays by about $2 billion annually for entitlement programs, such as Social Security, that have benefits adjusted by increases in the CPI.

The Bureau of Labor Statistics' quality controls provide a check on but are not sufficient to ensure the accuracy of the price data used to construct the CPI. The Bureau has requested funds to design a better system of controls. GAO recommends that the Congress approve that request.

GAO found no evidence that the reliability of the national CPI has been compromised by "bad" data. GAO and the Bureau are less confident, however, about the local area CPIs that are published as by-products of the national CPI. GAO recommends that the Bureau do a better job of alerting users to the concerns about the reliability of the local indexes.

GAO/GGD-83-32
APRIL 1, 1983

Request for copies of GAO reports should be sent to:

U.S. General Accounting Office
Document Handling and Information
Services Facility
P.O. Box 6015
Gaithersburg, Md. 20760

Telephone (202) 275-6241

The first five copies of individual reports are free of charge. Additional copies of bound audit reports are $3.25 each. Additional copies of unbound report (i.e., letter reports) and most other publications are $1.00 each. There will be a 25% discount on all orders for 100 or more copies mailed to a single address. Sales orders must be prepaid on a cash, check, or money order basis. Check should be made out to the "Superintendent of Documents".

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON D.C. 20548

B-201604

To the President of the Senate and the
Speaker of the House of Representatives

This report describes how the Bureau of Labor Statistics collects and processes data for the Consumer Price Index and points out the need for better controls over the quality of that data.

We are sending copies of this report to the Director, Office of Management and Budget and to the Secretary of Labor.

Shilton J. Hoolan

Acting Comptroller General
of the United States

COMPTROLLER GENERAL'S
REPORT TO THE CONGRESS

FUNDS NEEDED TO DEVELOP CPI QUALITY CONTROL SYSTEM

D I G E S T

GAO reviewed the Bureau of Labor Statistics' (BLS') controls over the collection and processing of consumer price data to determine if they provide reasonable assurance that the Consumer Price Index (CPI) is based on "good" data. Although GAO found no evidence that the reliability of the national CPI has been compromised by "bad" data, it did identify shortcomings in BLS' control mechanisms that could permit quality problems to go undetected or uncorrected.

WHY THE REVIEW WAS MADE

BLS calculates and publishes CPIs for the Nation, for each of the 4 regions into which it has divided the Nation, and for each of 28 cities. GAO's interest in reviewing BLS' controls over the data that are collected in support of those indexes is a consequence of the CPI's importance to policymakers at all levels of government.

At the national level, for example, the CPI is used to measure the success or failure of Government economic policy and to escalate income payments. In that regard, the Congressional Budget Office estimated in 1981 that a 1-percent increase in the CPI can increase annual outlays by about $2 billion for Federal entitlement programs, such as Social Security, that have benefits adjusted by increases in the CPI.

The importance of the CPI and the substantial amount of money tied to movements in that one statistic make it imperative that BLS assure CPI accuracy and reliability through an appropriate system of quality controls.

GAO/GGD-83-32
APRIL 1, 1983

Tear Sheet

BLS NEEDS BETTER CONTROLS TO ENSURE CPI ACCURACY

GAO's review covered that part of the CPI process most heavily laden with manual procedures, including (1) the collection of price information in the field, (2) the clerical and technical reviews directed at checking the accuracy and validity of that information, and (3) the keypunching of that information onto computer tape. GAO did not review the sampling procedures BLS uses to identify the goods and services it is going to price and the outlets at which it is going to price them. Nor did GAO review the computerized procedures that BLS follows to calculate the CPI using the keypunched price data.

BLS' procedures for collecting and processing CPI data seem to provide a good base for producing a reliable CPI. But without a system that allows management to assess data quality and identify any trouble spots, there is no guarantee that the procedures are being effectively implemented. Toward that end, BLS has a range of checks and reviews in place to identify data collection and processing errors. Although some of those checks and reviews are effective, the system as a whole provides inadequate control over CPI data quality. (See pp. 16 to 23.)

The most significant example of this inadequate control is BLS' Quality Assurance Program, which was intended to provide BLS headquarters with information on the type and extent of errors occurring in the data collection process. That program has fallen far short as an effective control mechanism because, among other things:

--The quality checks done under the program are performed in a way that subjects the results to bias.

--The information verified via the program does not represent a statistically valid sample of the total data collection effort.

--BLS does not have adequate procedures for compiling error data or standards against which to judge the data. (See pp. 18 to 21.)

BLS knows its controls over CPI data quality are inadequate. In its budget submission for fiscal year 1983, BLS asked for funds that would allow it to design a quality control system. Those funds were deleted from the President's budget. In its budget submission for fiscal year 1984, BLS again asked for funds (about $200,000) to design a quality control system. This time the funds were included in the President's budget which went to the Congress on January 31, 1983. (See pp. 24, 25, 31 and 32.)

GAO thinks the Congress should approve BLS' request. Given the significant uses of the CPI, reasonable steps need to be taken to ensure its reliability. BLS' plan to design a quality control system is a reasonable first step.

One reason it is difficult to generate a sense of urgency about quality controls in the CPI area is the fact that inadequate controls have no visible bad effect. If an automobile's reliability were to deteriorate due to inadequate quality controls, the effect would be readily apparent. But if the CPI's reliability were to deteriorate, how would one tell?

With that in mind, GAO considered two ways that it might independently assess CPI data quality but decided that both were infeasible. Instead, GAO analyzed quality related information already available in BLS to see if there was any cause for concern about CPI reliability. (See pp. 4 and 5.)

Errors are to be expected in any data gathering operation, and the BLS information GAO reviewed showed that the CPI is no exception. GAO saw nothing, however, that caused it to question the reliability of the national CPI. (See pp. 25 to 29.)

GAO does not feel as confident about the 28 city CPIs. Because the amount of data that goes into constructing these indexes is considerably less than goes into constructing the national CPI, the tolerance of local indexes to error has to be much lower. BLS readily admits that local CPIs do not provide the same level of reliability as the national CPI and that it is becoming "increasingly concerned" about the expanding uses to which those indexes are being put at the local level. (See pp. 29 to 30.)

Although BLS expresses concern about the use of local indexes, it is not effectively conveying that concern to users. Local CPIs are published via news releases issued by BLS' eight regional offices. In reviewing those releases, GAO noticed that only two of the regions included a warning about the reliability of local data. (See p. 30.)

RECOMMENDATION TO THE CONGRESS

GAO recommends that the Congress approve funding requests to design a quality control system for the CPI.

RECOMMENDATIONS TO THE SECRETARY OF LABOR

GAO recommends that the Secretary of Labor direct BLS to:

--Assess all aspects of the CPI program to identify ways it can better ensure CPI data accuracy, within the current budget environment, until a formal quality control system is implemented. In so doing,

BLS should consider such things as encouraging supervisors to make more use of the telephone to verify a data collector's work and devising online tests to assess the system's effectiveness in identifying and handling bad information that had been purposely injected into the data processing and review pipeline. (See p. 32.)

--Attach appropriate warnings to the local area CPIs so users are aware of BLS concerns about the reliability of those indexes.

AGENCY COMMENTS

BLS said it had "explicitly dedicated" some resources within the existing budget base to quality control. More specifically, BLS has set up a Quality Management Working Group, which will be responsible for designing and reviewing CPI quality control, and a Task Force, which will develop an interim Quality Assurance Program for the CPI. Establishment of these groups is an important step in the right direction.

BLS said also that it would develop "an appropriate statement" about the limitations of local area data for all its CPI releases. Such a statement has since been developed and is scheduled to be in use by the time this report is issued. (See pp. 33 and 34.)

Contents

		Page
DIGEST		i
CHAPTER		
1	INTRODUCTION	1
	How the CPI is used	1
	Objectives, scope, and methodology	3
2	THE CPI DATA COLLECTION AND PROCESSING SYSTEM	6
	Data collection	6
	Data processing	10
	Technical review	11
3	BLS NEEDS TO STRENGTHEN ITS CONTROLS OVER CPI DATA QUALITY	16
	BLS' quality control system: its strengths and weaknesses	16
	BLS has asked for funds to start developing a CPI quality control system	23
	No evidence that reliability of national CPI is being compromised by bad data; less confidence about lower level indexes	25
	Conclusions	31
	Recommendation to the Congress	33
	Recommendations to the Secretary of Labor	33
	Agency comments and our evaluation	33
APPENDIX		
I	January 31, 1983, letter from the Office of Management and Budget	36
II	January 10, 1983, letter from the Department of Labor	37

ABBREVIATIONS

BLS	Bureau of Labor Statistics
CPI	Consumer Price Index
GAO	General Accounting Office
QA	Quality Assurance

CHAPTER 1

INTRODUCTION

The Consumer Price Index (CPI) is a statistical measure of change, over time, in the prices of goods and services purchased by urban consumers. The Bureau of Labor Statistics (BLS) publishes two CPIs monthly. The first, known as CPI-W, measures the price changes associated with a market basket that represents the goods and services bought by urban wage earners and clerical workers--about 40 percent of the Nation's noninstitutionalized civilian population. Until 1978, CPI-W was the only consumer price index BLS published. In 1978, BLS began publishing a second index, known as CPI-U, which measures the price changes associated with a market basket that represents the goods and services bought by all urban consumers including the self-employed, the unemployed, and the retired--about two times the population covered by CPI-W. [1]

HOW THE CPI IS USED

The CPI is used to measure the success or failure of Government economic policy, to translate other economic indicators into inflation-free dollars, and to escalate income payments. The CPI has its most noticeable impact as an escalator.

--Millions of workers are covered by collective bargaining contracts which provide for cost-of-living increases that are based on increases in the CPI.

--Millions of beneficiaries of federally administered retirement programs, including Social Security, receive annuity increases that are based on increases in the CPI.

--Components of the food stamp program, involving outlays of about $9 billion in fiscal year 1980, are indexed to the CPI.

---

[1] We use the generic abbreviation CPI, by itself, whenever it is unnecessary to distinguish between CPI-W and CPI-U.

--The officially defined poverty level, which is the basis of eligibility in many government health and welfare programs, is updated periodically to keep in step with inflation as measured by the CPI.

--Payments for children who eat lunch at school under the National School Lunch Act and the Child Nutrition Act are adjusted on the basis of changes in a component of the CPI.

--An unknown number of rental, royalty, and child support agreements contain escalator clauses that are tied to the CPI.

BLS calculates and publishes CPIs for the Nation, for each of the 4 regions into which it has divided the Nation, and for each of 28 cities. Although many escalated income payments of the type mentioned above are tied to one of the national CPIs, several others are tied to one of the regional or city indexes. Two examples are leases and child support agreements, which often make reference to a local CPI. BLS and local government officials provided us with other examples.

--The Department of Housing and Urban Development uses regional and city indexes to help it compute Federal payments for low income housing assistance.

--Various area transit authorities, such as those in New York City, Atlanta, Chicago, and Washington, D.C., use the appropriate local CPI to escalate employee wages.

--The Kansas City, Missouri, city government uses the local CPI to compute cost-of-living adjustments for its retirement system.

--The local CPI for San Francisco/Oakland is used in drawing up wage contracts for Reno, Nevada, firefighters.

BLS estimated in 1978 that about one-half of the national population was affected directly by changes in the CPI. The Congressional Budget Office estimated in 1981 that a 1-percent increase in the CPI can increase annual outlays by $2 billion for Federal entitlement programs, such as Social Security, that have benefits adjusted by increases in the CPI. And the

CPI's impact will increase significantly in 1985 when the Government begins using it to adjust income tax brackets and personal exemption amounts for inflation.

Because the CPI has such a far-reaching impact, accuracy of the data behind it is vital. This report describes how CPI data are collected and processed and examines whether BLS has adequate controls to reasonably ensure that the CPI is based on "good" data.

OBJECTIVES, SCOPE, AND METHODOLOGY

We conducted our study in accordance with generally accepted Government audit standards. Our purpose was to evaluate BLS' practices and procedures for collecting and processing consumer price data, with emphasis on assessing how BLS controls the quality of that data. Our review covered that part of the CPI process that is most heavily laden with manual procedures, including (1) the collection of price information, (2) the clerical and technical reviews directed at checking the accuracy and validity of that information, and (3) the keypunching of that information onto computer tape.

We did not review the beginning of the CPI process at which point BLS, through various sampling procedures, identifies the specific goods and services it is going to price and the specific outlets at which it is going to price them. Nor did we review the end of the process at which point BLS takes the keypunched price data and, through various computerized applications, calculates the CPI.

We did our work at BLS headquarters in Washington, D.C., and at two of its eight regional offices--Atlanta, Georgia; and Philadelphia, Pennsylvania. We selected Atlanta because that region is one of the three largest in terms of the number of States it covers and because it provided us with the best diversity of pricing areas ranging from small urban locations to large metropolitan areas. We selected Philadelphia because that region covers the Washington, D.C. metropolitan area, which made it more convenient for us to observe the data collection process.

We observed various aspects of the data collection process in Washington; Atlanta; Philadelphia; Albany, Georgia; Baltimore, Maryland; Manassas, Virginia; and West Palm Beach, Florida. We familiarized ourselves with BLS' system for collecting and processing CPI data by reviewing operating manuals; interviewing personnel at all stages of the system, from

the data collector in the field to top management; attending BLS training courses for data collectors; observing various aspects of the system, such as price gathering; and reviewing administrative records. To assess the adequacy of BLS' procedures for controlling CPI data quality we (1) developed minimum criteria for an effective quality control system through a review of literature [2] and through discussions with various officials in the Federal statistical community including officials from BLS, the Census Bureau, the Department of Agriculture, the National Center for Health Statistics, and the Office of Management and Budget and (2) judged BLS' procedures against those criteria. The criteria are discussed in chapter 3.

Because a statistical organization can have a poor system of quality controls yet still produce good quality data, we explored ways that we might independently assess the quality of CPI data being collected in the field. We wanted to see if we could develop information that would enable us to comment on CPI reliability, irrespective of the adequacy of controls over CPI data. We considered two approaches in that regard, neither of which proved feasible.

We first tried observing data collectors during their pricing visits and recording any errors we noticed, but we abandoned that approach after accompanying several collectors. We felt, among other things, that our presence could make the collector more careful or tense and thus produce biased results.

We then considered "shadowing" data collectors--going to a sample of outlets to gather price data shortly after the collector had completed his or her visit and then comparing our results to the data collector's. We decided against shadowing after considering various negative aspects of that approach, such as (1) the burden it could impose on respondents which might jeopardize their voluntary participation in the CPI program, (2) the amount of shadowing we would have to do to produce reliable results, and (3) the very real problems we could encounter in trying to do all that shadowing without data collectors becoming aware of it.

---

[2] Some of the literature was identified through a computerized search system (called DIALOG) maintained by Lockheed Information Systems. Other literature was obtained from the Census Bureau and the National Center for Health Statistics.

Absent any workable technique for directly assessing quality, we did what we thought was the next best thing. We reviewed quality related information already available in BLS to see if it provided any cause for concern about CPI data quality. The information we reviewed and our results are discussed in chapter 3.

CHAPTER 2

THE CPI DATA COLLECTION AND PROCESSING SYSTEM

Around the 23rd of each month, like clockwork, BLS publishes the Consumer Price Index for the previous month. To meet that schedule month after month, BLS has developed a system that enables it to collect and process a large mass of price information in a relatively short time frame. This chapter describes that system.

DATA COLLECTION

Between 1972 and 1974, BLS, through the Bureau of the Census, conducted a Consumer Expenditure Survey that involved about 40,000 households (families and single persons). Through the use of questionnaires and diaries, that survey generated extensive information on how people spend their money. Using that information, BLS compiled a market basket representing the goods and services purchased by the Nation's consumers. The month-to-month pricing of the many items in that market basket forms the backbone of the CPI and marks the beginning of the data collection and processing cycle that is the subject of this report.

Some CPI pricing is done by headquarters personnel using information obtained from secondary sources. For example:

--Used car prices are obtained from the National Automobile Dealers Association's "blue book."

--Electricity prices are gathered by the Department of Energy and submitted to BLS each month.

--Paperback book prices are computed using information from the New York Times best seller list and from major paperback book publishers.

Most CPI pricing is done by BLS' eight regional offices. For about 3 weeks every month, data collectors working out of those offices visit or telephone the thousands of businesses and persons (hereafter referred to as respondents) who have agreed to participate in the CPI program. The process is a repetitious one--the same respondents are contacted and the same goods and services are priced month after month. Over a year's time, about 320 data collectors will record about

70,000 rent charges, about 28,000 property tax quotations, and about 1 million prices for food and sundry other goods and services ranging from socks to automobiles and from the cleaning of teeth to the installation of plumbing fixtures.

Those 320 data collectors are part-time employees hired to gather CPI data in 85 urban areas ranging in size from small cities to major metropolitan areas. [1] The bigger the urban area, the larger the collection workload and the larger BLS' workforce. So, while BLS can get by with only 1 data collector in places like Canton, Ohio, it needs about 18 data collectors to handle pricing activities in New York City.

Although BLS' regional offices are responsible for directly supervising data collectors, the tight time frames involved in producing the CPI demand direct lines of communication between collectors and BLS headquarters. Most significant in that respect is the flow of pricing schedules directly from Washington to the data collectors and back again.

Pricing schedules are computer generated forms that describe the items to be priced and on which data collectors record prices and price-related information. The type of information data collectors gather and the methods they use to gather that information differ for each of the three component surveys of the CPI--commodities and services, rent, and property tax.

Commodities and services

Of the items priced in the field, the Commodities and Services Survey covers all but two--rent and property tax. The many different products and services covered by this survey present data collectors with a myriad of pricing situations. The collector's work can be as easy as going into a supermarket and recording a price off a can of peas or as difficult as questioning a busy automobile mechanic about the price of parts and labor included in a carburetor overhaul.

Besides checking prices, data collectors have to pay attention to sizes and quantities. If nails were 4 cents

---

[1] BLS selected the 85 areas through a probability sample of the Nation's urban areas based on the results of the 1970 Decennial Census adjusted to 1973.

apiece last month but are selling now at 3 for 10 cents, for example, the data collector needs to record both pieces of information--the price and the number of nails offered for that price. The computer will perform the necessary mathematics to determine the per nail price for comparison with last month's figure.

Data collectors record price, size, or quantity changes by crossing out the prior month's entry which is printed on the pricing schedule and inserting the new price, size, or quantity. They are required to query respondents about any price change of 10 percent or more and to record any explanation on the pricing schedule, generally through the use of a code. According to a BLS official, that requirement serves two purposes--it acts as a check on accuracy by forcing data collectors to verify large price changes and it helps Washington assess the data coming in from the field. One code a data collector might enter on the pricing schedule, for example, would tell Washington that the price change was due primarily to a change in labor or overhead costs at the retail level.

Data collectors must also pay attention to the item description printed on the pricing schedule to ensure that the item being priced this month is the same one priced last month. In so doing, a collector may find that the outlet no longer sells the same commodity or provides the same service described on the pricing schedule. If that happens, the data collector will substitute to the most comparable commodity available in the outlet or the most comparable service provided by the outlet.

A substitution may involve only a minor change in the item being priced or it may involve a shift to a dramatically different item. For example:

--When appliances change model years, a particular refrigerator may change model numbers while remaining essentially the same refrigerator. In this case, the data collector would substitute the new model number refrigerator for the old, and BLS would proceed as if the same item were being priced. To compute the current month's CPI, BLS would follow its normal procedure--compare the current month's price of the new model refrigerator with the prior month's price of the old model.

--If a bridal dress is being priced and the outlet stops selling bridal dresses and sells no other formal dresses, the data collector may end up having to

substitute a street dress for the bridal dress. Because a street dress is so unlike a bridal dress, BLS would proceed differently than we described in the refrigerator example--it would not compare the price of the street dress with the price of the bridal dress. Instead, BLS would exclude this item from the current month's CPI computation.

A data collector might also find that although an outlet still sells the item described on the pricing schedule, the item is temporarily out of stock or out of season. If that happens, the collector will enter a code on the pricing schedule explaining why the item could not be priced and the computer will assign a price to the item through imputation--a process discussed later in this chapter.

Finally, a data collector might find that the item described on the pricing schedule can no longer be priced because the outlet has gone out of business or the respondent no longer wants to participate in the survey. In this situation, the collector will enter an explanatory code on the pricing schedule and CPI pricing activities at that outlet will cease. When Washington determines that cessation of pricing activities at outlets has adversely affected sample sufficiency, it will select new outlets to replace the discontinued ones.

Rent

For the Rent Survey, data collectors gather information on the amount of rent being paid and on the facilities and services being provided for that rent. Such information is obtained either in person or over the phone from the owners, managers, or tenants of thousands of rental units. Each of those units is priced once every 6 months, with the workload staggered to ensure pricing activity each month of the year. At pricing time, the data collector determines if there has been any change in the rent, facilities, or services between the current month and the preceding month.

If a data collector is unable to price a unit, he or she will enter a code on the pricing schedule explaining why. If the code indicates that the data collector's inability to price the unit was caused by some permanent condition, such as demolition, the unit will be dropped from the survey. If the code indicates that the condition might only be temporary, as would be the case if the data collector found the dwelling vacant, the unit will be kept in the survey and visited again either 6 months or a year later.

Property tax

The Property Tax Survey is designed to measure the taxes paid on a single-family residence which has had no change in its acreage, structure, or exemption eligibility between the tax years. If a capital change or certain exemption changes have occurred, the property is quality adjusted by subtracting the value of the change from the current assessed value of the property. This is done before comparing the current year's tax bill to the previous year's tax bill. Data collectors gather information on property tax rates and assessment values from the agency in the jurisdiction that handles property tax data. There is no interaction with property owners.

- - - -

To better ensure the reliability of the CPI as it relates to items having frequent price changes and to enhance operational efficiency, BLS has divided each month into three pricing periods, each consisting of about 6 work days. The data collection workload in each of the 85 urban areas is divided among those three periods. When one period's work is done, the related pricing schedules are mailed to Washington, where they move into the next phase of the cycle--data processing. [2]

DATA PROCESSING

Once price information has been collected, the next step is to get it into the system for use in computing the CPI. That process involves a series of clerical and mechanical reviews to make sure the information is in computer acceptable form before it is entered into the system and to make sure it is entered correctly.

As pricing schedules come into Washington from the field, those on which the item description and price have not changed since the last pricing period are separated from the rest. Only basic information from the no-change schedules, such as

---

[2] Although the information presented in the rest of this chapter generally applies not only to the price information coming in from the field but also to the price information collected in Washington, we have written it to specifically relate to the field-collected data.

outlet name and schedule number, is entered into the computer. Using that identifying information, the computer will automatically locate the previous month's price for any unchanged item and use that price in computing the current CPI.

Schedules involving some change in the item's price and/or description are reviewed by the clerical staff, a review that is directed basically at form rather than substance. The clerks, for example, would be looking for instances where the data collector (1) crossed out the previous month's price as printed on the pricing schedule but failed to record a new price, (2) indicated by a code that he or she had to make a substitution but recorded no data for the new item, or (3) wrote a message on the pricing schedule that was too long to fit in the computer. Clerks would contact commodity analysts, whose roles are discussed below, for advice on how to treat the pricing schedules in the first two examples. In the third example, the clerk would simply edit the message down to an acceptable length.

After the preliminary clerical review, schedule information is keypunched onto computer tape. The keypunched information is then computer reviewed for errors, such as incorrect codes or transposed figures, which might have been missed during the preliminary review or introduced during keypunching. After any errors identified during that mechanical review are corrected, the computer generates a Schedule Review Listing which shows all the information that had been keypunched. Clerks then compare each of the pricing schedules to the listing to make sure information on the schedules was entered into the system accurately.

TECHNICAL REVIEW

While clerks are comparing the pricing schedules to the Schedule Review Listing, other persons, known as commodity analysts, are conducting a more technical review--a review that can result in significant changes to the CPI data.

Commodity analysts are economists responsible for validating and analyzing price data and for explaining short-run and long-run price trends. Each analyst (BLS has about 30) is responsible for specific CPI items. One analyst, for example, might be assigned prescription drugs and new cars while another would have men's apparel and reading materials. The analysts are expected to learn as much as they can about the items assigned them--technical developments, price trends, and

the like. Drawing on that knowledge and considering the information supplied by data collectors, commodity analysts decide whether and to what extent certain price information will be used in computing the current month's CPI.

The primary tool used by commodity analysts in this decision-making process is the Commodity Review Listing--a computer generated document that contains information on items meeting certain criteria. An item will generally show up on a listing, for example, if its price changed by 10 percent or more since the last pricing period. The responsible commodity analyst will assess the reasonableness of the new price based on his or her knowledge of the product after taking into account the data collector's explanation for the price change.

Because there is not enough time in a pricing period to contact the field to verify questionable data, analysts will nullify prices they feel are incorrect. Say, for example, that the analyst responsible for bakery products sees that the reported price of a loaf of white bread at one outlet in Philadelphia has risen from 59 cents last month to 89 cents this month while the reported prices for similar loaves at other outlets have remained fairly constant. If the data collector has not satisfactorily explained the increase, the commodity analyst may nullify the reported price, thus excluding it from use in computing the current month's CPI.

When a reported price has been nullified or, in certain instances, when no price has been reported (such as when the item is temporarily out of stock or out of season), BLS' computer system is programmed to impute a price. Imputation involves assigning the item a price on the basis of the price movement of similar items in the same geographic area, so that a price will be available for comparison in the next pricing period.

In our white bread example, the imputed price would be based on the price movement of the other loaves of white bread that had been priced that month in Philadelphia. To oversimplify: if the prices of other loaves of white bread in Philadelphia rose an average of 2 percent since the last pricing period, the price of the loaf in question will be imputed at 60 cents (a 2-percent increase over the prior month's price of 59 cents). The price picked up by the data collector next month will be compared to this month's imputed price in calculating next month's CPI.

An item will also show up on a Commodity Review Listing if its description has changed since the previous month. The

commodity analyst is interested in whether the item priced this month is comparable to the one priced last month. To help the analyst in that regard, the Commodity Review Listing includes this month's and last month's description of the item.

The analyst's judgment as to comparability is important because it determines the extent to which price data will be used in computing the CPI. If the data collector had been pricing a 12-ounce package of Kraft American cheese, for example, but had to shift to a 12-ounce package of Borden American cheese because the outlet stopped stocking Kraft, the analyst would probably deem the substitution comparable. This month's price of Borden would then be compared to last month's price of Kraft in computing this month's CPI. If the data collector had been pricing a London Fog trench coat with a zip-out lining, however, and had to shift to a store brand trench coat with no lining, the analyst would probably deem the substitution noncomparable.

When a substitution is deemed noncomparable, the new item is introduced into the CPI sample through a process called linking. With linking, the price for the trench coat will not be used in computing this month's CPI. When a price is obtained for the store brand trench coat next month, it will be compared to this month's price in computing next month's CPI. The London Fog trench coat will be dropped from the sample, and the store brand trench coat will be priced as long as it is available.

Sometimes an item's description will change from one month to the next because of a change in the item's quality. A company might change its paper towels from 1 to 2 ply, for example, or a landlord might add window air conditioners to a rental unit, or an automobile towing service might change the number of miles included in its basic charge. Under ideal conditions, the commodity analyst would adjust the price data coming in from the field to account for any quality change so that the CPI will reflect pure price changes only. In actuality, however, it is impractical to expect analysts to place a value on every change in the quality of CPI goods and services.

Assume, for example, that the paper towel manufacturer increases the price of a roll of towels by 20 cents while changing the towels from 1 to 2 ply. The commodity analyst is faced with deciding how much of the price increase is due to the change in quality and how much is a pure and simple price

increase. Unless the analyst has some basis for making that decision, he or she would probably treat the change like an ordinary noncomparable substitution.

In some cases, the commodity analyst is able to put a price tag on a quality change. The introduction of new car models, for example, often involves items such as radial tires or rear-window defoggers being included in the base price where before they were offered as options. If those items were not included on the prior priced model car, the commodity analyst will consult automobile industry data to determine the value of the items and will use that value to compute a quality adjustment. The effect of a quality adjustment is to factor out the portion of the price change that is due to quality change but allow the portion that is pure price change to be reflected in the CPI.

Because commodity analysts have to make decisions based on little more than the information provided by data collectors, communication between the analysts and the collectors is vital. Much of that communication is in the form of messages on the pricing schedules.

A message from the commodity analyst will appear on the pricing schedule that goes out to the data collector at the beginning of the pricing period. For example, the data collector may see a message telling him or her to verify or explain the price change reported on the prior month's schedule. A commodity analyst would send that sort of message if the data collector had failed to explain a price change of 10 percent or more or if a price change was out of line with other prices reported for the same geographic area--as was the case in our white bread example on page 12.

A message from the data collector will appear on the completed pricing schedule that comes into Washington from the field. The message might be an explanation of why the data collector did something, an explanation of why an item could not be priced, a response to an earlier message from the commodity analyst, or a question on a pricing problem. The presence of any message on the pricing schedule is another criterion that will trigger an item's appearance on the Commodity Review Listing.

If a commodity analyst wants to nullify a price, make some other change to the information on the Commodity Review Listing, or send a message to the data collector, he or she provides the necessary information to the processing clerks

for entry into the computerized data base. After the information has been entered, General Review Listings are produced. These computer generated documents show the changes made to the data base as a result of the commodity analyst's review of the Commodity Review Listing as well as the clerk's review of the Schedule Review Listing. Clerks check the General Review Listings to verify that all changes were properly entered into the computer.

After corrections necessitated by reviews of the General Review Listing have been made, the data base is ready for use in constructing the CPI. It is at this point too that the computer, using the information in the data base, prints out the pricing schedules for next month. Those schedules are mailed to the field and the data collection and processing cycle begins all over again.

CHAPTER 3

BLS NEEDS TO STRENGTHEN ITS CONTROLS OVER CPI DATA QUALITY

For any statistical series, especially one so widely used as the CPI, accuracy is vital. BLS' quality control system provides a check on, but does not ensure, CPI accuracy. Although we found no reason to believe that the reliability of the national CPI has been compromised by inadequate quality controls, we are not as confident about local area indexes. BLS has taken steps toward upgrading its controls, but any significant stride in that direction depends on the availability of funding. Even absent full funding, however, BLS can still make some improvements.

BLS' QUALITY CONTROL SYSTEM: ITS STRENGTHS AND WEAKNESSES

BLS' procedures for collecting and processing CPI data seem to provide a good base for producing a reliable CPI. But without a system that allows management to assess data quality and identify any trouble spots, there is no guarantee that the procedures are being effectively implemented. Toward that end, BLS has a range of checks and reviews in place to identify data collection and processing errors. Although some of those checks and reviews are effective, the system as a whole provides inadequate control over CPI data quality.

Controls over data collection

BLS' basic check on the data collection process is the Quality Assurance (QA) Program operating out of each of the Bureau's eight regional offices. The Program, which began in 1977, was designed to serve two purposes: to help the regional offices evaluate data collector performance and to provide BLS headquarters with information on the type and extent of errors occurring in the data collection process.

The QA Program consists of two types of checks: (1) observation, in which a supervisor accompanies the data collector during his or her pricing visits and (2) reinterview, in which a supervisor repeats the work done by a data collector and compares the results. Observation enables the supervisor to watch a data collector in various pricing situations and to evaluate general qualities such as interviewing

skills and the ability to organize materials. Reinterview enables the supervisor to check the data collector's accuracy and thoroughness in filling out a pricing schedule. The reinterview is a vehicle for (1) checking that the data collector actually visited the outlet and priced the good or service, (2) ensuring that the pricing was done in accordance with technical instructions and procedures, and (3) identifying errors and their causes.

We concerned ourselves only with the reinterview aspect of the QA Program because that is the aspect BLS looks to for error-related information.

Supervisors in BLS' regional offices are responsible for scheduling and performing reinterviews following certain general guidance furnished by headquarters. For example, BLS requires that each data collector be checked by both observation and reinterview at least annually and that each reinterview cover a minimum number of specific type outlets, such as two food stores.

Supervisors perform reinterviews as soon as possible after the data collector has finished pricing the item to avoid differences caused by a lapse of time. The supervisor is supposed to check the original pricing schedule for completeness and accuracy, record any differences on a reinterview report, and discuss those differences with the respondent. If the supervisor sees, for example, that a 1-pound can of coffee is marked at $2.20 instead of the $2.09 reported by the data collector, he or she will check with the respondent (probably the store manager). The respondent can tell the supervisor whether or not $2.09 was the correct price when the data collector was in the store.

After completing a reinterview, the supervisor is supposed to meet with the data collector. That meeting serves two purposes: it provides the data collector with feedback on the results of the supervisor's checks and it gives the supervisor another chance to reconcile any differences that could not be resolved with the respondent.

On the basis of his or her discussions with the respondent and the data collector, the supervisor will record a reason for each of the differences noted during the reinterview. BLS' QA Program manual gives the supervisor five possible reasons to choose from: (1) the data collector made a mistake, (2) the supervisor made a mistake--which means the information recorded by the data collector on the pricing schedule was correct, (3) the respondent gave the data collector incorrect

information, (4) there was a price change between the time the data collector priced the item and the time of the reinterview--which means the information on the pricing schedule was correct, or (5) the reason is unknown.

In an attempt to determine what elements had to be present if a reinterview was to be an effective quality control mechanism, we reviewed pertinent literature and talked with statistical experts in BLS, the Census Bureau, the Department of Agriculture, the National Center for Health Statistics, and the Office of Management and Budget. On the basis of what we read and what we were told, BLS' QA Program falls short as an effective control mechanism for several reasons.

No surprise

Data collectors sometimes know when the supervisor will be checking their work, thus eliminating a necessary element of effective control--surprise. In that regard, BLS' QA manual provides conflicting guidance. One section of the manual says that data collectors "should not know * * * the timing of reinterviews" but another section says that the data collector should be contacted "at the start of the pricing period" and instructed to "retain completed price schedules until the supervisor arrives." Our field work in BLS' Philadelphia and Atlanta regional offices indicated that some supervisors were following the latter instruction--thus giving data collectors advance notice of their quality checks.

Dependent instead of independent price verification

Supervisors conduct their reinterviews with the data collectors' pricing schedules in front of them. That kind of dependent verification produces unreliable QA results because the supervisor, by looking at the data collector's pricing schedule, can be conditioned to expect to find certain things. Also, as we noticed while observing some reinterviews, a supervisor can be tempted to verify information simply by reciting it from the data collector's pricing schedule and asking the respondent to affirm it.

Independent verification should produce more reliable results from a quality control standpoint. Independent verification means that supervisors fill out their own pricing schedules as if they were data collectors and that respondents provide information without prompting from the supervisors. Only after completing his or her pricing schedule would the

supervisor look at the schedule completed by the data collector, identify any differences, and attempt to reconcile them.

No statistically valid sample

For reinterview results to be of any value to management in assessing the quality of its data collection effort, the items verified during the reinterviews need to represent a statistically valid sample of the total data collection effort. To meet that criterion on a nationwide basis, it seems reasonable to suggest that BLS headquarters would need to coordinate, if not dictate, the selection of items to be verified. In actuality, however, the items verified under the QA Program do not represent a statistically valid sample. Each region decides how many reinterviews it is going to do during the year on the basis of available resources, and each supervisor decides when he or she is going to conduct a particular reinterview and what he or she is going to verify during that reinterview with only minimal national office guidance.

The QA manual, for example, says that for each data collector being checked, the supervisor should select a minimum of two service outlets; however, it is left to the supervisor to decide on the specific outlets. That kind of discretion makes it easy for a supervisor to avoid "problem outlets" even though those are the outlets where data collectors may be having the most difficulty. As one supervisor told us: she tries to avoid outlets where the respondents are known to be difficult. She explained that some respondents do not want to spend much time with data collectors and that matters would only get worse if they were asked to spend even more time with a supervisor. Although avoiding such outlets may be pragmatic in that it lessens the risk of respondents dropping out of the CPI program, it is unacceptable from a quality control standpoint.

Insufficient coverage of nonresponse outlets

Closely related to the issue of a statistically valid sample is the fact that the QA Program does not provide sufficient coverage of nonresponse outlets. We use the term "nonresponse outlet" to identify any outlet in which the data collector was supposed to price one or more items but did not or could not. Such would be the case, for example, if an outlet was temporarily closed, if the respondent was unavailable

during the pricing period, if the respondent decided not to participate in the CPI program any more, or if the outlet went out of business or moved out of the area. An effective quality control system not only would check whether a data collector recorded certain price information correctly but also would verify that the data collector had a valid reason for not recording other information. Nevertheless, the QA Program manual does not specifically require supervisors to perform reinterviews at nonresponse outlets. Not surprisingly, then, our review of 1,813 reinterviews done in fiscal year 1980 throughout BLS' eight regional offices showed that only one involved a nonresponse outlet.

- - - -

Even if the QA Program were revised along the lines suggested above, the information generated by the reinterviews would still be useless from an analytical standpoint because BLS has not developed adequate procedures for compiling QA error data or standards against which to judge the data.

BLS attempted, through the use of quarterly reports, to compile information on the types of errors occurring in the data collection process as identified through the QA reinterviews. With that information, BLS hoped to identify problem areas so corrective action, such as additional training, could be implemented. At the same time, regional personnel hoped to use the quarterly reports for indications as to how well their region was performing compared to others.

According to BLS officials in Washington and in the field, however, the quarterly reports proved to be cumbersome and unnecessarily voluminous. As one regional official wrote:

> "The number of tables that one must wade through to find pertinent information is burdensome. For example, the property tax printout provides 18 different tables, many of which are variations on the same theme. The two inch thick rent report with 52 separate tables has proven to be even more unmanageable, while the three inch thick [Commodities and Services] report might quickly extinguish any interest in Q.A. results whatsoever."

And in the words of a headquarters official:

> "It appears that all offices are unhappy with the [QA] program as it is presently constituted. The

output of the program to date is in a form that is virtually unusable by any of the offices in Washington."

What made the program output unusable was not only the unmanageable volume of data in the quarterly reports but also the fact that not enough QA checks were being performed to support any type of quarterly analysis. And, more importantly, BLS never established any benchmarks against which to measure the data. Because no statistical program can be expected to operate error free, it is important for management to decide what level of quality it wants to achieve and what amount of error its statistical series can tolerate. BLS has not done that for the CPI, so it has no basis for analyzing QA results.

BLS stopped producing the quarterly QA reports in 1980, primarily because of their limited use.

Commmodity analysts provide a second important level of control over the quality of CPI data. Among other things, they review data meeting certain criteria, delete prices they feel are inappropriate, adjust prices to reflect changes in quality, and conduct special analyses. Although commodity analysts are generally recognized as experts and appear to serve a vital review function, their effectiveness from a quality control standpoint is impaired by the absence of any procedure within BLS for routinely compiling information on the changes made and questions raised by the analysts.

For example, there are a variety of standard messages that commodity analysts can send to data collectors after reviewing Commodity Review Listings. Several of those messages relate to data accuracy, such as:

--"Component prices do not add to total. Please check."

--"Explain price change." (Used when a price change of 10 percent or more was not explained by the data collector or when a price change is out of line with other prices reported for the area.)

--"Size not entered on schedule." (Used when the data collector failed to report the item's size in cases where size is a required entry--such as for food and alcoholic beverages.)

Even though the frequency with which analysts send those type messages to the field could be looked to as an indicator of quality, BLS does not compile that type of information.

Controls over data processing

As discussed in chapter 2, the clerical and mechanical processing stage includes checks to ensure that information is being accurately transcribed. The first check is a computer review of keypunched data. That review identifies syntax errors--errors such as entry of a numeric code when only an alphabetic code is allowed or entry of a two digit code when there is space on the computer tape for only one digit. A second important check involves clerical comparison of the pricing schedules with computerized printouts showing the data entered into the computer from those schedules.

Also, the private contractors who keypunch CPI data for BLS are required, by terms of their contracts, to maintain an error rate no higher than one-tenth of 1 percent of the total number of characters keyed in a batch of pricing schedules. According to a typical keypunch contract we reviewed, if BLS determines through inspection of a stratified random sample of a batch that the error rate exceeds one-tenth of 1 percent, the batch will be returned to the contractor for repunching.

Our inquiries, although not as thorough as our inquiries into the QA Program, were sufficient to satisfy us that the clerical and mechanical reviews are structured in a way that should provide adequate control over the data processing phase of the CPI cycle. We did not do the kind of analysis that would enable us to determine whether those reviews are actually providing the level of control suggested by their structure. Nor did we assess whether existing procedures, such as clerical review of all keypunched data, represent the most efficient way for BLS to control quality at this stage of the process. We were more interested in determining whether BLS had adequate controls rather than assessing the efficiency of those controls.

Controls over commodity analysts

Quality controls over the work done by commodity analysts are more difficult to devise than controls over the work done by data collectors, clerks, and keypunchers because much of what an analyst does involves judgment. Though difficult to devise, controls are vital because analysts, through their decisions to delete or revise field collected price information, can significantly affect CPI accuracy.

Price change or deletion decisions by commodity analysts are not subject to routine review. Analysts maintain no structured records on the decisions they make and BLS has no

system to gauge the appropriateness of analysts' actions or inactions. In that regard, we do not know how often commodity analysts delete, revise, or question field collected data because BLS maintains no such summary data. Supervisors of commodity analysts and a staff of economists review processed data after changes and deletions have been made by the analysts, but the review is to identify and verify abnormally large changes in index figures, not to assess the normal price change or deletion decisions made by the analysts.

BLS HAS ASKED FOR FUNDS TO START DEVELOPING A CPI QUALITY CONTROL SYSTEM

BLS knows its controls over CPI data quality are inadequate and has developed a plan for correcting the situation. Now it needs money to put its plan into action.

In April 1980, the then Assistant Commissioner for Survey Design, in commenting on quality management throughout BLS, noted that:

> "There is no Bureau-wide (or Office-wide) objective effort to measure the impact on final products of undetected errors, 'corrections' or imputations that later prove to be incorrect."
>
> "Quality management * * * will become and remain an integral part of the BLS program if and only if top management exercises the requisite oversight to assure that quality is not subordinated to production."
>
> "With the high dependence of important segments of public policy on BLS data and the potential impact of errors on the quality of life for large numbers of Americans, we can ill-afford to continue to exploit our charmed existence."

Soon thereafter, a Task Force was established to look specifically at quality assurance in the CPI area. It was evident to the Task Force that CPI quality assurance was inadequate. In the words of the group's Chairman:

> "The QA program should:
>
> 1. Maintain and improve the quality of the data collected * * *.

2. Insure the adequacy of the collection procedures and forms.

3. Identify areas of deficiencies * * *.

4. Provide a suitable format for promoting a high degree of confidence in the validity of the Bureau's statistics."

* * * * *

"It is the view of the program office that the present Q.A. program is not meeting the above goals."

Recognizing the need for a "comprehensive quality control effort to include all aspects of data collection, processing, and review," BLS outlined a general plan for implementing such a system in four phases over 4 years. The first phase called for work to begin on determining what data input points were to be measured and establishing priorities for measurement. BLS' plan also called for hiring a statistical consultant to help design the quality control system.

In its budget submission for fiscal year 1983, as sent to the Office of Management and Budget, BLS requested $5,461,000 to initiate a multiyear revision of the CPI which was to include the first phase of its quality control work.[1] According to a Resource Requirements Estimate prepared by BLS at the time it prepared its budget request, quality control work in fiscal year 1983 was expected to cost about $189,000 of which $100,000 was for the consultant. However, funds for the CPI revision, including the quality control aspects, were not included in the President's fiscal year 1983 budget as submitted to the Congress in February 1982.

In an August 1982 letter to the Chairman of the Joint Economic Committee, the Director of the Office of Management and Budget noted that:

---

[1] Other things BLS wanted to do as part of its CPI revision included the following: (1) update the market basket to reflect more recent consumer expenditure information, (2) reselect the sample of geographic areas in which CPI price data are collected to represent the urban population distribution as measured by the 1980 Census, and (3) modernize the computer system used to calculate the CPI.

"At present, there is no specific administration plan for funding a CPI revision. However, we recognize the importance of a well organized revision and we intend to work with the Commissioner of BLS to ensure that such a plan is completed, probably for inclusion in the FY 84 budget request."

In its budget submission for fiscal year 1984, BLS again included funds for CPI revision efforts including design of a quality control system. The Assistant Commissioner for Consumer Prices and Price Indexes told us that the current plan for establishing a quality control system contains more specifics than did the plan developed in support of the fiscal year 1983 budget request but that $200,000 still represents a reasonable estimate of the first year costs. We understand also that BLS no longer intends to rely heavily on a design consultant but instead expects to design the system in-house using consultant support only where necessary. The additional specificity in BLS' plan and the growing confidence in its in-house capabilities stem from the experience BLS has gained in developing a quality control system for the Producer Price Indexes. [2]

NO EVIDENCE THAT RELIABILITY OF NATIONAL CPI IS BEING COMPROMISED BY BAD DATA; LESS CONFIDENCE ABOUT LOWER LEVEL INDEXES

One reason it is difficult to make any strides toward improving quality control is the inability to demonstrate that the absence of adequate controls has had an adverse effect. If an automobile's reliability were to deteriorate due to inadequate quality controls, the effect would be readily apparent. But if the CPI's reliability were to deteriorate, how would one tell?

In many respects, the problem is one of the "Catch 22" variety. It is hard to generate a sense of urgency about the need for quality controls in the CPI area when there is no evidence that CPI data are unreliable; but to accumulate that kind of evidence, you almost have to have an effective quality control system--one that provides management with a reliable measure of quality.

---

[2] Producer Price Indexes measure average changes in prices received in primary markets of the U.S. by producers of commodities in all stages of processing--finished goods, intermediate goods, and crude materials.

Recognizing the problem and absent a feasible way to independently assess the quality of CPI data, we analyzed quality related information already available in BLS. Our purpose was not to scientifically assess CPI data quality but to see if we could get some feel for the extent to which errors might be creeping into the CPI data base--to see if there was any cause to feel the least bit uncomfortable about CPI reliability. Even though our results produced nothing to cause us to question the reliability of the national CPI or its major components, they did leave us less confident about specific item and local area indexes.

QA reinterview results

We analyzed the results of QA reinterviews conducted during fiscal year 1980 for the Commodities and Services Survey. We did not analyze the reinterview results for the Rent and Property Tax Surveys because those surveys account for a relatively small portion of the overall QA effort.

In analyzing the reinterview results, we were looking for errors which would or could affect the accuracy of the CPI. It should be remembered here that not every difference identified by a supervisor during a reinterview qualifies as an error. As we mentioned earlier, some differences are due to mistakes by the supervisor or to changes that occurred between the time the data collector visited the outlet and the time of the reinterview. The only differences we included in our definition of "error" were those that supervisors attributed to data collectors making mistakes or to respondents giving data collectors incorrect information.

With the above in mind and relying on advice from BLS personnel, we grouped errors into three categories: (1) those, such as misrecorded prices or improper substitutions, which would definitely have an impact on the index if they were not caught and corrected during the review process; (2) those, such as misspelled outlet names, that would have no impact on the index; and (3) those, such as a failure to fully or accurately describe the item being priced, which might affect the index under certain circumstances and might not under others.

In fiscal year 1980, BLS priced about 1,375,000 items in about 31,000 outlets for the Commodities and Services Survey. Supervisors checked 10,643, or less than 1 percent, of those items through QA reinterviews. Our analysis of those reinterviews showed that supervisors found at least one index impacting error in 4.2 percent of the items reinterviewed and at least one potential index impacting error in 14.9 percent of

the items (items with both actual and potential index impacting errors were included only in the actual category). As the following table shows, error rates varied greatly among BLS regions.

Region	Number of items checked	Items with index impacting errors		Items with errors that could affect the CPI		Totals	
		Number	Percent of items checked	Number	Percent of items checked	Number	Percent
Boston	715	69	9.7	63	8.8	132	18.5
New York	766	24	3.1	123	16.1	147	19.2
Philadelphia	484	7	1.4	71	14.7	78	16.1
Atlanta	2,203	81	3.7	270	12.3	351	15.9
Chicago	1,654	99	6.0	405	24.5	504	30.5
Dallas	655	9	1.4	16	2.4	25	3.8
Kansas City	1,672	75	4.5	297	17.8	372	22.2
San Francisco	2,494	82	3.3	343	13.8	425	17.0
Total	10,643	446	4.2	1,588	14.9	2,034	19.1

The error rates we computed are not indicative of the quality of CPI data in general nor of the relative quality of work in the various regions. As we mentioned earlier, BLS' QA Program does not represent a statistically valid sample. So the incidence of error identified during the reinterviews cannot be projected across the universe of CPI data. And the differences between regions can probably be attributed, in no small part, to the lack of standardization in the way reinterviews are performed. The relatively large error rates for Chicago, for example, could very well indicate better quality reinterviews rather than poorer quality data collection.

One can only surmise what the error rates might be if the QA Program were operating consistently across the country and if it were being administered like a true quality control vehicle (surprise, independent price verification, etc.).

Results of special analyses

Commodity analysts and others in BLS have, on occasion, conducted special reviews directed at assessing the accuracy of specific pieces of CPI data. Besides the QA reinterviews, those special reviews were the only useful source of CPI quality related data we could find within BLS.

In 1978 and 1982, for example, a commodity analyst conducted two reviews--one involving men's jeans, the other involving men's suits--which were directed at assessing whether identical items being priced in different areas of the country were being described the same on the pricing schedules. An item's description is important from a quality standpoint because it helps ensure that the same item is priced from month to month and because it is what the data collector uses to identify comparable items for substitution purposes. The analyst concluded, in both cases, that the data received from the field was "fairly good" but that there were some problems with the accuracy of the more technical price factors. For example, in 11 of 22 instances the data collectors had incorrectly described the jeans fabric and in 41 of 62 instances the data collectors had failed to specify the type of suit fabric.

BLS uses a scientific process to select outlets where CPI pricing is to be conducted and has established specific criteria that must be met before a selected outlet can be dropped from the program. An outlet can be dropped, for example, if the data collector finds, using guidance provided in the data collection manual, that the outlet is outside the pricing area. In 1979, BLS' Office of Survey Design looked at 332 outlets across the country that data collectors had identified as being outside the pricing area and found that the collectors were wrong in 176, or 53 percent, of the cases.

In an April 1981 memorandum to BLS' Assistant Commissioner for Field Operations, the Assistant Commissioner for Consumer Prices and Price Indexes noted that commodity analysts had found a number of recurring data collection problems with respect to fresh fruits and vegetables. He went on to say that:

> "Further research has resulted in the discovery that the problem is rather widespread throughout the various regions and involves a significant number of data collectors. I feel that there is sufficient evidence at this point that the problem is significant enough to warrant immediate attention in order to prevent any further impact on the accuracy of the data being used to calculate the CPI."

In January 1980, the Office of Survey Design expressed concern about the number of nonresponse outlets in the Rent Survey and asked that the field staff try to reduce the number. Of the 374 rental units that were supposed to be priced during the last 6 months of 1979 in one pricing area, for example, about 25 percent were coded as nonresponses.

There is cause to be concerned about lower level CPIs

Although we were able to document various CPI data errors, we were not able to identify any adverse effect on the CPI. We do not know (nor does BLS) what level of error it would take to cast a pall of unreliability on the national CPI figure that is reported in the media every month. We tend to believe, however, that the large amount of data that goes into constructing the national CPI makes it fairly tolerant to error and that the amount and kinds of errors discussed above would fall within acceptable limits. We are much less confident about lower level indexes and particularly the 28 city indexes that BLS publishes.

A city CPI's tolerance to error has to be much lower than the national CPI's because the amount of data that goes into constructing a local area index is considerably less than goes into constructing the national index. For example:

--A special BLS tabulation showed that while the September 1981 index figure for Women's and Girls' Apparel in the national CPI-U was computed using 1,682 price quotes, the figures for that same item in the local area indexes were supported by considerably fewer quotes--only 58 for the Philadelphia index, for example. Obviously, a couple of errors in the Philadelphia data would have much more of an effect on the reliability of the city index than on the reliability of the national index.

--A March 1980 BLS memorandum referred to an "* * * embarrassing incident [in which] half of the change in the Cincinnati medical care index was due to a change in prices for dental services for which we were pricing only one dentist."

BLS readily admits concern about the reliability of its local CPIs. For example, in its July 1981 response to questions about the CPI posed by a subcommittee of the Senate Appropriations Committee, BLS had several things to say about its local indexes, including:

--"In rough terms, the error on a given local area measure * * * is, on average, 5 times as great as the error on the U.S. * * * Index."

--"The local area CPI's are basically a by-product of the national CPI system and do not provide the same levels of reliability as the national program * * *. BLS is becoming increasingly concerned about the expanding uses to which these data are being put at the local level."

--"Clearly, in situations where relatively small changes in a statistic can trigger relatively large changes in dollar expenditures, a high degree of accuracy is needed. Most, if not all, of the local area measures, since they are based on small sample sizes, are subject to relatively larger sampling errors, and are therefore subject to somewhat more erratic movements on a short term basis. BLS cautions users on the use of local area data in escalation agreements."

Although BLS expresses concern about the use of local indexes, it is not effectively conveying that concern to users. Local CPIs are published via news releases issued by BLS' regional offices. In reviewing releases issued by each of the eight regions, we noticed that only two of the regions included a warning about the reliability of the local data. The Kansas City Regional Office's news release for January 1982, for example, included the following note:

--"Local area indexes are subject to more sampling and other measurement error than the national index because each, being a subset of the national index, has a smaller sample size. In light of the variability of local area indexes, users are advised to exercise caution in evaluating local data."

Although the kind of warning cited above puts the user on notice, its value is diminished by the absence of any specifics as to just how reliable the particular local index is. In that regard, another of the things BLS wants to do as part of the CPI revision discussed on page 24 is develop variance estimates which will quantify the sampling error in the CPI. In BLS' words:

--"The publication of variance estimates may provide for alteration of contractual arrangements using local area indexes since changes in these indexes may not be found to be statistically significant."

CONCLUSIONS

The importance of the CPI and the substantial amount of public and private sector money tied to movements in this one statistic make it imperative that BLS ensure CPI reliability through an appropriate system of quality control. BLS' current procedures do not provide that assurance. This is not a startling discovery on our part. BLS knows its controls are inadequate and has expressed a desire to improve matters. Unfortunately, development of a quality control system costs money.

We recognize that funds are scarce and that statistical agencies, including BLS, have had to eliminate or constrain various statistical series to meet budget cutbacks. We recognize also that in such a budgetary climate, it is easy to curtail quality controls or, as is the case here, to forego implementing controls because there is no visible bad effect. The reliability of a statistical series could surely deteriorate without controls: but who would know?

BLS' plan to design a quality control system (which could cost about $200,000) is a reasonable first step to ensuring CPI reliability and should be one of the CPI revisions approved for funding in fiscal year 1984. What happens after that and what costs are incurred beyond the $200,000 will depend on the type of quality control system designed and the extent to which the design is implemented. In that regard, BLS, the Office of Management and Budget, and the Congress need information that will provide a sound basis for making future funding decisions concerning CPI quality control. They will not get that kind of information without taking the first step--a design effort structured to (1) identify points in the CPI process that need to be controlled, (2) determine the kinds of controls that need to be implemented, (3) estimate the costs associated with those controls, and (4) assign priorities.

In a draft of this report sent to the Office of Management and Budget on December 7, 1982, we proposed that BLS' request for funds to design a quality control system for the CPI be approved. In commenting on our draft (see app. I), the Director of the Office of Management and Budget indicated that our proposal had been adopted. He said (and we verified) that the President's budget for 1984, which went to the Congress on January 31, 1983, contained a request for funds to start a major revision of the CPI which is to include "designing a

comprehensive quality control program to cover data collection, processing, and reviewing activities." We think the Congress should appropriate the funds for that design effort.

Even if quality control design efforts do get funded and everything proceeds smoothly, it could be several years before any quality control system is implemented. With that in mind, BLS needs to consider what it can do now, within the current budget environment, to help it keep better tabs on CPI data quality.

BLS could, for example, revise the procedures associated with its QA Program to strengthen the elements of control while continuing to operate the program as a regional tool for evaluating data collectors. In that respect, BLS could emphasize the value of surprise in assessing a data collector's work and could stress the importance of checking all aspects of a collector's performance including his or her handling of nonresponse outlets. Along those same lines, BLS could encourage supervisors to make more use of the telephone to verify a data collector's work. We are not suggesting that telephones take the place of onsite reinterviews and observations. What we are suggesting is that BLS would have better assurance of CPI data quality if supervisors were able to check on a data collector's work more than once a year. The telephone provides a relatively quick and inexpensive way to do just that.

There are things BLS can do in Washington, too. We think, for example, that management would be better served if it had summary information on the extent to which commodity analysts are questioning and/or revising field collected data. A quality check point, especially one as important as the technical review by commodity analysts, is not being used to full advantage if management does not get regular feedback on the check point's results. BLS might also consider checking on the adequacy of its quality control points in Washington by devising online tests to assess the system's effectiveness in identifying and handling bad information that had been purposely injected into the CPI data processing and review pipeline.

Our purpose in discussing the above procedures is to suggest how BLS might be able to improve its controls, if only minimally, within the current budget environment. It should be obvious that the kinds of procedures we are talking about, even in the aggregate, would in no way approximate a true quality control system. We think it is equally obvious, however, that BLS needs to do whatever it can now to better assure itself that CPI data are accurate. We are not saying that the above procedures represent the best course of action

in that regard or that they provide for the most efficient use of available funds. We offer them only as ideas for BLS' consideration.

Absent an effective quality control system, one can never be certain the CPI is founded on sufficiently accurate data. We saw nothing during our review, however, that caused us to doubt the reliability of the national CPI. But we are less confident about the local indexes, and so is BLS. We think BLS needs to do a better job of sharing with users its concerns about local indexes. The warning that two BLS regions are already including in their news releases at least puts users on notice, and, as a minimum, should be adopted by other regions. The warnings would be more meaningful, however, if they were less general; if they included more specific indicators of accuracy. Then the user of a particular index would have some basis for deciding whether that index is reliable enough for the use to which it is being put.

RECOMMENDATION TO THE CONGRESS

We recommend that the Congress approve funding requests to design a quality control system for the CPI.

RECOMMENDATIONS TO THE SECRETARY OF LABOR

We recommend that the Secretary of Labor direct BLS to:

--Assess all aspects of the CPI program to identify ways it can better ensure CPI data accuracy, within the current budget environment, until a formal quality control system is implemented. In so doing, BLS should consider the kinds of things discussed on page 32 of this report.

--Attach appropriate warnings to the local area CPIs so users are aware of BLS concerns about the reliability of those indexes.

AGENCY COMMENTS AND OUR EVALUATION

The BLS Commissioner, in commenting on a draft of this report for the Secretary of Labor (see app. II), noted that BLS

--will "continue to place the highest priority on insuring that the CPI is compiled in an accurate and timely fashion,"

--has "explicitly dedicated" some resources within the existing budget base to quality control,

--has undertaken "significant public education about the limitations of local area data in the CPI," and

--will develop "an appropriate statement" for all its CPI releases.

In discussing those comments with BLS' Assistant Commissioner for Consumer Prices and Price Indexes, we learned that all CPI news releases, starting with those reporting January 1983 price data, are to include the following note:

> "Local area CPI indexes are by-products of the national CPI program. Because each local index is a small subset of the national index, it has a smaller sample size and is, therefore, subject to substantially more sampling and other measurement error than the national index. As a result, local area indexes show greater volatility than the national index, although their long-term trends are quite similar. Therefore, the Bureau of Labor Statistics strongly urges users to consider adopting the national average CPI for use in escalator clauses."

The Assistant Commissioner also told us that BLS had established a Quality Management Working Group and a Quality Assurance Task Force (not the same task force mentioned on page 23). The Working Group will be responsible for designing and reviewing CPI quality controls. The Task Force's mission is to develop an interim QA program for the CPI. As part of its effort, the Task Force is to recommend minimum reinterview requirements, develop selection criteria for the types of outlets, goods, and services to be checked, and develop a new QA manual. Establishment of those groups, especially the Quality Management Working Group, is an important step in the right direction.

The BLS Commissioner, in further commenting on our draft, expressed concern that certain language, particularly our use of the term "inadequate controls," could seriously mislead readers. The Commissioner explained that (1) the CPI process "has multiple, built-in controls to ensure accuracy in compilation and publication," (2) what is lacking is a "formal quality control system that is both comprehensive and statistically based," and (3) the "implication that no control system at all exists is clearly erroneous."

We are not saying or implying that the CPI process is devoid of controls. We say just the opposite, in fact, on page 16. Also, because we only reviewed a part of the CPI process, albeit a major part, we have no way of knowing how adequately the rest of the process is being controlled. What we are saying is that BLS' controls are not sufficient to ensure the accuracy of the price data used to construct the CPI. With that in mind, we think it is fair to describe BLS' controls as inadequate. To avoid the possibility that our meaning might be misinterpreted, however, we have revised the cover summary, page ii of the digest, and page 16 of the report to describe the situation in more precise terms.

BLS also wanted us to recognize that "formal quality control of statistical surveys is not a well-established process which BLS can simply adopt" and that the $200,000 figure cited in the report represents initial startup costs only. We explained on page 31 that $200,000 is just the beginning and that future resource requirements would depend on the kind of system designed and the extent to which the design is implemented. In that regard, we recognize that development of a comprehensive system of quality controls will not be simple. The significant strides BLS claims to have made in developing controls in the Producer Price Index Program, however, would seem to provide a basis for more optimism about its prospects for the CPI than might have been possible a few years ago.

APPENDIX I

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

JAN 31 1983

Mr. Morton A. Myers
Director
Program Analysis Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Myers:

Thank you for the opportunity to comment on your draft report "Funds Should Be Provided to Develop Controls Over CPI Data Quality."

The draft contains a recommendation to the Congress and to the Office of Management and Budget to approve funding requests to design a quality control system for the Consumer Price Index.

The President's Budget for 1984 contains a request for funds to start a major revision of the CPI. The revision plan calls for designing a comprehensive quality control program to cover data collection, processing, and reviewing activities and for future publishing of estimates of non-sampling error.

Thank you again for the opportunity to comment.

Sincerely,

Joseph R. Wright, Jr.
Deputy Director

cc: Honorable Raymond Donovan
Secretary of Labor

APPENDIX II

U. S. Department of Labor

Commissioner for
Bureau of Labor Statistics
Washington, D.C. 20212

JAN 10 1983

Mr. Philip A. Bernstein
Director
Human Resources Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Bernstein:

This is in reply to your letter to the Secretary of Labor requesting comments on the draft GAO report entitled, "Funds Should Be Provided to Develop Controls Over CPI Data Quality." The Department's response is enclosed.

The Department appreciates the opportunity to comment on this report.

Sincerely yours,

Janet L. Norwood

JANET L. NORWOOD
Commissioner

Enclosure

U.S. Department of Labor's Response To
The Draft General Accounting Office Report
Entitled--

FUNDS SHOULD BE PROVIDED TO DEVELOP CONTROLS OVER CPI DATA QUALITY

Our comments on the proposed report on the need for controls over CPI data quality deal with specific points made in the findings portion as well as with the report's recommendations.

Our most serious concern is that certain language in the report could seriously mislead readers. In particular, the use of the term "inadequate controls" in the cover summary and pages ii and 16 suggests a more negative situation than that warranted by the more measured expression of the report in general. The CPI process has multiple, built-in controls to ensure accuracy in compilation and publication. What is lacking is a formal quality control system that is both comprehensive and statistically based. We believe that such a formal control system is highly desirable. Such a system could enhance and replace portions of the existing control system, but the implication that no control system at all exists is clearly erroneous. (See GAO note 2.)

In addition to the extensive control processes built into the CPI, BLS has long recognized the appropriateness of a formal quality control program. Efforts to fund such a program go back over the last decade, well beyond the FY '83 budget cycle cited in the proposed report.

It is also important to recognize, however, that formal quality control of statistical surveys is not a well-established process which BLS can simply adopt. Our research and experience in this area indicate that successful application of quality control processes to statistical surveys is rare, often not germaine, and that new as well as conventional approaches need to be used. Development of comprehensive formal quality control for the CPI requires substantial investment in an essentially new field. Finally, it should be realized that as important as a formal quality control program is for the CPI, it is only one of a number of competing tasks that are needed to improve the quality of the CPI, such as more up-to-date samples and expenditure weights, improved methods for selecting and pricing substitutes, pricing unpriced items, and replacing out-dated software.

GAO Note: Page references in this appendix have been changed to agree with the page numbers in the final report.

For completeness in discussing budget issues related to formal quality control, it might be useful for the report to highlight a few other facts. One is that the $200,000 figure represents an initial start-up cost. Additional funding for installation and operation of the program would be required. Additional funding would also be required for variances.

On page 6, there is reference to about 400 items for which BLS collects prices. Presumably this reference relates to the 382 Entry Level Items (ELI's) used in the calculation of the index. The expenditure weights in the index are held constant at the level of 265 so-called item strata. Within these strata, the proportion of price quotations is fixed for ELI's. The term "item" as used in the CPI really relates to classes of items, not unique products as the casual reader might assume. The 265 item strata constitute an exclusive and exhaustive classification of all consumption expenditures. A possible substitute for that sentence might be, "Using that information, BLS classified all consumption expenditures into 265 categories. For each of these categories, BLS prices a number of different individual specifications selected within each urban area included in the CPI area sample." (See GAO note 2.)

Although the last sentence in the second paragraph on page 19 is theoretically correct, it does over simplify the real world. A comprehensive quality control program would include a valid sampling method for selection of outlets to be included for reinterview or other quality control; nevertheless, it would still be necessary to balance the relative impacts on the overall quality of the index between potentially losing survey respondents, with resultant index bias, and biasing the quality control sample. (See GAO note 3.)

Response to Recommendations:

The proposed report recommends that BLS be directed to "Assess all aspects of the CPI program to identify ways it can better assure CPI data accuracy, within the current budget environment, until a formal quality control system is implemented."

The BLS has and will continue to place the highest priority on insuring that the CPI is compiled in an accurate and timely fashion. Within the existing budget base, some resources are being explicitly dedicated to quality control. BLS is constantly trying to improve index procedures, samples and concepts but has not been successful in securing

GAO Note 2: The language in question has been revised.

GAO Note 3: This is a valid concern--one that warrants BLS' attention during its design of a quality control system.

funds to carry out a comprehensive quality control program. As briefly noted in the report, BLS has made significant strides in quality control methods and procedures in the Producer Price Index (PPI) program, where resources were made available under the PPI revision program. Building on that expertise, a comprehensive quality control plan is being developed for the CPI. Following that planning effort, BLS expects to undertake specific process studies of quality. New on-going controls and additional process studies will be implemented as test results dictate and resources are made available.

The proposed report also recommends that BLS "Attach appropriate warnings to the local area CPI's...." BLS has undertaken significant public education about the limitations of local area data in the CPI and other program areas. The Bureau will continue these efforts and will develop an appropriate statement for all its CPI releases. BLS believes, however, that the expansion of local area samples and the calculation of index variances--unfortunately, costly projects for which funds are not currently available--would be of even greater importance to users of local CPI's.

24683

AN EQUAL OPPORTUNITY EMPLOYER

UNITED STATES
GENERAL ACCOUNTING OFFICE
WASHINGTON, D.C. 20548

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE,$300

POSTAGE AND FEES PAID
U. S. GENERAL ACCOUNTING OFFICE

THIRD CLASS