Livingston Survey Documentation

Federal Reserve Bank of Philadelphia

May 2010

The Livingston Survey data set contains forecasts of economic variables from a survey of forecasters. The survey was begun in June 1946 by Joseph Livingston, a columnist for the *Philadelphia Inquirer*. Livingston continued the survey until his death in 1990, at which time the Federal Reserve Bank of Philadelphia took it over.

The survey is conducted twice a year, in June and December. The survey newsletter, consisting of the median forecast of each variable, is available from the Federal Reserve Bank of Philadelphia.[1] The survey consists of forecasts of 18 different variables describing national output, prices, unemployment, and other macroeconomic data. Files containing the mean and median responses, as well as the individual responses, are available on the Philadelphia Fed's web page.

Section 1 discusses how we have reorganized the survey's data set. The data set is now much easier to use. Section 2 summarizes the structure of the survey. Section 3 documents how the data are organized on our web page. Section 4 discusses how to construct growth rates from the forecasted levels and our base values. Section 5 discusses the survey's forecasts for the Consumer Price Index, one of the most important variables in the survey. Please read Section 6, which discusses very important features of each variable in the survey.

1. Reorganization of the Data Set

In June 2004, in response to numerous problems with the structure of the Livingston Survey data set, we decided to change the way in which we present the data on our web page. This section describes the changes.

Previously, we presented the data in text files, using a complicated nomenclature system to label the variable and forecast horizon. Often, the labels did not appear directly above the data they described, and the labels differed from those described in the documentation. This caused users quite a bit of confusion and increased the cost of working with the data. We now simplify the presentation considerably by using Excel workbooks, and we have abandoned the old, complicated nomenclature system. Users of the old Livingston data will remember that each variable had a root mnemonic that was prefixed to a code that described the forecast horizon. For example, the forecasts for the level of the CPI were labeled as CPIZ (the monthly base value), CPI0 (the forecast for the month of the survey), CPI6 (the six-month-ahead forecast), CPI12 (the 12-month-ahead forecast), CPIJZ (the annual base value for a June survey), CPIJ0 (the annual average forecast for the year of the survey), and CPI1J (the annual average for the next year). The new system features the following improvements:

[1] The first survey to use the median forecasts in the write-ups was the survey of June 2000. All previous write-ups, including those written by Joseph Livingston, used the mean forecast.

- All median forecasts are stored in an Excel workbook containing multiple worksheets, corresponding to the variable being forecast. (The mean forecasts are stored in a different workbook.) The worksheets are labeled descriptively, to indicate the variable being forecast. For example, the median CPI forecasts are stored in the workbook *Median.xls*, in the worksheet labeled, *CPI*.

- There is one worksheet for each variable, and each worksheet uses the same generic column headers to indicate the forecast horizon—for example, the column header *Forecast0Month* appears in each worksheet and denotes the forecast for the month (or quarter) of the survey and the column header *Forecast6Month* denotes the six-month-ahead (or two-quarter-ahead) forecast.

- Problems with misaligned column headers have been eliminated.

- We have eliminated the confusing distinction between the column headers used for the annual forecasts in a June survey and those for a December survey.

2. Overview of the Livingston Survey

The survey consists of forecasts of 18 different variables describing national output, prices, unemployment, and other macroeconomic data. It features forecasts for quarterly variables (real GNP/GDP, nominal GNP/GDP, real nonresidential fixed investment, and corporate profits after taxes), variables available at a monthly frequency (industrial production, total private housing starts, the PPI, the CPI, unemployment, average weekly earnings, retail sales, and automobile sales), and variables available at a daily frequency (prime rate, 10-year Treasury bond rate, 90-day Treasury bill rate, and S&P 500 stock price index).

For variables available **quarterly**, the forecast dates are:

- The value of the variable in the current quarter (beginning with the June 1992 survey);
- The value of the variable two quarters from now;
- The value of the variable four quarters from now;
- In June surveys, two annual average forecasts:
  - For the current year;
  - For the following year.
- In December surveys, three annual average forecasts:
  - For the current year (beginning with the June 1992 survey);
  - For the next year;
  - For the year after.

For variables **monthly**, the forecast dates are:

- The value of the variable in the current month (beginning with the June 1992 survey);

- The value of the variable six months from now;
- The value of the variable 12 months from now;
- In June surveys, two annual average forecasts:
    - For the current year;
    - For the following year.
- In December surveys, three annual average forecasts:
    - For the current year (beginning with the June 1992 survey);
    - For the next year;
    - For the year after.

For variables available **daily**, the forecast dates are:

- The value of the variable at the end of the current month (beginning with the June 1992 survey);
- The value of the variable at the end of the month six months from now;
- The value of the variable at the end of the month 12 months from now;
- In a June survey: the value of the variable at the end of December of the next year;
- In a December survey: the value of the variable at the end of December two years from now.

Since June 1990 (excluding December 1990), two other questions ask for long-term forecasts: the annual average CPI inflation rate and the annual average real GDP growth rate for the next 10 years.

Two Examples. In the survey of June 1992, we asked for forecasts for the quarters and months shown in the table below:

Forecasts in the June 1992 Survey

Variable Frequency	*High Frequency Forecasts*	*Annual Forecasts*
Quarterly	1992:Q2, 1992:Q4, and 1993:Q2	1992 and 1993 (annual average)
Monthly	1992:June, 1992:Dec, and 1993:June	1992 and 1993 (annual average)
Daily	1992:June, 1992:Dec, and 1993:June (last day of the month indicated)	1993:Dec (last day of the month indicated)

In the survey of December 1992, we asked for forecasts for the quarters and months shown in the table below:

Forecasts in the December 1992 Survey

Variable Frequency	*High Frequency Forecasts*	*Annual Forecasts*
Quarterly	1992:Q4, 1993:Q2, and 1993:Q4	1992, 1993, and 1994 (annual average)
Monthly	1992:Dec, 1993:June, and 1993:Dec	1992, 1993, and 1994 (annual average)
Daily	1992:Dec, 1993:June, and 1993:Dec (last day of the month indicated)	1994:Dec (last day of the month indicated)

Any researcher seriously thinking about using the data from the Livingston survey should read John Carlson's article "A Study of Price Forecasts."[2] Carlson's article explores the pros and cons of the Livingston Survey's data set. This document also contains a section that examines the Consumer Price Index in great detail and explains some of the known problems with this particular variable.

[2] Carlson, John. "A Study of Price Forecasts," *Annals of Economic and Social Measurement*, Vol. 6, No. 1, Winter 1977, pp. 27-56.

For further information about the Livingston Survey, please contact:

Calvin Price and Tom Stark
Research Department
Federal Reserve Bank of Philadelphia
Ten Independence Mall
Philadelphia, PA 19106 – 1574
E-mail: phil.liv@phil.frb.org

3. File and Variable Descriptions

Three Excel workbooks contain the survey's responses. Each workbook contains 16 worksheets, and each worksheet uses the same generic column headers to describe the forecasts. There is one worksheet for each variable, and the worksheet contains a different forecast step in each column (described by the column header). This section describes the workbooks, the worksheets, and the column headers.

The workbooks and a short description are shown in the table below.

Description of Livingston Survey Files (Excel Workbooks)

Livingston Survey Workbook Name	*Description of Contents*
Means.xls	Time series of mean responses.
Medians.xls	Time series of median responses.
RawData.xls	Panel data of individual responses.

Each of the Excel workbooks listed above contains 16 worksheets, named to describe the variable being forecast, as shown in the table below.

Description of the Worksheets in Each Workbook

Worksheet Name	*Description*	*Important Comments*
Quarterly Variables		
RGDPX	Real Gross Domestic Product	Billions of real dollars. Prior to 1992, real GNP.
GDPX	Nominal Gross Domestic Product	Billions of dollars. Prior to 1992, nominal GNP.
BFIX	Real Business Fixed Investment	Billions of real dollars. Prior to 1990, new plant and equipment expenditures.
CPAT	Corporate Profits After Tax	Billions of dollars. With inventory valuation adjustment (IVA) and capital consumption adjustment (CCAdj).
Monthly Variables		
IP	Industrial Production Index	Index level.
TPHS	Total Private Housing Starts	Millions of units. Annual rate. Prior to 1969, nonfarm housing starts.
PPI	Producer Price Index for Finished Goods	Seasonally adjusted index level. Prior to December 2004, not seasonally adjusted index level.
CPI	Consumer Price Index	Seasonally adjusted index level. Prior to December 2004, not seasonally adjusted index level.
UNPR	Civilian Unemployment Rate	
WMFG	Average Weekly Earnings in Manufacturing	Dollars. Not seasonally adjusted.
RTTR	Retail Trade	Billions of dollars. Monthly rate.
AUTODF	Auto Sales, Domestic and Foreign	Millions of units. Annual rate.
Daily Variables		
PRIME	Prime Interest Rate	
TBOND	Rate on 10-Year Treasury Bond	Prior to the survey of December 2002, rate on 30-year Treasury bond.
TBILL	Rate on 3-Month Treasury Bill	
SPIF	S&P 500 Stock Price Index.	Index level.

The worksheets share the following generic column headers, which describe the horizon of the forecast values in the column or, in some columns, the historical values.

Column Headers in the Worksheets[3]

Column Header	*Description*
Date	Date of the survey (either June or December).
BasePeriod	Last quarterly or monthly historical value known at the time the survey questionnaire was mailed.
Forecast0Month	Forecast for the month or quarter in which the survey is taken (for monthly variables: June or December; for quarterly variables: Q2 or Q4).
Forecast6Month	Forecast for the month that is six months ahead or for the quarter that is two quarters ahead of the survey date.
Forecast12Month	Forecast for the month that is 12 months ahead or for the quarter that is four quarters ahead of the survey date.
BaseYear	Last annual-average historical value known at the time the survey questionnaire was mailed.
Forecast0Year	Annual average forecast for the year in which the survey is taken.
Forecast1Year	Annual average forecast for the next year.
Forecast2Year	Annual average forecast the year after.
Forecast10Year	Forecast for annual average growth over the next 10 years (RGDPX and CPI worksheets only).

Two additional columns appear in the panel data set of individual responses (RawData.xls). The header *ID* gives a number that identifies each individual. The same ID number is used across surveys, so you can use it to track how a forecaster's projection changes over time. The header *Category* classifies the forecaster's affiliation, according to the codes listed in the table below.

[3] See the section below on "Growth Rates and Base Values" for a detailed description of some problems with the variables, *BasePeriod* and *BaseYear*. For variables available at a daily frequency (PRIME, TBOND, TBILL, and SPIF), the forecasts and base values are those for the last day of the indicated period. See the tables above ("Forecasts in the June 1992 Survey" and "Forecasts in the December 1992 Survey") for two examples of how the nomenclature system works.

Participants' Affiliation Codes in RawData.xls

Code	*Affiliation*
A	Academic Institution
B	Commercial Banking
C	Consulting
F	Federal Reserve
G	Government
T	Industry Trade Group
R	Insurance Company
I	Investment Banking
L	Labor
S	Nonfinancial Business
U	Other / Unknown

The affiliation codes for each participant change over time. Before the survey of June 1981, the participants have the same code throughout their association with the survey. If a participant worked for an investment banking group in the 1960s and then acquired employment in commercial banking in the 1970s, the survey's code lists this participant as an investment banker. We apologize for this inconsistency in the data, but the past records of the participants are incomplete. The codes after June 1981 capture the participants' changing affiliations over time.

Note that in the survey of December 2009, two new category codes were introduced: *Consulting* and *Industry Trade Group*. Panelists were allowed to retroactively switch their category code into one of these categories for previous surveys. The panelists who have done so correspond to ID numbers 264, 366, 383, 384, 394, 427, and 444. The currently available data set reflects these changes.

4. Growth Rates and Base Values

The Livingston survey data set contains forecasts for the *levels* of variables. For many purposes, however, forecasts for growth rates are better. If you are interested in computing the implied forecast for the annualized growth of a variable between two forecast dates, say between six

months ahead of the survey date and 12 months ahead for industrial production, you can use a formula such as

$$\left[\left(\frac{Forecast12Month}{Forecast6Month}\right)^{\frac{12}{6}}-1\right]*100$$

where *Forecast12Month* is the forecast for the level 12 months after the survey date, and *Forecast6Month* is the forecast for the level six months after the survey date. You may choose, of course, to use either the mean forecast or the median forecast. Beginning with the survey of June 1992, we also provide forecasts for the level of the variable in the month of the survey (June or December). So you could compute an implied forecast for annualized growth between the month of the survey and, say, 12 months later, using a formula such as

$$\left[\left(\frac{Forecast12Month}{Forecast0Month}\right)^{\frac{12}{12}}-1\right]*100$$

where *Forecast0Month* is the forecast for the level of the variable in the month of the survey. In a June survey, *Forecast0Month* is the forecast for the level in June. In a December survey, it is the forecast for the level in December.

If you are interested in computing the implied rate of forecasted growth in the annual average index level between the year of the survey and the next year, a formula such as

$$\left[\left(\frac{Forecast1Year}{Forecast0Year}\right)-1\right]*100$$

will do the trick. In this formula, *Forecast0Year* and *Forecast1Year* are the forecasts for the annual average level of the industrial production index in the year of the survey and the next year, respectively.

The survey's data set also provides information on the last historical monthly (or quarterly) value known. The last known annual values are also provided. These are the variables denoted *BasePeriod* and *BaseYear*, respectively. With these base values, one could compute an implied forecast for annualized growth over the period between the period of the last known historical value and the period of a forecast value, using a formula such as

$$\left[\left(\frac{Forecast6Month}{BasePeriod}\right)^{\frac{12}{k}}-1\right]*100$$

where *k* represents the number of periods between *BasePeriod* and *Forecast6Month*.

In the previous version of this documentation, we described the base values as the last historical values known at the time the survey questionnaire was mailed. For example, in a June survey, the questionnaires are mailed in May, and the last known values are those for April (monthly variables) or Q1 (quarterly variables). For the daily variables, these values would be those for the last day in April. In a December survey, the questionnaire is mailed in November, and the base values would be those for October or Q3.

However, in the process of reorganizing the data set, we discovered many discrepancies between our description of the base values and the actual base values listed in the data set. The only exception is the case of the CPI. For this variable, we have verified that the base values correspond to the value of the CPI in April (June survey) and October (December survey).

A Caveat. The Philadelphia Fed offers the following advice concerning the base values in the data set (*BasePeriod* and *BaseYear*) corresponding to all surveys through the survey of December 2003: You should not use these values to compute a forecast growth rate**.** The only exception is the CPI: For this variable, the base values are accurate.

Data Set Partitioning. For the reasons (and some others) described in this section, we are partitioning the data on our web page into data before the survey of June 2004 and data after. For all surveys starting with June 2004, the base values (*BasePeriod* and *BaseYear*) are the last historical values known at the time the survey questionnaires are mailed. The questionnaire for a June survey is mailed in May, and the December questionnaire is mailed in November. That means the base values are defined as the values for the months, quarters, and years shown in the table below.

Base Values in Surveys from June 2004 - present

Survey Date	*BasePeriod*		
	Daily Variable	Monthly Variable	Quarterly Variable[4]
June Survey	Last Day of April	April	Q1
December Survey	Last Day of October	October	Q3

The annual base values (*BaseYear*) are those for the annual average over the year preceding the year of the survey (monthly and quarterly variables) or for the last day of that year (daily variables).

[4] Because of a one-month delay in the release of data on corporate profits, the quarterly base value for this variable is that for the quarter preceding the quarter shown in the table. In a June survey, this is Q4; in a December survey, this is Q2.

5. Discussion of the Consumer Price Index (CPI)

The CPI variable is the most requested variable of the Livingston Survey's data set; it is, therefore, also the most scrutinized. Several minor inconsistencies characterize this variable. The problems usually arise because Joseph Livingston requested forecasts for CPI levels rather than inflation rates. Users should also know that these forecasts are for the not-seasonally-adjusted level of the CPI in surveys prior to December 2004. (The same is true for the forecasts of the Producer Price Index [PPI].) Beginning with the survey of December 2004, we asked the panelists to provide forecasts for the seasonally adjusted index level.

The first concern revolves around the data that the economists used in the completion of the survey. (See also the section, "Growth Rates and Base Values," for a related discussion.) On every survey response form, Livingston would list the most recent monthly or quarterly data to assist economists in making their forecasts. For the June surveys, the CPI figure was either the March or April number, and for the December surveys, the CPI figure was either the September or October number. One criticism of the survey is that Mr. Livingston never asked the participants the base data they used to make their forecasts: the printed data on the survey forms (supplied by Livingston) or more recent data. To construct the inflation rates from forecasts in levels, some assumptions had to be made about which monthly figure the individual participants used at the time they made their forecasts. It was decided that economists used the April CPI figure for the June surveys and the October CPI figure when making predictions for the December survey. (For further details on these choices, see the aforementioned article by Carlson.)

The CPI base data were collected from the *Survey of Current Business* (SCB) in the month and year that the survey was sent to the economists. For the June surveys, the May SCB was used, and for the December surveys, the November SCB was used.

For surveys after December 1991, participants are asked to give the current-month (June and December) forecasts and then base their six-month and 12-month forecasts on their current-month predictions.

The second problem was with rounding. In some years, Livingston rounded the CPI numbers to the nearest one (338.5 would be rounded to 338). Since these rounded CPI figures were listed on the survey forms, the forecasters used rounded CPI figures when making their forecasts. The surveys in which we are positive the CPI figures were rounded run from the June survey of 1985 to the June survey of 1988, a total of seven surveys. There were probably more surveys in which Livingston used rounded data. We can say with confidence that the rounded data affected only a small percentage of the surveys.

6. Special Notes on the Variables

In this section, we describe any special features of the variables in the survey.

GDPX (Nominal Gross Domestic Product)

- Billions of dollars. Seasonally adjusted. Annual rate.
- Prior to 1992, this was nominal gross national product.
- This variable was omitted from the following surveys: June 1950, December 1950, June 1951.

RGDPX (Real Gross Domestic Product)

- Billions of real dollars, with varying base years. Seasonally adjusted. Annual rate.
- Prior to 1992, this was real gross national product.
- Included in the survey since June 1971.
- Forecasts for the 10-year annual average rate of growth included since June 1990.

BFIX (Real Business Fixed Investment)

- Billions of real dollars, with varying base years. Seasonally adjusted. Annual rate.
- Included in the survey since December 1990.

CPAT (Corporate Profits After Tax)

- Billions of dollars. Seasonally adjusted. Annual rate.
- Included in the survey since June 1971.
- Prior to June 2006, this variable is corporate profits after tax, *without* inventory valuation adjustment (IVA) and capital consumption adjustment (CCAdj).. We have, however, noticed some surveys in which there appears to have been an arbitrary change, to the after-tax measure that <u>includes</u> IVA and CCA. From the survey of June 2004 to the survey of December 2005, this is the measure that <u>excludes</u> IVA and CCA
- From June 2006 forward, the variable for corporate profits after tax will **<u>include</u>** inventory valuation adjustment (IVA) and capital consumption adjustment (CCAdj). This change has been implemented in response to requests by both panelists and users of the data.
- The quarterly base value (*BasePeriod*) for the quarter preceding the survey month is not available at the time the questionnaire is mailed to the participants. For this variable, we use, instead, the value in the quarter that is two quarters prior to the quarter in which the survey is taken—that is, Q4 in a June survey and Q2 in a December survey.
- Because we use the measure that excludes IVA and CCA, the historical values of this variable are subject to large, discrete jumps when there is a change in tax law affecting depreciation provisions. For information on a recent example, see the Bureau of Economic Analysis's press release of May 26, 2005, available at www.bea.gov. The time series of projections for this series in the Livingston Survey may or may not capture the jump in the historical values, depending on whether the forecasters anticipated the change in tax law.

IP (Industrial Production Index)

- Index level, with varying base years.

TPHS (Total Private Housing Starts)

- Millions of units. Seasonally adjusted. Annual rate.
- Prior to 1969, nonfarm housing starts.

PPI (Producer Price Index for Finished Goods)

- Index level, with varying base year. Not seasonally adjusted in surveys prior to December 2004. Seasonally adjusted beginning with the survey of December 2004.

CPI (Consumer Price Index)

- Index level, with varying base year. Not seasonally adjusted in surveys prior to December 2004. Seasonally adjusted beginning with the survey of December 2004.
- Forecasts for the 10-year annual average rate of growth included since June 1990.

UNPR (Civilian Unemployment Rate)

- Percentage points. Seasonally adjusted.
- Included since June 1961.

WMFG (Average Weekly Earnings in Manufacturing)

- Dollars. Not seasonally adjusted.

RTTR (Retail Sales)

- Billions of dollars. Seasonally adjusted. Monthly rate.
- Included since December 1959.

AUTODF (Auto Sales, Domestic and Foreign)

- Millions of units. Seasonally adjusted. Annual rate.
- Included since December 1966.

PRIME (Prime Interest Rate)

- Percentage points.
- Included since June 1981.
- Forecasts are for the last day of the period, not monthly averages.
- Beware: Early in the history of this variable, we have noticed some surveys in which the annual forecasts are for the annual average over the year, not the last day of the year. In the surveys beginning with June 2004, the annual forecasts are for the last day of the year.

TBOND (Interest Rate on 10-Year Treasury Bonds)

- Percentage points.
- Included since June 1992.
- Prior to the survey of December 2002, this was the rate on 30-year Treasury bonds.

TBILL (Interest Rate on Three-Month Treasury Bills)

- Percentage points.
- Included since June 1992.

SPIF (S&P 500 Stock Price Index)

- Index level.
- Included since December 1990.