Table of Contents

United States
Department of
Agriculture

Food and
Nutrition
Service

3101 Park Center Drive
Alexandria, VA 22302

FORECASTING FOOD STAMP PROGRAM PARTICIPATION AND BENEFITS

Mark Dynarski
Anuradha Rangarajan
Paul Decker

A product of
Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08536

August 1991

Table of Contents

United States Department of Agriculture

Food and Nutrition Service

3101 Park Center Drive
Alexandria, VA 22302

AUG 1 9 1991

The Honorable Quentin N. Burdick
Chairman
Subcommittee on Agriculture, Rural Development,
and Related Agencies
Committee on Appropriations
United States Senate
Washington, D.C. 20510

The Honorable Jamie L. Whitten
Chairman
Subcommittee on Rural Development, Agriculture,
and Related Agencies
Committee on Appropriations
U.S. House of Representatives
Washington, D.C. 20515

Dear Mr. Chairmen:

Over the last 2 years ending this past April, the number of persons participating in the Food Stamp Program (FSP) has risen by 4 million. Although the magnitude of this increase is not unprecedented, it began in a period when neither changes in the program nor unemployment could account for the increase. In fact, the unemployment rate had declined through 1989 to a low of 5.1 percent, and even though the rate began to rise in 1990, it never approached the historic level associated with earlier peaks in FSP participation. As a result, existing forecasting models did not predict the sizeable increase in participation that occurred.

This led the Food and Nutrition Service (FNS) to seek an independent assessment of our existing forecasting procedures and recommendations to improve our forecasts of participation and benefits.

To conduct the evaluation, we contracted with Mathematica Policy Research, Inc. (MPR), a firm that is widely regarded as one of the preeminent social science research organizations in the country. They brought to this task both nationally recognized technical skills and an extensive knowledge of and insight into the Food Stamp Program.

We took several important actions to ensure that the study results reflected the advice of a broad range of experts and incorporated state-of-the-art forecasting techniques. In the spring of 1990, MPR and FNS convened a symposium of practitioners and scholars from Federal agencies, Congressional staff and the research community to explore this issue. MPR supplemented their own extensive resources and engaged two additional independent consultants: Dr. Burt Barnow, a senior researcher at Lewin/ICF and a distinguished practitioner in the fields of labor and welfare economics and econometrics; and Dr. Nancy Kirkendall, Chief Mathematical Statistician with the U.S. Department of Energy, Adjunct Professor in the Department of Statistics at George Washington University, and a nationally recognized expert in pure time-series models.

FNS also sought further independent reviews of the analysis from three nationally recognized scholars in the field of applied econometrics and time-series analysis: Dr. John Geweke, Professor of Economics at the University of Minnesota and a Consultant with the Research Department of the Federal Reserve Bank of Minneapolis; Kenneth D. West, Professor of Economics at the University of Wisconsin; and Sidney Saltzman, Professor of Planning and Regional Science at Cornell University. These scholars reaffirmed our opinion that the report met high and accepted standards for the application of econometric methods to a complex and difficult forecasting problem.

Finally, all stages of this major effort benefitted from the program and technical expertise of senior economists in the Economic Research Service of the Department of Agriculture.

We believe that this study successfully met all the objectives we set forth. We have learned that FNS forecasting procedures meet credible standards but that our models and procedures for monitoring and diagnosing the quality of forecasts can be modestly improved. We obtained the assurance that, without the benefit of hindsight, no alternative model could have projected the 1989 turning point in participation. This stems from both the limitation of econometric models to detect new relationships without the benefit of more time-series data and our new understanding from a companion study on participation (<u>A Study of the Increase in Food Stamp Program Participation Between 1989 and 1990</u>) that many factors, of differing emphasis in different States and at different times over the recent period, caused the increase in food stamp participation.

As the 1990 Economic Report of the President emphasized, "forecasting is an imprecise science. Unanticipated events with economic consequences ... occur from time to time. In addition, the reactions of businesses and households to changes in economic conditions or policy may shift over time." Thus, good forecasting is a dynamic process of ongoing development and refinement. The work produced under this study provides us with an enhanced knowledge of statistical tools and practices, as well as many invaluable insights for meeting the challenge of both current and future forecasting needs. Among the noteworthy recommendations are a mechanism for measuring the quality of forecasts and a process for creating a historical record of our progress in updating and testing new models.

Sincerely,

Betty Jo Nelsen

Betty Jo Nelsen
Administrator

Table of Contents

ACKNOWLEDGMENTS

This report has benefited greatly from the comments of many individuals. In particular, we wish to thank Christine Kissmer, Marian Lewin, Robert Dalrymple, Steven Carlson, and Christy Schmidt (the Food and Nutrition Service, Office of Analysis and Evaluation), Catherine Squires and David Burr (the Food and Nutrition Service, Program Information Division), David Smallwood and William Levedahl (U.S. Department of Agriculture, Economic Research Service), Thomas Fraker and Walter Corson (Mathematica Policy Research, Inc.), Nancy Kirkendall (U.S. Department of Energy), Burt Barnow (Lewin/ICF, Inc.), John Geweke (University of Minnesota), Kenneth West (University of Wisconsin), and Sidney Saltzmann (Cornell University). West Addison provided programming support. The report was produced by Denise Dunn, Monica Capizzi, and Debra Jones, and was edited by Thomas Good.

MPR Project Number: 7890-015
FNS Contract Number: 53-3198-9-31

CONTENTS

CONTENTS (continued)

Table of Contents

TABLES

TABLES (continued)

FIGURES

EXECUTIVE SUMMARY

The Food Stamp Program (FSP) enables low-income households to achieve and maintain nutritious diets by increasing their food purchasing power. Over time, the number of program participants can change, often dramatically, in response to the legislative tightening and broadening of eligibility requirements, fluctuations in economic activity and changes in the structure of the economy, State and Federal improvements in the accessibility of program benefits, and changes in the behavior of households.

The sensitivity of participation and benefits to economic conditions and to program changes poses a challenge for budgetary planning. The Food and Nutrition Service (FNS) submits budget estimates of future program benefits to the Office of Management and Budget at the end of the first quarter of each fiscal year. Future program benefits are based on the forecasts of an econometric model that links participation and benefits to macroeconomic conditions and to variables that indicate program changes.

In late 1989, FSP participation began to grow. The remarkable feature of this growth is that it occurred at a time when neither the unemployment rate, which had been declining, nor major changes in the program could account for the increase. Consequently, the existing forecasting models did not predict the sizeable increase that occurred. Thus, FNS sought an independent assessment of its existing forecasting procedures, as well as recommendations for improving its forecasts of participation and benefits.

CHANGES IN FSP PARTICIPATION AND BENEFITS OVER TIME

Since 1977, the number of persons who receive food stamp benefits in a given month has fluctuated between 14 and 22 million (excluding Puerto Rico). The number of participants was 14 million in 1978 prior to the elimination of the purchase requirement, which was a feature of the FSP up to that time. After the purchase requirement was eliminated by the Food Stamp Act of 1977 (PL 95-113)--a change that was implemented fully by January 1979--participation increased sharply, to 16 million in the second quarter of fiscal year 1979. Participation peaked at 22 million in 1983 during a deep economic recession, and fell almost continuously from 1983 to 1989. Beginning in 1989, participation increased at a rapid rate, rising to 23 million in April 1991.

The average value of food stamp benefits per participant has also varied since January 1979. However, in contrast to participation, average benefits have increased steadily since 1980. Beginning at $33 per month in the first quarter of fiscal year 1980, average monthly benefits increased to $52 by the first quarter of fiscal year 1989. The increase in monthly benefits reflects regular cost-of-living adjustments that increase benefits annually, the effect of legislative changes, and the lower countable incomes of recipients. After adjustment for inflation, average monthly benefits increased by $4 over the nine-year period.

FORECASTING FSP PARTICIPATION AND BENEFITS

Changes in household circumstances and program features over time give rise to changes in participation at the aggregate level. In turn, changes in household circumstances are driven by macroeconomic and demographic factors. If relationships between household decisions to receive food stamps and underlying factors could be estimated, the number of FSP participants could be forecast for various values of the underlying factors. However, the relationships between household decisions, program features, and macroeconomic and demographic factors are likely to be highly complex and not amenable to direct modeling.

The forecasting approach used in this report is based on multivariate statistical models of aggregate participation and benefits, whose parameters were estimated with least-squares regression techniques. Such models play a prominent role in many forecasting contexts, and are a key tool for government agencies and businesses that require forecasts of important variables for planning and budgeting purposes.

However, for two reasons, forecasting participation is more problematic than forecasting macroeconomic quantities, such as Gross National Product or personal income. These factors suggest that expectations about the accuracy of participation forecasts should be tempered. First, the forecasting period for participation--two years ahead--is long relative to the period over which macroeconomic quantities are typically forecast. The lengthy forecasting period is due to the Federal budgeting cycle. At the start of each fiscal year, FNS submits forecasts to the Office of Management and Budget for inclusion into the President's proposed budget for the following fiscal year. A consequence of the lengthy forecasting period is that forecast errors are larger, because random factors that can throw the forecast off target are more apt to surface, and because forecasts of explanatory variables are less accurate for more distant future periods.

Second, forecasting models work best when the relationships that determine the forecasted variables are stable over time and amenable to estimation. However, three factors suggest that the context for forecasting participation is not likely to be stable. First, the FSP is a relatively new program, and the period for which participation data are available is short relative to such macroeconomic quantities as the Gross National Product, for which decades of data may be available. Second, FSP participation is affected directly by changes in program statutes, regulations, and procedures, which have occurred frequently. Third, the recent dramatic increase in participation, which began in a period marked by relatively low and stable unemployment, suggests that previous participation patterns may be changing. Each of these factors argues for viewing the participation forecasting model as a component of a forecasting process in which forecasting performance is monitored closely and the forecasting model updated regularly.

The explanatory factors that were tested in the multivariate statistical models included combinations of macroeconomic variables, program parameters, major legislative changes, and demographic variables. The choice of explanatory variables for testing was constrained in two ways. First, for an explanatory variable to be suitable for inclusion in the forecasting model, independent forecasts of the explanatory variable had to be available. Second, FNS relies on forecasts of macroeconomic quantities--such as the unemployment rate, price levels, personal income, and the Gross National Product--that are consistent with the President's budget proposal. In practice, the necessity of using independent forecasts of macroeconomic and demographic variables means that FNS forecasts of participation and benefits can be no more accurate than the independent forecasts.

THE RESULTS FROM A COMPARISON OF ALTERNATIVE FORECASTING MODELS

Alternative forecasting models for FNS applications were judged according to their ability to forecast participation and benefits in time periods that were outside the sample period used for

Table of Contents

estimation. For example, alternative models were estimated with data through 1987, and the estimated models were then used to generate forecasts of participation and benefits for 1988 and 1989. Because the true values of participation and benefits for 1988 and 1989 were known, it was possible to assess the forecasting accuracy of alternative models by comparing forecasted participation and benefits with actual levels of participation and benefits in those time periods. Models were also judged according to whether their estimated parameters were reasonable, and whether the models tracked participation and benefits accurately during the sample period used for estimation.

Several alternative models were estimated and their forecasting accuracy assessed. The key results were as follows:

- The participation models that were estimated generally yielded two-year-ahead participation forecasts that were accurate to within plus or minus 6 to 7 percent per month, or, equivalently, to within plus or minus roughly 1 million participants per month. Even if future average benefits were known with certainty, this level of forecasting accuracy implies that models may have forecast errors on the order of plus or minus $840 million annually. Larger errors may occur if the forecasts of macroeconomic quantities on which the participation forecasts are based are inaccurate.

- The forecasting performance of some participation models was marginally better than others. Forecasts generated by a participation model whose explanatory variables included the number of unemployed workers, variables for seasonality, and a correction for the correlation of random factors over time were the most accurate.

- The forecasting accuracy of the limited number of regression models of average benefits that were estimated was generally inferior to the accuracy of a formula approach for forecasting average benefits that relied on parameters estimated with a simulation methodology.

- A regression model of total program benefits provided forecasts whose accuracy was similar to the accuracy of forecasts from two-equation participation and average benefit models.

CONCLUSIONS

The purpose of the independent assessment was to determine whether FNS forecasting models were adequate, and to suggest improvements if they were not. The results indicate that the forecasting accuracy of FNS forecasting models might be modestly improved. However, in general, neither the alternative models nor the existing FNS models yielded forecasts that could be characterized as highly accurate. Moreover, none of the alternative models would have captured the increase in participation that began in 1989.

The reasons for the increase in participation that began in 1989 are not yet completely understood, and in retrospect the period may be viewed as a point at which new relationships emerged that should be reflected in the forecasting model. For this reason, we recommend that FNS continue to monitor the accuracy of future forecasting efforts and modify the forecasting model

Table of Contents

appropriately. To assist in these efforts, a forecasting process was specified for evaluating the performance of forecasting models over time. The forecasting process that was developed will provide a mechanism for tracking the quality of forecasts and for updating the empirical model to reflect new information.

Table of Contents

I. INTRODUCTION

The Food Stamp Program (FSP) enables low-income households to achieve and maintain nutritious diets by increasing their food purchasing power. Over time, the number of program participants has changed, often dramatically, in response to legislative changes in eligibility requirements, fluctuations in economic activity and changes in the structure of the economy, and changes in the behavior of households. Since 1977 the number of persons who receive food stamp benefits in a given month has fluctuated between 14 and 22 million (Figure I.1).[1] The number of participants was 14 million in 1978 prior to the elimination of the purchase requirement, which was a feature of the FSP up to that time. After the purchase requirement was eliminated by the Food Stamp Act of 1977 (PL 95-113)--a change that was implemented fully by January 1979--FSP participation increased sharply, to 16 million in the second quarter of fiscal year 1979. Participation peaked at 22 million in 1983 during a deep economic recession, and fell almost continuously after 1983. However, participation has recently been increasing at a rapid rate, crossing the 20-million-person threshold in March 1990 and rising to over 22 million in February 1991.

Since the Food Stamp Act of 1977, numerous legislative changes to the Food Stamp Program have affected participation (U.S. House of Representatives, 1991). Figure I.1 indicates five points at which significant pieces of legislation were passed.[2] The Omnibus Budget Reconciliation Act of 1981 (OBRA 1981) reduced eligibility and delayed benefit increases in a number of ways. The Food Security Act of 1985 relaxed benefit and eligibility rules, established categorical eligibility for households comprised entirely of Aid to Families with Dependent Children (AFDC) recipients or Supplemental Security Income (SSI) recipients, and established an employment and training program

---

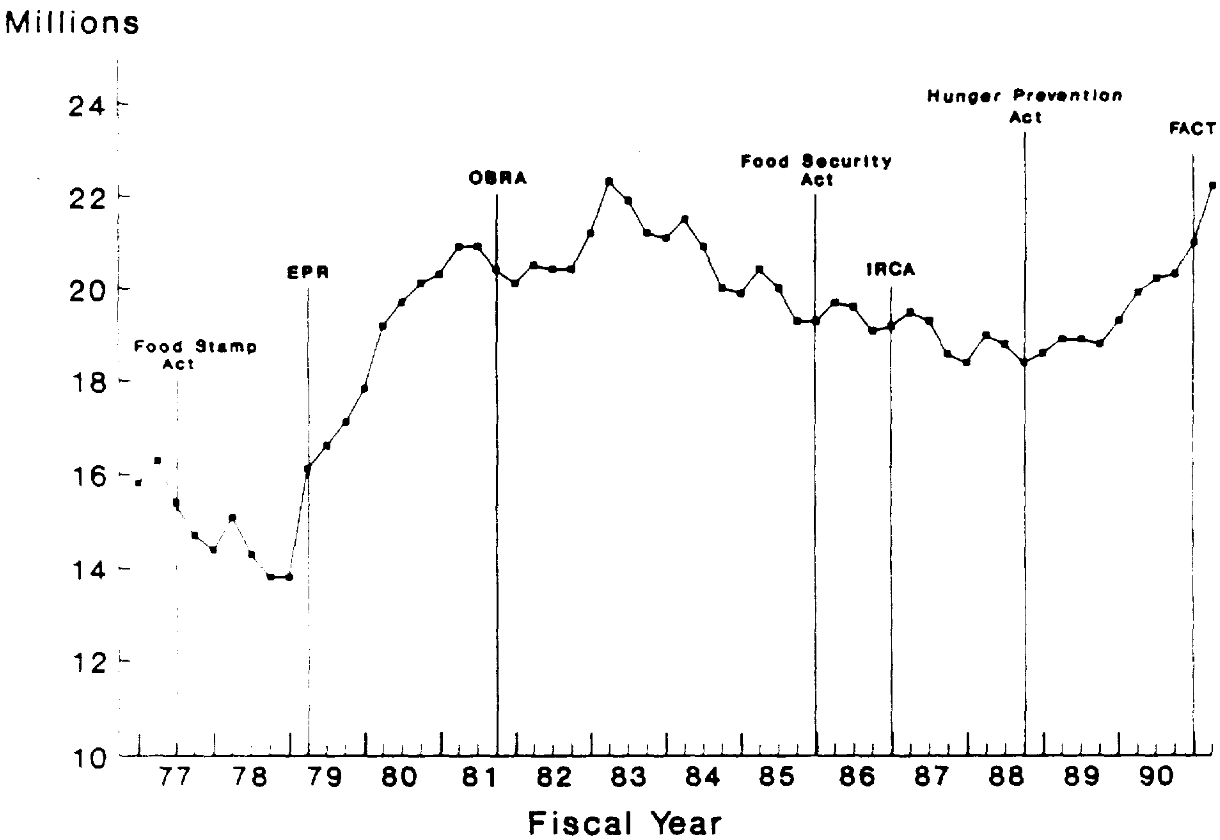
[1]Participation data exclude FSP participants in Puerto Rico.

[2]The figure does not show the points at which specific features of the legislation were actually implemented.

Table of Contents

FIGURE I.1

Food Stamp Participation



SOURCE: Public Information Data Bank, Food Stamp Program, USDA Food and Nutrition Service.

NOTE: Excludes food stamp participants in Puerto Rico. Food stamp participation refers to average monthly participation within fiscal year quarters.

Table of Contents

for able-bodied food stamp recipients. The Immigration Reform and Control Act of 1986 (IRCA) authorized that special categories of aliens could be eligible for food stamp benefits. The Hunger Prevention Act of 1988 increased food stamp benefits, liberalized eligibility and benefit rules, and authorized new funding for program outreach. The 1990 Food, Agriculture, Conservation, and Trade Act (FACT) reauthorized FSP appropriations through fiscal year 1995.

The average value of food stamp benefits per participant has also changed since 1980. However, in contrast to participation, average benefits have increased steadily since 1980 (see Figure I.2).[3] Average benefits adjusted for inflation have also increased since 1980, though by a much smaller amount. Beginning at $33 per month in the first quarter of fiscal year 1980, average monthly benefits increased to $52 by the first quarter of fiscal year 1989. However, after adjustment for inflation, average monthly benefits increased only modestly, from $33 to $37. Figure I.2 shows that increases in average benefits are generally evident in the first quarter of each fiscal year, when adjustments to maximum benefit allotments and other program parameters take effect. After the first quarter, average benefits within a fiscal year generally decline, due to the income growth and seasonal changes in the composition of FSP participants.

The sensitivity of FSP participation and benefits to economic conditions and to program changes poses a challenge for budgetary planning. The Food and Nutrition Service (FNS) submits budget estimates of future program benefits to the Office of Management and Budget (OMB) at the end of the first quarter of each fiscal year. However, the variability in FSP participation levels means that simple estimates of program benefits based on current participation levels may differ significantly from future program benefits.

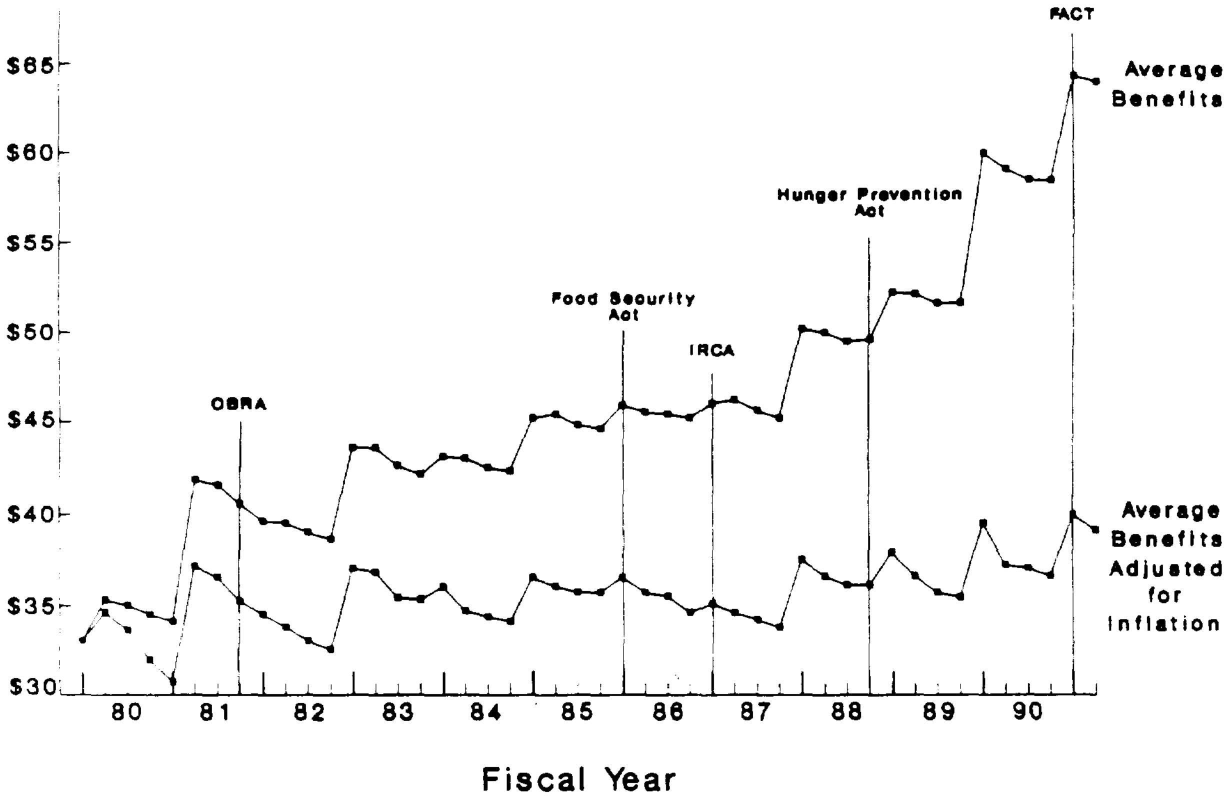
To make its budget estimates more accurate, FNS forecasts the total number of persons participating in the Food Stamp Program and the average value of food stamp benefits received.

[3]Data on average benefits exclude benefits received by FSP participants in Puerto Rico.

Table of Contents

FIGURE I.2

Average Monthly Food Stamp Benefits Per Participant



SOURCE: Public Information Data Bank, Food Stamp Program, USDA Food and Nutrition Service.

NOTE: Excludes food stamp benefits in Puerto Rico. Food stamp benefits refer to the average monthly benefits received by FSP participants within fiscal year quarters. Average benefits are adjusted for inflation using the CPI for food at home.

Forecasting is currently a two-step process. In the first step, FNS uses regression models to predict aggregate program participation and the average value of benefits received per participant. In the second step, FNS calculates total program benefits as the product of forecasted aggregate program participation and the forecasted average benefit per participant.[4] The choice of explanatory variables used by FNS is constrained in two ways. First, for a variable to be included in the forecasting model, independent forecasts of the variable must be available. Second, FNS relies on forecasts of macroeconomic quantities--such as the unemployment rate, price levels, personal income, and the Gross National Product--that are consistent with the President's budget proposal.

The recent forecasting experience of FNS econometric models suggested that an examination of alternative models would be useful. The FNS model in use in 1987 did not capture the dramatic increase in participation that began in 1989 and which continued through 1990, and concerns were raised within FNS that the forecasting model may not reflect the state of the art in econometric forecasting. To address these concerns, FNS contracted for an independent assessment of FNS forecasting models and asked for recommendations about alternative approaches for forecasting participation and benefits.[5]

A. OVERVIEW OF THE REPORT

This report provides an econometric analysis and evaluation of the current FNS forecasting model, and discusses alternative forecasting models for FSP participation and benefits. An additional goal of the report is to set out a forecasting *process* to serve as an organizing framework for future forecasting efforts.

---

[4]This description of the forecasting process applies as of April 1990. Prior to this date, FNS forecast the average value of benefits received per participant using a formula based on simulations of benefits received by different types of households.

[5]In conjunction with its efforts to improve its forecasting capabilities, FNS has also conducted studies to examine the reasons for the recent dramatic increase in participation (Corson and McConnell, 1990; McConnell, 1991).

A preliminary and necessary step before modeling FSP participation empirically is to discuss the conceptual factors underlying the FSP participation decision. Chapter II delineates a conceptual framework for FSP participation that illuminates the factors underlying the participation decision. The discussion in Chapter II emphasizes the importance of the interaction among program changes, evolving macroeconomic conditions, and demographic trends in explaining changes in FSP participation over time.

Chapter III presents an evaluation of several alternative forecasting models of FSP participation, average benefits, and total program benefits. The models were evaluated primarily according to their ability to provide accurate forecasts outside the sample period used for estimation. A regression model of participation whose explanatory variables include the number of unemployed workers and the number of female-headed households with children showed good forecasting accuracy among the alternative participation models. However, even the most accurate of the alternative models yielded two-year-ahead forecasts of program participation that were reliable only to within 6 percent of actual participation, and the degree of reliability would be lower if the explanatory variables used to generate forecasts of participation were themselves forecast with error. A regression model of total program benefits provided forecasts whose quality was similar to the quality of forecasts from two-equation participation and average benefit models.

Chapter IV discusses the elements of a forecasting process that will be useful for evaluating the performance of forecasting models over time. A uniform system of forecast reporting may facilitate accumulating forecasting experience and assessing how the forecasting model can be improved in the future. The chapter also presents a prototype recordkeeping form that may be a useful tool for organizing information used in the forecasting process.

II. CONCEPTUAL FRAMEWORK FOR FORECASTING FOOD STAMP PROGRAM PARTICIPATION

This chapter discusses a conceptual framework for forecasting Food Stamp Program participation. Participation is viewed as a dynamic event, with changes in participation driven by underlying changes in economic conditions, program parameters, and household circumstances.

From a static analytic perspective, an eligible household chooses whether to receive food stamps by comparing the value of food stamp benefits (that is, the enhanced consumption of food and other goods) with the costs of food stamp receipt (that is, the time and monetary cost of applying for and receiving food stamps, as well as any stigma or embarrassment associated with receiving and using food stamps). Households decide to participate if the benefits exceed the costs.

For example, eligible households with the lowest incomes may place the highest value on the enhanced consumption provided by food stamps, and would receive a greater amount of food stamp benefits as well, because food stamp benefits are larger for households whose income is lower. Both factors imply that households with lower incomes would be more likely to participate. The organization and activities of the FSP may also affect whether households are knowledgeable about the program and whether the program is accessible to them. Households that are unaware of the program or have limited access to program offices would be less likely to participate than would similar households that are aware of the program or have more convenient access to program offices.

A more dynamic perspective on program participation is possible if the static model is viewed as applying at a point in time, but allows household circumstances and program features to change over time. For example, a household that currently does not receive food stamps may later experience a reduction in income that makes program participation more desirable. Alternatively, program features may be altered in a way that makes participation more desirable. Conversely, a household that currently receives food stamps may later experience an increase in income that makes

Table of Contents

participation less desirable, or program features may be altered in a way that makes participation less desirable.

The dynamics of program participation can be expressed in terms of two transition rates: the *continuation rate*, which is the probability that a household on the program in one month continues on the program in the following month, and the *intake rate*, which is the probability that a household

not on the program in one month is on the program in the following month. At any point in time, the continuation rates of program participants depend on the level and variability of income and other household circumstances, and on the features of the FSP. The intake rates of nonparticipants also depend on the level and variability of income and household circumstances, and on program features.

In the simplest analytic case, with a fixed population and with intake and continuation rates that are constant across the population, the number of participants is determined solely as a function of the total size of the population and the two transition rates. FSP participation in a time period t consists of the proportion of program participants from the previous period who decided to remain

Simple differentiation of equation (2) shows that participation will rise when either c or k rises. Under simple assumptions about the effects of factors on transition rates, one can also determine whether changes in the factors will increase or reduce program participation.

Here, it is useful to sketch the channels through which various factors affect the transition rates and hence program participation. For example, macroeconomic factors affect transition rates by affecting employment and earnings levels. During economic downturns, some nonparticipants may experience unemployment or other forms of earnings reductions, such as cutbacks in the number of hours worked. In turn, these reductions increase FSP intake rates by increasing the number of households that meet program eligibility criteria. For program participants, the economic downturn may lower the possibilities of finding jobs or of finding better jobs, thus increasing continuation rates because fewer participants are able to move off the program. Hence, economic downturns increase the flow onto the program while reducing the flow off the program, thereby increasing the number of participants. Economic upturns would affect transition rates in the opposite direction, reducing the number of participants.

Increases in the value of maximum benefit allotments or allowable deductions from gross income will increase the food stamp benefits for which households are eligible, thus increasing the continuation rate (participants will be less likely to leave the program in the future) and the intake rate (eligible nonparticipants will be more likely to participate in the future). The number of participants will rise accordingly. The size of the increase will be determined by the sensitivity of the two transition rates to changes in food stamp benefits. Other changes in the program, such as increased outreach activities, may affect the intake rate by promulgating the program more widely, but will not affect the continuation rate.

Changes in the features of other social assistance programs, such as AFDC or SSI, may also affect FSP participation. Households consisting solely of AFDC and SSI recipients are categorically eligible to receive food stamp benefits. Hence, increases in AFDC benefit levels or greater program outreach activities may attract more eligible households to the AFDC program, which may in turn increase FSP participation.[1] However, because the benefits of other social assistance programs are frequently counted as income toward food stamp benefit determinations, increases in benefits from other programs may also mean that some households will receive smaller food stamp benefits, which may induce some of the households to leave the Food Stamp Program.

Demographic shifts may also affect average transition rates.[2] For example, the growing number of female-headed families may increase program participation because their earnings are generally lower than those of two-parent households, and their earnings may also fluctuate more widely due to changes in economic conditions.

If the relationship between continuation and intake rates and other factors were known or could be estimated, the number of program participants given by equation (2) could be forecast for various values of the underlying factors.[3] However, the form of the relationships among transition rates, program features, and macroeconomic and demographic factors is likely to be highly complex and not amenable to direct modeling. Moreover, the data demands for estimating models of transition rates

---

[1]An increase in AFDC participation may increase FSP participation because (1) once an individual is applying for AFDC, the cost of applying for food stamps is very low (a single application form and a single interview apply to both programs), (2) AFDC eligibility workers inform recipients about their eligibility for food stamps, and (3) once the individual has entered the welfare system, it may reduce the psychological burden of receiving additional welfare benefits.

[2]Demographic shifts would not affect program participation if transition rates were equal across households. However, because transition rates onto the Food Stamp Program are greater for some types of households than for others, increases in the number of households whose transition rates are larger would increase FSP participation.

[3]For example, in 1987, there were 90 million households in the United States. Monthly continuation and intake rates for the Food Stamp Program have been estimated to be about 94 percent and .5 percent, respectively (Burstein and Visher, 1989; and Carr et al., 1984). Substituting into equation (2), the equilibrium number of households receiving food stamps in a month is calculated to be 6.9 million. Because the average number of households in the Food Stamp Program in 1987 was actually 7.1 million, the simple model performs adequately in approximating the number of participating households. However, the equilibrium value is sensitive to the values of the transition rates, which were estimated by the preceding authors with data from 1979 to 1983. Transition rates that are estimated with more current data may differ from those estimated with data from earlier years.

would be imposing.[4] During periods when transition rates may be changing, the challenges of estimating these relationships to forecast participation are even more formidable.

An alternative forecasting approach used in this report is to construct multivariate models of aggregate program participation and benefits, and to use these models to forecast future participation and benefits. The aggregate multivariate models can be considered representative of the data relationships generated by changes in economic conditions, program parameters, and household characteristics which work their way toward changes in program participation. For example, it has been noted that rising unemployment should increase transition rates, thereby increasing program participation. Thus, the data would show a positive relationship between the number of unemployed workers and the number of program participants, and the strength of the data relationship could be estimated with multivariate methods. In general, information on the magnitudes of the causal effects of unemployment on the continuation rate and the intake rate cannot generally be disentangled in the estimated data relationship. However, a knowledge of the magnitudes of the direct causal relationships through which unemployment affects participation is unnecessary for budgeting and forecasting purposes.[5]

The choice of other types of variables to be included in the aggregate models can be motivated from within the conceptual framework. Variables that represent benefit parameters, major legislative changes, and the demographic composition of households will affect transition rates as discussed, and it is thus reasonable to test their inclusion in the aggregate models. However, some factors may affect transition rates only modestly, which makes it difficult for a statistical analysis to uncover a significant effect of those factors on participation. Other factors might have a strong relationship with

[4]Previous research on the dynamics of FSP participation has been based on multivariate statistical methods, but these efforts focused on understanding the reasons for program entry and exit rather than on building forecasting models (Burstein and Visher, 1989).

[5]Of course, estimating the causal relationships would be the primary goal of the analysis if our purpose were to gain a better understanding of the program participation decisions of households. In econometric terms, a model of causal relationships is considered to be a "structural" model, whereas a model of data relationships is considered to be a "reduced-form" model.

Table of Contents

transition rates, but may not have exhibited sufficient variation during the sample period to affect participation. Alternatively, several factors may have moved together, thereby making it difficult to uncover their separate influences on participation. For these reasons, it is possible that the forecasting ability of some aggregate models that include only a few of the variables suggested by the conceptual framework is similar to the forecasting ability of models that include a fuller range of variables suggested by the conceptual framework. Much of the analysis in the next chapter focuses on uncovering aggregate models that provide forecasts of reasonable quality with only a modest number of explanatory variables.

III. REGRESSION MODELS OF FOOD STAMP PROGRAM PARTICIPATION AND BENEFITS

This chapter examines a number of alternative regression-based forecasting models of FSP participation and benefits. The alternative models are evaluated according to the reasonableness of their estimated coefficients and their goodness of fit, and more importantly according to their ability to forecast outside the sample period. The chapter also examines an alternative approach for forecasting total program benefits directly with a regression model, an approach that provides forecasts that are comparable to forecasts of program benefits from the two-equation model currently in use.

Section A discusses the basic regression models and specification issues, and the model evaluation criteria used to assess alternative forecasting models. Section B examines a number of alternative specifications of FSP participation forecasting models. Section C examines several alternative specifications of models of average food stamp benefits. Section D presents a total food stamp benefit forecasting model. Section E compares the forecasts generated by the alternative approaches. Section F summarizes the key findings.

A. USING REGRESSION MODELS FOR FORECASTING

Because the analysis relies heavily on regression models, it is useful to begin by discussing the basic issues associated with forecasting with regression models.[1] A typical linear regression model of a variable Y can be expressed as:

$$(1) \quad Y_t = a + b_1X_{1t} + b_2X_{2t} + \cdots + b_nX_{nt} + u_t \qquad t = 1, 2, \cdots, T$$

[1]An alternative forecasting methodology was developed on the basis of a *time-series analysis* of FSP participation, in which current participation levels were determined by previous participation levels and by random movements of participation. Appendix C presents an analysis of pure time-series models of participation. In general, participation forecasts from the time-series models were inferior to those from the regression models.

Table of Contents

where t is the time period of observation, X_{it} are explanatory variables, a and b_i are the parameters to be estimated, and u_t is an error term that represents the influence of factors that affect the level of Y but are not themselves included in the model as explanatory variables.

If the error term u_t has a mean of zero and a constant variance, if errors are uncorrelated with each other over time, and if the parameters of equation (1) are estimated with ordinary least-squares techniques, an optimal forecast of Y at any period k in the future is given by:

$$(2) \quad Y^f_{T+k} = \hat{a} + \hat{b}_1 X^f_{1(T+k)} + \hat{b}_2 X^f_{2(T+k)} + \cdots + \hat{b}_n X^f_{n(T+k)} ,$$

where $\hat{a}, \hat{b}_1, \cdots, \hat{b}_n$ are the values of the parameters from equation (1) estimated with data through period T, and $X^f_{1(T+k)} \cdots X^f_{n(T+k)}$ are values of the forecasts of the explanatory variables k periods in the future (Pindyck and Rubinfeld, 1981). The forecasts of Y generated by equation (2) are optimal in the sense that the expected value of Y^f will equal the true value of Y (the forecasts are *unbiased*), and the average distance between the forecasted value and the actual value of Y will be smaller than the average distance generated by other forecasting methods (the forecasts have *minimum variance*). However, the optimal forecast property of equation (2) depends on the accuracy of the forecasts of the explanatory variables. That is, the forecasts of Y from equation (2) will generally be closer on average to the true values than will forecasts generated with other methods, but if the true values of the explanatory variables are far from the forecasted value the true value of Y may nevertheless be far from the forecasted value of Y.

The regression forecast of Y is more complicated when the error terms are *serially correlated.* Errors are said to be serially correlated when the values of the error in the current period are correlated with values in future periods. For example, for many time series of economic variables, a positive error in the current period is correlated with a positive error in the subsequent period. In most contexts, failing to allow for serial correlation when it exists leads to overstated estimates of the statistical significance of a model's estimated coefficients (Johnston, 1984).

Table of Contents

A simple representation of serial correlation is a *first-order autoregressive* error structure, which can be written as:

(3) $u_t = \rho u_{t-1} + \nu_t$

where ν_t is a random error term that has a mean of zero and a constant variance, and is uncorrelated over time. If ρ is positive, then positive error terms in current periods will tend to be followed by positive error terms in subsequent periods, and negative errors will tend to be followed by negative errors.

If the structure of the error term in equation (3) is incorporated into the regression model, forecasted values of Y will depend not only on the forecasted values of the explanatory variables, but also on the forecasted values of the error term (Pindyck and Rubinfeld, 1981). Because the mean value of ν_t in equation (3) is zero, a sensible forecast of the error one period ahead is $u^f_{T+1} = \hat{\rho}\hat{u}_T$, and in general a sensible forecast of the error k periods ahead is $u^f_{T+k} = \hat{\rho}^k\hat{u}_T$, where $\hat{\rho}$ is the estimated value of ρ. This information on future error terms can be incorporated into the basic forecasting equation (2) to generate more accurate forecasts, yielding:

(4) $Y^f_{T+k} = \hat{a} + \hat{b}_1 X^f_{1(T+k)} + \cdots + \hat{b}_n X^f_{n(T+k)} + \hat{\rho}^k\hat{u}_T .$

Because the estimated value of ρ is generally less than one, the effect of serial correlation on future forecasts of Y diminishes exponentially as forecasts are generated further into the future. The forecast of Y from equation (4) is optimal only when serial correlation of the error term has the particular structure represented by equation (3). More complex patterns of serial correlation would require modifying equation (4). The results of tests to determine whether equation (3) is the appropriate structure of serial correlation for regression models of FSP participation are discussed below.

Table of Contents

1. Criteria for Evaluating Models

Before presenting the estimation results for alternative forecasting models, it is appropriate to discuss the criteria that were used to evaluate the models.

Though models need not have reasonable estimated coefficients to forecast well, as a practical matter it would not be sensible to rely on forecasting models whose estimated coefficients were clearly at odds with intuition. For this reason, the estimated coefficients from alternative models were examined to determine whether the signs and magnitudes of the estimates were reasonable. The goodness-of-fit statistics of the estimated models were also examined, to verify whether the models tracked the data well. The alternative models generally fit the data exceptionally well, with R^2 statistics usually greater than .97.[2]

The primary purpose of considering the models examined here is to forecast future FSP participation and benefits. Thus, an important criterion for evaluating alternative models is their ability to forecast accurately outside the sample period. A variety of summary measures exist for out-of-sample forecast error. The *root mean square error* (*rmse*) is the most commonly used summary measure of forecasting quality (Pindyck and Rubinfeld, 1981). The *rmse* measures the average distance between forecasted values and actual values in a specific time period, and is expressed as:

$$(5) \qquad rmse = \sqrt{\frac{1}{q}\sum_{t=1}^{q}(Y_t - Y_t^f)^2}$$

where q is the number of time periods for which forecasts are made.[3] Larger values of the *rmse* imply less accurate forecasts of Y. Positive and negative forecast errors are weighted equally in

---

[2] R^2 measures the proportion of the variation in the dependent variable which is explained by the regression equation.

[3] The *rmse* should be distinguished from the *forecast variance*. In theory, a different forecast variance is associated with each future time period, with more distant periods having larger variances (see Section III.4).

Table of Contents

calculating the *rmse*, but larger forecast errors are weighted more heavily than smaller forecast errors because errors are squared before they are summed.[4]

The mean forecast error (*mfe*) is also a useful measure of out-of-sample forecast error. The *mfe* measures the extent to which forecasted values overpredict or underpredict actual values on average over a specific time period, and can be expressed as:

$$(6) \quad mfe = \frac{1}{q}\sum_{t=1}^{q}(Y_t - Y_t^f) .$$

The *mfe* is a useful criterion here because models that have a small *mfe* can generate accurate fiscal-year forecasts over the course of the year even if they do not forecast accurately for specific periods of the year.[5]

---

[4]As an example of this property of the *rmse*, consider two alternative models; the first forecasts values for two future periods of 90 and 90, and the second forecasts values of 80 and 100. Suppose that the actual values are 100 and 100. The *rmse* is 10.0 for the first model, which erred by 10 and 10, and 14.1 for the second model, which erred by 20 and 0.

[5]Several other summary measures of out-of-sample forecast errors can be defined. For example, the *rmse* and the *mfe* can be expressed in percentage terms as--

$$\text{rms percent error} = \sqrt{\frac{1}{q}\sum_{t=1}^{q}\left(\frac{Y_t - Y_t^f}{Y_t}\right)^2}$$

and

$$\text{mean percent error} = \frac{1}{q}\sum_{t=1}^{q}\left(\frac{Y_t - Y_t^f}{Y_t}\right),$$

and the *mean absolute error* can be defined as

$$\text{mean absolute error} = \frac{1}{q}\sum_{t=1}^{q} | \, Y_t - Y_t^f \, | .$$

Because the models estimated in Section III.B generally underpredicted or overpredicted FSP participation consistently during a time period, the mean absolute errors were numerically close to the mean forecast errors. For this reason, mean absolute errors are not presented in this report.

2. Computing Out-of-Sample Forecast Errors

The basis for evaluating the ability of alternative models to forecast future FSP participation and benefits was to compute out-of-sample forecasts and compare them with observed values of participation and benefits. The general approach for computing out-of-sample forecast errors was to partition the available data into a time period for estimating the model and a time period for generating forecasts from the estimated models and comparing the forecasts with actual values of the variable being forecast. If ample data were available, the out-of-sample forecasting procedure could have been repeated for many time periods, and the average quality of forecasts from different models could have been compared. However, as explained below, the time series for FSP participation is relatively short. For this reason, summary measures of out-of-sample forecast errors were computed for two time periods only, and the results for the two periods were averaged.

The approach for computing out-of-sample forecast errors was designed to simulate the forecasting procedure as it would be performed in practice. The FNS budgeting cycle requires that fiscal-year participation and benefit forecasts be made two years ahead of the current year. Because data were available through FY 1989, models were estimated over two sample periods--the first ending in FY 1986 and the second ending in FY 1987. Forecasts of participation (or average benefits) for 1988 and 1989 were then computed from the estimated regression models, based on actual values of the explanatory variables in 1988 and 1989.[6] Forecasted values of participation and benefits for FY 1988 and FY 1989 were then compared with actual values in FY 1988 and FY 1989. Out-of-sample forecast errors were calculated as the difference between forecasted and observed

[6]Using the actual values of the explanatory variables to generate forecasts is akin to ascertaining the forecasting accuracy of alternative models when the true state of the economy is *known*. It is also possible to use forecasted values of the explanatory variables to generate forecasts, which is akin to ascertaining the forecasting accuracy of alternative models when the true state of the economy is *unknown*. It was not possible to compare alternative models that contained forecasted explanatory variables, because historical forecasts of the many explanatory variables that were tested were not available. However, Appendix A shows that forecast errors associated with one key explanatory variable--the unemployment rate--may substantially increase the forecast errors associated with FSP participation.

Table of Contents

values. This procedure essentially simulated the results of two years of forecasting experience with each of the models.

Two other forecasting horizons are also important within the FNS budgeting cycle. FNS requires fiscal-year forecasts one year ahead of the current fiscal year, to review the adequacy of the budget for the upcoming fiscal year, and for the last two quarters within a fiscal year, to determine whether it must request a supplemental appropriation. Out-of-sample forecast errors for one year ahead and two quarters ahead were computed with a procedure similar to the procedure for two-year-ahead forecast errors. To generate one-year-ahead forecast errors for 1988 and 1989, models were estimated with data through 1987 and 1988, respectively, and forecasts were calculated and compared with observed values in those years. To generate two-quarter-ahead forecasts, models were estimated with data through FY 1988.2 and FY 1989.2, and forecasts were calculated for FY 1988.3 and 1988.4, and FY 1989.3 and 1989.4.

B. REGRESSION MODELS OF FSP PARTICIPATION

This section discusses the estimation results and forecasting accuracy of alternative forecasting models of FSP participation. The section first compares the forecasting models recently used by FNS. It then discusses several tests that were conducted to determine the robustness of the estimation results in terms of serial correlation, the method used to adjust for seasonal movements in FSP participation, the sample periods analyzed, the choice of quarterly versus monthly data as the basis for analysis, and the inclusion of lagged dependent variables. A number of model variants whose explanatory variables differ are then compared, and forecast confidence intervals are calculated to assess the statistical accuracy of FSP participation forecasts.

1. FNS Participation Models

Recent FNS participation models provide a natural benchmark for analyzing alternative forecasting models. The specific characteristics of FNS participation models have changed over time.

The first model, which is termed Model P.1, was used through FY 1989. The dependent variable for Model P.1 was the seasonally adjusted monthly average of FSP participants in a fiscal-year quarter, beginning in 1977. The explanatory variables were the seasonally adjusted unemployment rate, the unemployment rate lagged one quarter, the seasonally adjusted consumer price index for food used at home, and two variables to capture the effects of the elimination of the purchase requirement (denoted EPR) on participation and the effects of changes in program eligibility under OBRA 1981 (denoted REC) on participation. The value of the EPR variable was increased gradually over four quarters beginning in FY 1979.1 to reflect the phasing-in of the legislative change, and was set equal to one after FY 1980.1. The values assumed by the REC variable were increased from zero to one in 1982, held steady at one through FY 1983, fluctuated between zero and one through FY 1988.2, and stabilized at zero thereafter. The fluctuations of the REC variable were intended to reflect program changes mandated by OBRA 1981 and subsequent offsetting changes.

The participation model used by FNS in FY 1990 differs in several ways from Model P.1.[7] The dependent variable for the more recent model, referred to as Model P.2, is the monthly average of FSP participation in a fiscal-year quarter, seasonally unadjusted, since 1982. The explanatory variables are the seasonally adjusted lagged unemployment rate, the lagged FSP participation level, three dummy variables for fiscal-year quarters (to account for the seasonality pattern in FSP participation), and a dummy variable equal to 0 prior to FY 1989.3 and equal to 1 thereafter to reflect changes in participation due to the 1988 Hunger Prevention Act. Because the sample period for Model P.2 begins in 1982, the dummy variables to capture the program changes for EPR and for OBRA 1981 were unnecessary because the entire sample period followed these program changes.

[7]FNS routinely assesses and updates its forecasting models. For this reason, models used in various years generally differ.

Table of Contents

Columns (1) and (3) of Table III.1 show regression estimates for Models P.1 and P.2.[8] The estimated coefficients and other in-sample statistics for both models were reasonable. For Model 1, the estimated coefficient for the EPR variable indicates that participation rose by 4.6 million after the EPR was implemented. The estimated coefficient for the REC variable indicates that participation fell by about 2 million participants after the program changes mandated by OBRA 1981 were implemented. The estimated coefficients for the unemployment rate indicate that higher unemployment rates are correlated with higher levels of program participation.[9] A one percentage point increase in the unemployment rate is correlated with an increase in program participation of more than one million persons. Model P.1 fits the data well within the sample period, as indicated by the high R^2 value. The estimated Durbin-Watson statistic also indicates that the error term is not serially correlated.[10]

As noted earlier, the REC variable that represents changes in the program mandated by OBRA 1981 assumed several different values between zero and one. The expressed intent of using different values was to capture later legislative changes that offset provisions of OBRA 1981, but an examination of the pattern of values for REC also supports the view that they were chosen to improve the fit between the model and the data. Because the values for REC were determined judgmentally, it is reasonable to ask whether the estimated results were sensitive to the particular

---

[8] Models are labelled *P* if the dependent variable was FSP participation, *B* if the dependent variable was average benefits, and *T* if the dependent variable was total benefits.

[9] As noted in Chapter II, the estimated coefficients in aggregate models cannot be interpreted as representing causal relationships. For example, the increase in FSP participation following the implementation of the EPR may also have been due to other factors that occurred in the same time period.

[10] The Durbin-Watson statistic (d) is a measure of first-order serial correlation (Johnston, 1984). In the Durbin-Watson test, the estimated Durbin-Watson statistic is compared with two critical values from existing statistical tables. Positive first-order serial correlation is evident if the estimated value is less than the smaller of the two critical values. Positive first-order serial correlation is not evident

Table of Contents

TABLE III.1

RECENT FNS REGRESSION MODELS OF FSP PARTICIPATION

	(1)[a] Model P.1	(2)[a] Model P.1′	(3)[b] Model P.2
Unemployment rate, seasonally adjusted	355.8 (82.6)	25.6 (150.9)	--
Unemployment rate lagged one quarter, seasonally adjusted	730.3 (71.4)	754.1 (139.3)	327.9 (44.2)
FSP participation lagged one quarter (thousands)	--	--	0.55 (0.06)
CPI for food at home, seasonally adjusted	5.9 (1.5)	9.7 (4.7)	--
Elimination of purchase requirement (EPR)	4,605.4 (154.9)	4,543.3 (350.3)	--
OBRA 1981 (REC)[c]	-2,070.5 (156.4)	--	--
OBRA 1981 (REC2)[d]	--	-946.7 (271.9)	--
FY 1989.3 dummy	--	--	503.5 (129.3)
1st quarter dummy	--	--	451.1 (85.3)
2nd quarter dummy	--	--	922.2 (84.3)
3rd quarter dummy	--	--	375.6 (81.2)
Constant	6,305.8 (500.6)	7,534.5 (1,267.1)	6,057.3 (1,049.4)
$\bar{R}^2$	0.99	0.98	0.98
Standard error of the regression	175.7	342.8	159.3
Durbin-Watson statistic	2.1	0.6	1.4
Sample period	FY 1977.1-FY 1989.4	FY 1977.1-FY 1989.4	FY 1982.1-FY 1989.4

TABLE III.1 (continued)

NOTE: Standard errors are in parentheses. The data used for estimation are shown in Appendix D.

[a]The dependent variable is the seasonally adjusted monthly average of FSP participants (in thousands) during a fiscal-year quarter.

[b]The dependent variable is the monthly average of FSP participants (in thousands) during a fiscal-year quarter (not seasonally adjusted).

[c]REC increases from zero to one in FY 1982, remains at one through FY 1983, fluctuates between zero and one through FY 1988.2, and stabilizes at zero thereafter.

[d]REC2 is a dummy variable equal to zero prior to FY 1982 and equal to one thereafter.

Table of Contents

values chosen. The robustness of the results was tested by specifying a new OBRA 1981 dummy variable, REC2. The values of REC2 were assumed to be zero before the second quarter of 1982, and one thereafter.

Column 2 of Table III.1 shows the estimation results from a model in which REC2 is substituted for REC, denoted as Model P.1′. The estimated coefficient of REC2 is less than half the value for REC. The estimated coefficient for the current unemployment rate is much smaller and statistically insignificant. The standard error of the regression is larger, and the Durbin-Watson statistic is much smaller and indicates positive serial correlation of the error term.[11] These results suggest that several important characteristics of the model were sensitive to how the OBRA 1981 variable was defined.

Column 3 of Table III.1 shows the estimation results from Model P.2, the FY 1990 version of the FNS participation model. The estimated coefficients for Model P.2 are reasonable, and the goodness-of-fit statistics indicate that the model fits the data well. The estimated coefficient for lagged participation is highly statistically significant, as is the estimated coefficient for the lagged unemployment rate. The pattern of seasonality in the estimated coefficients of the quarterly dummy variables indicates that program participation tends to be greater by almost a million participants in the second fiscal quarter (January through March) than in the fourth fiscal quarter (July through September). The estimated coefficient for the Hunger Prevention Act dummy variable indicates that participation rose by a half million after the third quarter of FY 1989. The presence of a lagged dependent variable invalidates the standard Durbin-Watson test for serial correlation, but the Durbin-

[11]The computed value of the Durbin-Watson statistic for Model P.1′ was 0.60. The critical lower-bound value of the Durbin-Watson test statistic at the 95 percent confidence level for a model with five explanatory variables and a sample size of 50 is 1.34 (Johnston, 1984, Table B-5).

h test for serial correlation, which accounts for the presence of a lagged dependent variable, suggests positive serial correlation of the error term.[12]

Table III.2 shows the out-of-sample forecasting accuracy of the three models (Models P.1 and P.2, and Model P.1′), and reports *rmse* and *mfe* statistics for the three forecast horizons. It should be recalled that the procedure for calculating two-year-ahead out-of-sample forecasts errors was to truncate the sample in 1986 and 1987 and to estimate the models with the truncated sample periods, and then to calculate forecasts of participation and forecast errors in 1988 and 1989. Similar procedures were followed for one-year-ahead and two-quarter-ahead forecast errors. The coefficients for the models estimated with the truncated sample periods are not reported, but in general they differed from the estimated coefficients reported in Table III.1 due to the different sample periods. The Hunger Prevention Act variable was not used to calculate out-of-sample forecast errors for Model P.2 because the variable assumed a value of one beginning only in 1989.

The average two-year-ahead *rmse* was 236 for Model P.1, 378 for Model P.1′, and 467 for Model P.2. The units of the *rmse* are in thousands of participants per month, and from FY 1988 through

---

[12]The Durbin-h statistic is defined as $h = \left(1 - \frac{d}{2}\right)\sqrt{\frac{n}{1 - n\hat{V}}}$ where $\hat{V}$ is the estimated variance of the coefficient of Y_{t-1}, n is the sample size, and d is the Durbin-Watson statistic. Based on the calculated Durbin-Watson statistic of 1.4 from Table III.1, the value of the Durbin-h statistic is 1.8, and the critical value of the Durbin-h statistic at the 95 percent confidence level is 1.65 (Johnston, 1984). An alternative test for serial correlation, also suggested by Durbin, entailed regressing the residuals from Model 2 on the lagged value of the residuals and all other explanatory variables. In this case, the Durbin-h test is a test of the statistical significance of the estimated coefficient of the lagged residual. The t-statistic of the estimated coefficient of 1.6 does not indicate positive serial correlation at the 95 percent confidence level. However, small-sample Monte Carlo experiments have shown that the alternative test for serial correlation is significantly less likely to detect serial correlation when it exists (Park, 1975). There is no consensus about the best test for serial correlation in lagged dependent variable models. Thus, some ambiguity exists about whether serial correlation is present in the FNS model. However, because the econometric consequences of ignoring serial correlation are severe in lagged dependent variable models, a conservative strategy would be to allow for serial correlation even if there is only weak evidence that it exists.

TABLE III.2

MEASURES OF OUT-OF-SAMPLE FORECAST ERROR FOR RECENT FNS REGRESSION MODELS OF FSP PARTICIPATION
(Participants in Thousands)

	(1) Model P.1	(2) Model P.1′	(3)[a] Model P.2
Two Years Ahead			
mfe			
FY 1988	34.0	-25.4	9.3
FY 1989	247.0	542.8	639.8
Average	141.5	258.7	324.6
rmse			
FY 1988	153.7	173.3	232.4
FY 1989	317.3	582.5	700.6
Average	235.5	377.9	466.5
One Year Ahead			
mfe			
FY 1988	17.8	100.0	106.3
FY 1989	247.5	474.5	435.8
Average	132.7	287.3	271.0
rmse			
FY 1988	150.1	205.3	247.0
FY 1989	318.4	516.3	528.9
Average	234.3	360.8	388.0
Six Months Ahead			
mfe			
FY 1988.3-FY 1988.4	-124.0	263.9	348.0
FY 1989.3-FY 1989.4	399.7	561.3	607.5
Average	137.8	412.6	477.8
rmse			
FY 1988.3-FY 1988.4	171.1	293.1	371.1
FY 1989.3-FY 1989.4	422.5	576.6	647.4
Average	296.8	434.9	509.3

NOTE: Models were estimated through FY 1986 and FY 1987 to generate two-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; through FY 1987 and FY 1988 to generate one-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; and through FY 1988.2 and FY 1989.2 to generate six-month-ahead out-of-sample forecasts for FY 1988.3-FY 1988.4 and FY 1989.3-FY 1989.4, respectively. The estimation results for the models used to generate out-of-sample forecasts are included as supporting tables in a separate volume.

[a]The FY 1989.3 dummy variable representing the Hunger Prevention Act was excluded from the regression models used to generate out-of-sample forecasts.

FY 1989 FSP participation averaged roughly 18.6 million participants per month. Thus, the two-year-ahead *rmse* statistics are, respectively, 1.2 percent, 2.0 percent, and 2.5 percent of average monthly participation over the period. The *mfe* statistics indicate that each of the three models would have underforecast participation over the period, by 142,000 participants per month in Model P.1 and by 325,000 participants per month in Model P.2.[13] The underforecasts were more severe in 1989, which as noted in Chapter I was the year in which participation began to rise after five years of steady decline. The forecasting accuracy of the models did not improve as the forecast horizon became shorter. For Model P.1, the two-year-ahead *rmse* was 236, the one-year-ahead *rmse* was 234, and the two-quarter-ahead *rmse* was 297. A similar pattern was evident for the other two models, and for mean forecast errors.

2. Tests of the Model Specification

Five characteristics of the FNS participation models examined in the previous section were investigated further: (1) the effect of serial correlation, (2) the method used to adjust for seasonality, (3) the use of quarterly data, (4) the choice of the sample period, and (5) the inclusion of lagged participation as an explanatory variable. The purpose of investigating these issues further was to determine the sensitivity of the estimated results to the particular specifications chosen.

a. The Effect of Serial Correlation

As noted in Section III.A, the random error term in time-series settings frequently exhibits positive serial correlation. The econometric consequence of ignoring serial correlation is that ordinary least-squares techniques yield inefficient parameter estimates and, in most cases, an underestimate of the error variance, thus yielding overstated significance tests and excessively narrow confidence intervals (Johnston, 1984).

---

[13]Because forecast errors were calculated as actual participation minus forecasted participation, positive values of *mfe* correspond to underpredictions of participation.

Table of Contents

Durbin-Watson tests (or Durbin-h tests) for serial correlation for Models P.1′ and P.2 were positive, meaning that the hypothesis that the errors were not serially correlated could be rejected. Standard econometric practice was used to modify Model P.1′ whereby the error term was given a simple first-order autoregressive structure (see equation (3)), and the model was re-estimated.[14]

Column (2) of Table III.3 shows the estimated coefficients for the model in which serial correlation was corrected (denoted Model P.3). For comparison, column 1 of Table III.3 shows the estimated coefficients for the model without the serial correlation correction (Model P.1′). The estimated value of the serial correlation coefficient for Model P.3 was .74, and the estimate was statistically significant. Comparing the estimated coefficients of the two models shows that the serial correlation correction yielded a higher value of the estimated coefficient for current unemployment and a lower value of the estimated coefficient for lagged unemployment. The standard error of the regression for Model P.3 is much smaller than for Model P.1′, indicating that Model P.3 fit the data better than did Model P.1′. The Model P.3 Durbin-Watson statistic of 2.0 indicates that there is no evidence of remaining serial correlation.

Table III.4 indicates that the out-of-sample forecasting performance of Model P.1′ improved when serial correlation was accounted for. The two-year-ahead *rmse* was 378 for Model P.1′ and 297 for Model P.3, which included the serial correlation correction.

b. Seasonal Adjustment Methods

FSP participation exhibits a strong seasonal pattern, with a significant decline in participation generally occurring between the second and fourth fiscal quarters. Unemployment exhibits a similar

---

[14]The Cochrane-Orcutt procedure in MicroTSP was used to estimate the coefficients of the modified model that included the serial correlation parameter ρ. The Cochrane-Orcutt procedure first uses ordinary least squares to estimate the model without accounting for serial correlation. The residuals from the estimated model are then regressed on the lagged residuals to estimate ρ. This estimate is then used to transform the variables of the regression model to obtain coefficient estimates that can be shown to be more efficient than OLS coefficient estimates (the transformation subtracts the lagged value of each variable times ρ from the current-period value). The process is then iterated until the estimated value of ρ does not change.

TABLE III.3

ALTERNATIVE SPECIFICATIONS OF REGRESSION MODELS OF FSP PARTICIPATION

		Serial Correlation	Seasonal Adjustment	Monthly Model	Sample Period	
	(1)[a] Model P.1'	(2)[a] Model P.3	(3) Model P.4	(4)[b] Model P.5	(5) Model P.6	(6) Model P.7
Unemployment rate	25.6 (150.9)	136.1 (127.1)	451.8 (126.0)	248.9 (102.6)	257.4 (139.4)	54.4 (166.9)
Unemployment rate lagged once	754.1 (139.3)	554.3 (113.8)	251.8 (110.5)	81.0 (102.5)	267.4 (124.7)	521.5 (167.6)
CPI for food at home	9.7 (4.7)	9.4 (5.2)	13.0 (5.7)	8.2 (12.5)	-4.3 (7.7)	6.6 (13.1)
Elimination of purchase requirement (EPR)	4,543.2 (350.3)	4,736.0 (481.9)	4,659.8 (478.7)	3,056.7 (557.5)	--	--
OBRA 1981 (REC2)[c]	-946.7 (271.9)	-705.0 (268.1)	-806.6 (269.3)	-385.3 (278.2)	-823.7 (270.6)	--
1st quarter dummy	--	--	239.4 (75.7)	--	178.6 (81.5)	199.3 (93.1)
2nd quarter dummy	--	--	429.2 (118.7)	--	563.5 (134.2)	739.1 (173.0)
3rd quarter dummy	--	--	237.3 (93.0)	--	283.5 (93.1)	150.5 (124.4)
Constant	7,534.6 (1,267.1)	7,897.7 (1,502.0)	6,645.8 (1,767.7)	12,493.1 (4,180.1)	17,944.7 (2,794.7)	13,235.1 (4,800.8)
ρ	--	0.74 (0.11)	0.77 (0.10)	0.97 (0.02)	0.75 (0.07)	0.74 (0.13)
$\bar{R}^2$	0.98	0.99	0.99	0.98	0.96	0.95
Standard error of the regression	342.8	239.1	225.1	262.5	211.8	230.1
Durbin-Watson statistic	0.6	2.0	1.6	2.4	1.4	1.1
Sample period	FY 1977.1- FY 1989.4	FY 1977.1- FY 1989.4	FY 1977.1 FY 1989.4	FY 1976.10 FY 1989.09	FY 1980.1- FY 1989.4	FY 1982.1- FY 1989.4

Table of Contents

TABLE III.3 (continued)

NOTES: Standard errors are in parentheses. The dependent variable is the monthly average of FSP participants (in thousands) during a fiscal-year quarter. The data used for estimation are shown in Appendix D.

[a]FSP participation, the unemployment rate, and the CPI for food at home were seasonally adjusted prior to the model estimation.

[b]Monthly time dummies were used during the model estimation but are not reported in the table.

[c]REC2 is a dummy variable that equals zero prior to FY 1982 and equals one thereafter.

Table of Contents

TABLE III.4

MEASURES OF OUT-OF-SAMPLE FORECAST ERROR FOR ALTERNATIVE SPECIFICATIONS OF REGRESSION MODELS OF FSP PARTICIPATION
(Participants in Thousands)

		Serial Correlation	Seasonal Adjustment	Monthly Model	Sample Period	
	(1) Model P.1'	(2) Model P.3	(3) Model P.4	(4) Model P.5	(5) Model P.6	(6) Model P.7
Two Years Ahead						
mfe						
FY 1988	-25.4	-271.0	-324.0	-1,002.2	41.3	-110.8
FY 1989	542.8	242.8	92.9	-393.3	696.3	635.0
Average	258.7	-14.1	-115.6	-697.7	368.8	262.2
rmse						
FY 1988	173.3	298.6	349.4	1011.8	186.5	261.2
FY 1989	582.5	295.6	146.2	435.1	766.4	747.4
Average	377.9	297.1	247.7	723.5	476.5	504.3
One Year Ahead						
mfe						
FY 1988	100.0	-37.3	-92.5	230.4	68.5	29.8
FY 1989	474.5	151.5	105.5	-144.3	447.5	413.0
Average	287.3	57.1	6.5	43.1	258.0	221.4
rmse						
FY 1988	205.3	121.8	146.3	266.2	190.8	236.9
FY 1989	516.3	235.4	157.7	228.1	544.2	553.3
Average	360.8	178.6	152.0	247.2	367.5	395.1
Six Months Ahead						
mfe						
FY 1988.3-FY 1988.4	263.9	143.5	189.0	19.0	289.5	354.5
FY 1989.3-FY 1989.4	561.3	325.5	254.0	180.3	579.0	721.5
Average	412.6	234.5	221.5	99.7	434.3	538.0
rmse						
FY 1988.3-FY 1988.4	293.1	165.8	191.2	116.7	296.2	374.5
FY 1989.3-FY 1989.4	576.6	355.4	260.7	228.0	599.5	734.3
Average	434.9	260.6	226.0	172.4	447.9	554.4

NOTE: Models were estimated through FY 1986 and FY 1987 to generate two-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; through FY 1987 and FY 1988 to generate one-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; and through FY 1988.2 and FY 1989.2 to generate six-month-ahead out-of-sample forecasts for FY 1988.3-FY 1988.4 and FY 1989.3-FY 1989.4, respectively. The estimation results for the models used to generate out-of-sample forecasts are included as supporting tables in a separate volume.

seasonal pattern. One approach for modeling seasonality is to deseasonalize variables prior to estimating the model. The pattern of seasonality is then imposed on the forecasts from the estimated model. Another approach to account for seasonal patterns is to enter seasonal dummy variables in the regression model directly. Model P.1′ was estimated with data that were deseasonalized prior to estimating the model. Model P.2 was based on seasonally unadjusted data but included quarterly dummy variables to capture seasonal patterns.

The two approaches for adjusting for seasonality are conceptually related, and in this context the choice between them depends on which of the approaches yields better forecasts.[15] Column (3) of Table III.3 shows the estimated coefficients for Model P.4, which was based on seasonally unadjusted data and included quarterly dummy variables to capture seasonality. All other features of the model are identical to those of Model P.3. The pattern of the two estimated unemployment coefficients for the two models was reversed: Model P.4 had a larger coefficient for current versus lagged unemployment, and Model P.3 had a larger coefficient for lagged unemployment. The magnitudes of the other coefficients were similar.

Table III.4 indicates that the out-of-sample forecasting performance of Model P.4 was superior to that of Model P.3. The two-year-ahead *rmse* was 247 for Model P.4 and 297 for Model P.3. One-year-ahead and two-quarter-ahead forecasts were also more accurate with Model P.4. For this reason, seasonality was modeled with the dummy variable approach in later models.

[15]The conceptual link between the two methods can be clarified via the properties of ordinary least squares coefficient estimates. As can be shown in a two-variable regression model, if program participation and unemployment are each regressed on seasonal dummy variables, and the residuals from the participation equation are regressed on the residuals from the unemployment equation (the residuals are deseasonalized participation and unemployment), the resulting estimated coefficient for unemployment would equal the coefficient estimate for unemployment that would be obtained in a model which contained unemployment and seasonal dummy variables together. In this case, deseasonalizing prior to estimation yields the same results as entering seasonal dummy variables directly into the regression model. However, the results will generally differ if a different method is used to deseasonalize variables prior to estimation.

Table of Contents

c. Quarterly Data

The models estimated to this point have used the monthly average of aggregate FSP participation in a fiscal-year quarter as the dependent variable. These data are also available by month. For comparison purposes, Model P.4 was re-estimated with monthly data. Column 4 of Table III.3 shows that the monthly model (Model P.5) has smaller unemployment coefficients and a larger serial correlation coefficient than does the quarterly model.

The two-year-ahead out-of-sample forecasting performance of the monthly model was inferior to the forecasting performance of the quarterly model. The two-year-ahead *rmse* was 724 for the monthly model, compared with 248 for the quarterly model. However, the monthly model provided more accurate six-month-ahead forecasts, with an *rmse* of 172, compared with an *rmse* of 226 for the quarterly model. This increase in precision suggests that monthly models be used for short-term forecasting, and that quarterly models be used for longer-term forecasting. However, using two separate forecasting models may not be cost-effective, considering that the improvement in the short-term forecasting accuracy of the monthly models is modest.

d. Choice of Sample Period

Using a longer time series is generally preferred in regression model estimation because a longer time series typically exhibits greater variation among the explanatory variables, and will thus generally yield more precise coefficient estimates. However, this reasoning is invalid if the underlying model changes over time. For example, legislative changes to the FSP may change the underlying data relationships, and in this case it may be appropriate to break the time series and estimate a new model for the post-legislation period. An alternative approach is to incorporate variables in the model to adjust for the effects of legislative changes. Model P.1 incorporated variables to capture legislative changes, whereas Model P.2 incorporated data only from after 1981 and contained no variables for legislative changes.

Two major legislative changes to the program have occurred since 1977. The first was the Food Stamp Act of 1977, which, among many changes, required that states eliminate the purchase requirement for food stamps. The second was OBRA 1981, which reduced eligibility and delayed benefit increases. An examination was undertaken to assess whether the Food Stamp Act of 1977 changed the relationship between FSP participation and the explanatory variables, in which the coefficients from a model that was estimated over the entire sample period with dummy variables to capture the effects of the legislation (Model P.4) were compared with estimated coefficients from a model that excluded the first three years of the sample period and the EPR dummy variable (Model P.6).[16] The pattern of the estimated unemployment coefficients differed for the two models, and the sign of the coefficient for CPI-food was reversed. These coefficient differences were not statistically significant, but tests for structural changes have little power to detect differences when sample sizes are small, as they are here. For this reason, models were compared on the basis of their out-of-sample forecasting properties. Table III.4 indicates that the out-of-sample forecasts of the 1980 model are less accurate than for the 1977 model, with a two-year-ahead *rmse* for the 1980 model of 477, compared with 248 for the 1977 model. The evidence thus suggests that truncating the sample period at 1980 would reduce the forecasting accuracy of the model.[17]

A similar result was found when the sample period was truncated at 1982, to reflect the legislative changes of OBRA 1981. Column 6 of Table III.3 shows the results for this model (Model P.7). A comparison between Model P.7 and Model P.4 shows that truncating the sample period skews the unemployment coefficients, with greater weight placed on lagged unemployment. The coefficient for current unemployment in the post-OBRA model is small and statistically insignificant. However, the out-of-sample forecasting accuracy of the post-OBRA model was inferior to the

[16]Because the purchase requirement was eliminated by the last quarter of 1979, the sample period was truncated so that the first quarter of 1980 became the first observation.

[17]In some contexts, it is possible to test statistically for structural change (Johnston, 1984). However, in this case, the tests have very little power to detect structural change due to the relatively brief periods being compared.

Table of Contents

accuracy of both the post-EPR model and the basic model that used 1977 as the starting point. The two-year-ahead *rmse* was 504 for the post-OBRA model, compared with 477 for the post-EPR model and 248 for the basic model.

In the remainder of the analysis, dummy variables are included to capture the effects of legislative changes on participation. However, as more data become available, it may be appropriate to evaluate alternative models, to examine whether the earlier period of legislative changes can be excluded from future model estimation.

e. Lagged Program Participation as an Explanatory Variable

Lagged dependent variables are sometimes used as explanatory variables in aggregate models. In Model P.2, lagged program participation was entered as an explanatory variable, and the estimated coefficient for lagged participation was highly statistically significant.

However, the results of the Durbin-*h* tests reported in the previous section indicated that serial correlation may be present in a model that includes lagged participation as an explanatory variable. When the error term is serially correlated and the lagged dependent variable is used as an explanatory variable, ordinary least squares estimates of all regression coefficients are biased and inconsistent, as are forecasts from regression models estimated with ordinary least squares (Johnston, 1984). The econometric difficulty lies in the correlation between lagged participation and the error term. Errors in the previous period affect both participation in the previous period, which is an explanatory variable, and errors in the current period. Correlation between explanatory variables and the error term violates the conditions under which ordinary least squares yields unbiased and consistent coefficient estimates.

Hatanaka (1974) developed instrumental variables techniques that yield consistent estimates of the regression coefficients when serial correlation is present with a lagged dependent variable. However, when these techniques were applied to Model P.2, the estimated coefficient of lagged program participation was negative and did not differ statistically from zero. This result suggests that the significant effect of lagged participation in Model P.2 earlier may be due to the correlation between lagged participation and the error term, rather than to lagged participation directly. For this reason, lagged participation is excluded as an explanatory variable in the remainder of the analysis here. However, as more data become available in the future, it may be useful to explore other lagged participation models.

3. Other Explanatory Variables

The results of the specification checks indicated that a reasonable regression model for forecasting program participation exhibits several characteristics. The model is based on quarterly data beginning in 1977, and includes unemployment and lagged unemployment variables, seasonal dummy variables, several variables to incorporate the effects of legislative changes, and a correction for serial correlation. This section reports the results from a number of models of this type that included additional explanatory variables.

On the basis of the conceptual framework, other explanatory variables for the regression models were classified as representing general economic conditions, the demographic composition of households, FSP parameters, or legislative changes to the program. Table III.5 lists the variables that were included in each of the categories.

The demands of forecasting FSP participation on the basis of other forecasted variables clearly place a premium on parsimonious models that include explanatory variables for which forecasts are readily obtained. In recognition of these demands, the modeling strategy used to examine additional explanatory variables was to introduce one additional variable at a time to the basic model specified earlier. The estimated coefficients for the additional variable were examined to determine whether their sign and magnitude were reasonable. For models whose coefficients were reasonable, out-of-sample forecast errors were calculated and compared with the out-of-sample forecasts errors from

Table of Contents

TABLE III.5

EXPLANATORY VARIABLES FOR FSP PARTICIPATION AND BENEFIT REGRESSION MODELS

1. **General Economic Conditions**	Number of unemployed workers[a]
	Unemployment rate[a]
	CPI[b]
	CPI for food consumed at home[b]
	Per-capita disposable income[c]
	Number of workers exhausting UI benefits[d]
	Number of first UI payments[e]
	Number of workers employed in the personal services industry[a]
	Number of workers employed in retail trade[a]
	Number of workers employed in nonagricultural jobs[a]
	Hourly wages[a]
	Hourly wages in the personal services industry[a]
	Hourly wages in retail trade[a]
	Weekly earnings[a]
	Weekly earnings in the personal services industry[a]
	Weekly earnings in retail trade[a]
2. **Demographic Characteristics**	Number of female-headed households with children under 18[f]
	Number of female-headed households below poverty with children under 18[f]
	Number of AFDC recipients[d]
	Number of SSI recipients[d]
3. **Program Parameters**	Maximum allotment
	Standard deduction
	Excess shelter and child care deductions
4. **Legislative Changes**	Elimination of purchase requirement (FY 1979)
	Legislative changes in 1981-82 (OBRA 1981, Agricultural and Food Act of 1981, and OBRA 1982)
	Food Security Act (FY 1986)
	IRCA (FY 1987)
	Hunger Prevention Act (FY 1988)

[a]U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*.

[b]U.S. Department of Labor, Bureau of Labor Statistics, *Monthly Labor Review*.

[c]Council of Economic Advisors, *Economic Report of the President*.

[d]U.S. Department of Health and Human Services, Social Security Administration, *Social Security Bulletin*.

[e]Unpublished data, U.S. Department of Labor, Employment and Training Administration, Unemployment Insurance Service.

[f]U.S. Department of Commerce, Bureau of the Census, *Current Population Reports*, Series P-60.

Table of Contents

In a number of estimated models, the estimated coefficients for the additional explanatory variables were numerically small or statistically insignificant. The magnitudes of the out-of-sample forecast errors for these models were usually similar to those of the out-of-sample forecast errors from the basic model estimated without additional variables. Because the text here focuses primarily on potentially interesting alternative models, Appendix B reports the results from the estimated models that did not improve significantly on the basic model.

a. General Economic Conditions

The additional variables that represented general economic conditions included the aggregate number of unemployed workers, the aggregate number of workers who exhaust their unemployment insurance (UI) benefits, and per-capita disposable income. The number of unemployed workers directly captures the size of a population that may need food stamp assistance, and may be correlated with several other general economic conditions. The number of workers who exhaust their UI benefits is a potentially useful explanatory variable because such workers are generally unemployed for at least six months, and may thus need food stamp assistance. Per-capita disposable income is a summary measure of economic prosperity that may also have an impact on program participation. Wage rates and employment levels in the personal services industry were considered, because conditions in this industry may greatly reflect the economic conditions facing the population that is eligible to receive food stamps.

Table III.6 reports the results of two specifications with the added economic variables. Column 1 of Table III.6 reports the estimated coefficients for the basic model (Model P.8), in which the number of unemployed workers was substituted for the unemployment rate. As mentioned earlier, the number of unemployed workers directly captures the size of a population that may need food stamp assistance. Because the unemployment rate is the number of unemployed workers divided by the size of the labor force, changes in the size of the labor force may change the unemployment rate

Table of Contents

TABLE III.6

REGRESSION MODELS OF FSP PARTICIPATION THAT CONTAIN ADDITIONAL EXPLANATORY VARIABLES

	(1) Model P.8	(2) Model P.9	(3) Model P.10	(4) Model P.11
Number of unemployed workers (thousands)	0.35 (0.11)	0.32 (0.11)	0.44 (0.08)	0.44 (0.10)
Number of unemployed workers lagged one quarter (thousands)	0.25 (0.10)	0.24 (0.11)	0.09 (0.08)	0.23 (0.10)
Number of workers employed in the personal services industry (thousands)	--	-1.59 (1.18)	--	--
AFDC recipients (thousands)	--	--	1.56 (0.23)	--
Number of female-headed households with children under 18 (thousands)	--	--	--	2.90 (1.06)
Elimination of purchase requirement (EPR)	5,099.8 (449.7)	5,156.9 (450.1)	4,591.0 (380.2)	4,637.6 (488.8)
OBRA 1981 (REC2)	-675.8 (232.9)	-545.5 (66.3)	-163.1 (210.5)	-666.7 (255.2)
1st quarter dummy	206.1 (73.8)	216.1 (74.4)	198.8 (50.5)	235.5 (65.8)
2nd quarter dummy	509.7 (116.0)	585.1 (128.3)	226.6 (88.5)	476.6 (103.0)
3rd quarter dummy	239.8 (90.6)	265.8 (91.4)	207.3 (64.5)	268.3 (83.2)
Constant	10,099.2 (646.5)	11,881.6 (1,518.2)	-6,183.0 (2,594.6)	-11,764.1 (8,854.0)
ρ	0.78 (0.09)	0.79 (0.10)	0.95 (0.05)	0.95 (0.02)
$\bar{R}^2$	0.99	0.99	0.99	0.99
Standard error of the regression	231.4	228.4	166.3	217.0
Durbin-Watson statistic	1.4	1.3	1.9	1.7
Sample period	FY 1977.1- FY 1989.4	FY 1977.1- FY 1989.4	FY 1977.1- FY 1989.4	FY 1977.1- FY 1989.4

NOTES: Standard errors are in parentheses. The dependent variable is the monthly average of FSP participants (in thousands) during a fiscal-year quarter. The data used for estimation are shown in Appendix D.

in a direction opposite from the direction of the change in the number of unemployed workers. The estimated coefficients for Model P.8 indicate that FSP participation would increase by roughly 6 persons when unemployment grows by 10 workers.[18] The out-of-sample forecasting accuracy of Model P.8 was somewhat lower than the accuracy of Model P.4, which contained the unemployment rate; in Model P.8, the two-year-ahead *rmse* was 272, compared with 248 in Model P.4 (see Table III.7). In general, the overall results were not affected by the inclusion of unemployment levels rather than unemployment rates, but using unemployment levels is preferred because they will increase as the population and FSP participation increase.

Column 2 of Table III.6 reports the estimated coefficients for Model P.9, which contains the number of persons employed in the personal services industry. The estimated coefficient for the personal services variable was negative, but only marginally significant. The value of the estimated coefficient indicates that FSP participation drops by 1.6 persons for each worker employed in the personal services industry. However, the out-of-sample forecasting accuracy for Model P.9 was low (see Table III.7), with a two-year-ahead *rmse* of 442, compared with 272 for Model P.8, which did not include the personal services variable.[19]

b. Demographic Composition

Aggregate FSP participation may be affected by shifts in the demographic composition of the population, as discussed in the conceptual framework. The variables used to account for demographic

---

[18]The estimated coefficients cannot be interpreted to mean that the FSP participation rate among unemployed workers is 60 percent. The unemployment variable is correlated with a host of economic factors, each of which may affect FSP participation.

[19]The relatively low Durbin-Watson statistics for Models P.8 and P.9 indicate that these models may contain higher-order serial correlation. However, statistical tests for second-order serial correlation yielded ambiguous results, depending on the model that was specified. This pattern suggests that the second-order serial correlation is due to variables that are excluded from the model, rather than to the structure of the error term.

Table of Contents

TABLE III.7

MEASURES OF OUT-OF-SAMPLE FORECAST ERROR OF REGRESSION MODELS OF FSP PARTICIPATION THAT CONTAIN ADDITIONAL EXPLANATORY VARIABLES
(Participants in Thousands)

	(1) Model P.4	(2) Model P.8	(3) Model P.9	(4) Model P.10	(5) Model P.11
Two Years Ahead					
mfe					
FY 1988	-324.0	-133.0	276.3	334.0	-302.5
FY 1989	92.9	316.5	556.0	613.8	-100.3
Average	-115.6	91.8	416.2	473.9	-201.4
rmse					
FY 1988	349.4	191.0	305.3	364.6	311.6
FY 1989	146.2	352.3	577.8	620.7	142.0
Average	247.7	271.7	441.6	492.7	226.8
One Year Ahead					
mfe					
FY 1988	-92.5	10.0	171.5	230.8	-94.8
FY 1989	105.5	226.3	349.8	155.8	-59.0
Average	6.5	118.1	260.1	193.3	-78.7
rmse					
FY 1988	146.3	116.5	212.7	269.2	120.4
FY 1989	157.7	277.0	393.6	173.7	113.1
Average	152.0	196.8	303.2	221.5	116.8
Six Months Ahead					
mfe					
FY 1988.3-FY 1988.4	189.0	202.0	191.0	250.5	179.0
FY 1989.3-FY 1989.4	254.0	338.2	378.0	157.5	168.0
Average	221.5	270.1	284.5	204.0	173.5
rmse					
FY 1988.3-FY 1988.4	191.2	202.2	191.3	253.6	179.2
FY 1989.3-FY 1989.4	260.7	353.0	392.7	168.4	179.7
Average	226.0	277.6	292.0	211.0	179.5

NOTE: Models were estimated through FY 1986 and FY 1987 to generate two-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; through FY 1987 and FY 1988 to generate one-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; and through FY 1988.2 and FY 1989.2 to generate six-month-ahead out-of-sample forecasts for FY 1988.3-FY 1989.4 and FY 1989.3-FY 1989.4, respectively. The estimation results for the models used to generate out-of-sample forecasts are included as supporting tables in a separate volume.

Table of Contents

composition include the number of female-headed households with children younger than age 18, the number of households with children younger than 18 that are below the poverty line, the number of AFDC participants, and the number of SSI participants.

Column 3 of Table III.6 shows the results from a model that includes the number of AFDC participants (Model P.10). The estimated coefficient for the AFDC variable was positive and statistically significant. The value of the coefficient indicates that FSP participation rises by 1.6 persons for every AFDC participant. The magnitude seems reasonable, considering the categorical eligibility of AFDC participants for food stamp benefits and the correlation between AFDC participation and the size of the FSP-eligible population. However, the out-of-sample forecasting accuracy of the AFDC model was low, with a two-year-ahead *rmse* of 493, compared with 272 for Model P.8, which did not include AFDC participation.

Column 4 of Table III.6 shows the results when the number of female-headed households with children younger than 18 is included (Model P.11). The estimated coefficient for the female-headed household variable was positive and statistically significant. The coefficient value indicates that FSP participation increases by 2.9 persons for every female-headed household with children younger than 18. The out-of-sample forecasting accuracy of Model P.11 was high, with a two-year-ahead *rmse* of 227, compared with 272 for Model P.8, which did not include the female-headed household variable.

However, the greater forecasting accuracy of Model B.11 may be due to the particular time period chosen for evaluating forecasts (that is, 1988 and 1989). The uniform upward trend of the female-headed household variable means that a model which includes female-headed households with an estimated positive coefficient will generally predict greater FSP participation two years ahead. Because actual FSP participation rose in 1989, Model P.11 predicts participation accurately for that particular year, and thus appears to perform the most effectively among the set of alternative models. However, if participation had fallen in 1989, Model P.11 might have forecast participation poorly for that year. For this reason, the more accurate forecasts from Model P.11 should be viewed with

Table of Contents

caution. Model P.11 must be tested further with future sample periods to determine whether the model does in fact forecast FSP participation accurately in a wider variety of circumstances.

c. Program Parameters

Since FSP participation is affected by the benefits received by households, the maximum benefit allotment for a family of four, the standard deduction, and the excess shelter and child care deductions were entered as explanatory variables. However, when these variables were included in Model P.8, the signs of the estimated coefficients were usually counterintuitive, or the estimated coefficients were not statistically significant. Appendix B shows the results from these models.

d. Legislative Changes

The effects of the EPR and the 1981 OBRA program changes were included in the basic model. In addition to these acts, the Food Security Act of 1985, the Immigration Reform and Control Act of 1986 (IRCA), and the Hunger Prevention Act of 1988 also changed the benefits from or eligibility for the Food Stamp Program. The effects of these three acts were included in the basic model via dummy variables that assumed a value of zero prior to the effective date of the legislated change, and a value of one after the date of the legislative change. However, these variables generally had only small effects on forecast errors relative to the basic model. Appendix B shows the results from models that included additional legislative variables.

4. Confidence Limits for Program Participation Forecasts

The reliability of forecasts of future program participation can be assessed by computing forecast confidence intervals. Confidence intervals for forecasts are statistical estimates of the range of values within which the variables being forecast are likely to fall.

Forecasted values from a regression model could differ from subsequently observed values for four reasons: (1) the model could be specified incorrectly, (2) the values of the explanatory variables used to compute the forecasts may differ from subsequently observed values, (3) the estimated

Table of Contents

coefficients of the model may differ from the true coefficients due to sampling variability, and (4) random variation could affect program participation and benefits. The confidence intervals computed in this section take into account the third and fourth sources of variation.[20]

Table III.8 shows forecasts of program participation for 1988 and 1989 from the basic participation model (Model P.8), and the upper and lower 95 percent confidence limits for each forecasted value.[21] The confidence intervals for the 1989 quarterly forecasts of participation are plus or minus 1.1 million participants for the first quarter of 1989, growing to 1.2 million participants by the fourth quarter of 1989. In percentage terms, the confidence intervals for participation in 1988 and 1989 are plus or minus 6 to 7 percent of the average forecasted value for participants per month.

Figures III.1 and III.2 plot program participation forecasts, confidence limits for the forecasts, and actual program participation for 1988 and 1989. The vertical lines in the figures indicate the points at which the sample periods end and the out-of-sample forecasting periods begin. The confidence intervals display a modest widening as forecasts are made further into the future. The increase in the width of the confidence intervals as forecasts are made further into the future is due to the diminishing serial correlation effect, which was discussed earlier in Section III.A. Moreover, because forecasts of the explanatory variables are themselves less accurate further into the future, the true confidence limits will be wider than those shown in Figures III.1 and III.2.

[20]The second source of variation is also important. In the context of FSP participation models, the primary explanatory variable is unemployment, which is forecast by the Office of Management and Budget. Appendix A discusses the accuracy of OMB unemployment-rate forecasts from 1980 to 1989 and the effects that unemployment-rate forecast error would have on participation forecast confidence intervals. Illustrative calculations in Appendix A suggest that the forecast errors for the unemployment rate substantially increase the forecast errors for FSP participation.

[21]The confidence limits were calculated with SAS Proc AUTOREG, which incorporates the improvement in forecasting accuracy due to the first-order autoregressive structure of the error term. However, AUTOREG does not account for the additional variance introduced from using an estimated value of the autocorrelation parameter when participation is forecast. The confidence limits shown in Table III.8 are thus somewhat more narrow than the theoretically correct limits. In practice, the increase in the width of the confidence interval due to the estimation of the autocorrelation parameter is small (Harvey, 1981).

Table of Contents

TABLE III.8

CONFIDENCE INTERVALS FOR TWO-YEAR-AHEAD OUT-OF-SAMPLE FORECASTS OF FSP PARTICIPATION

	(1) Participation Forecast (millions)	(2) Lower Bound of 95 Percent Confidence Interval (millions)	(3) Upper Bound of 95 Percent Confidence Interval (millions)
FY 1988			
Quarter 1	18.7	17.5	19.8
Quarter 2	19.2	18.0	20.4
Quarter 3	18.8	17.6	20.0
Quarter 4	18.4	17.1	19.6
FY 1989			
Quarter 1	18.3	17.2	19.4
Quarter 2	18.8	17.7	20.0
Quarter 3	18.6	17.4	19.7
Quarter 4	18.2	17.0	19.4

NOTE: Model P.8 in Table III.6 was estimated through FY 1986 and FY 1987 to calculate the two-year-ahead out-of-sample forecasts and their confidence limits for FY 1988 and FY 1989, respectively. SAS Proc AUTOREG was used to compute the confidence limits.

Table of Contents

FIGURE III.1

CONFIDENCE LIMITS FOR FSP PARTICIPATION MODEL P.8

OUT-OF-SAMPLE FORECAST: 1987-1988

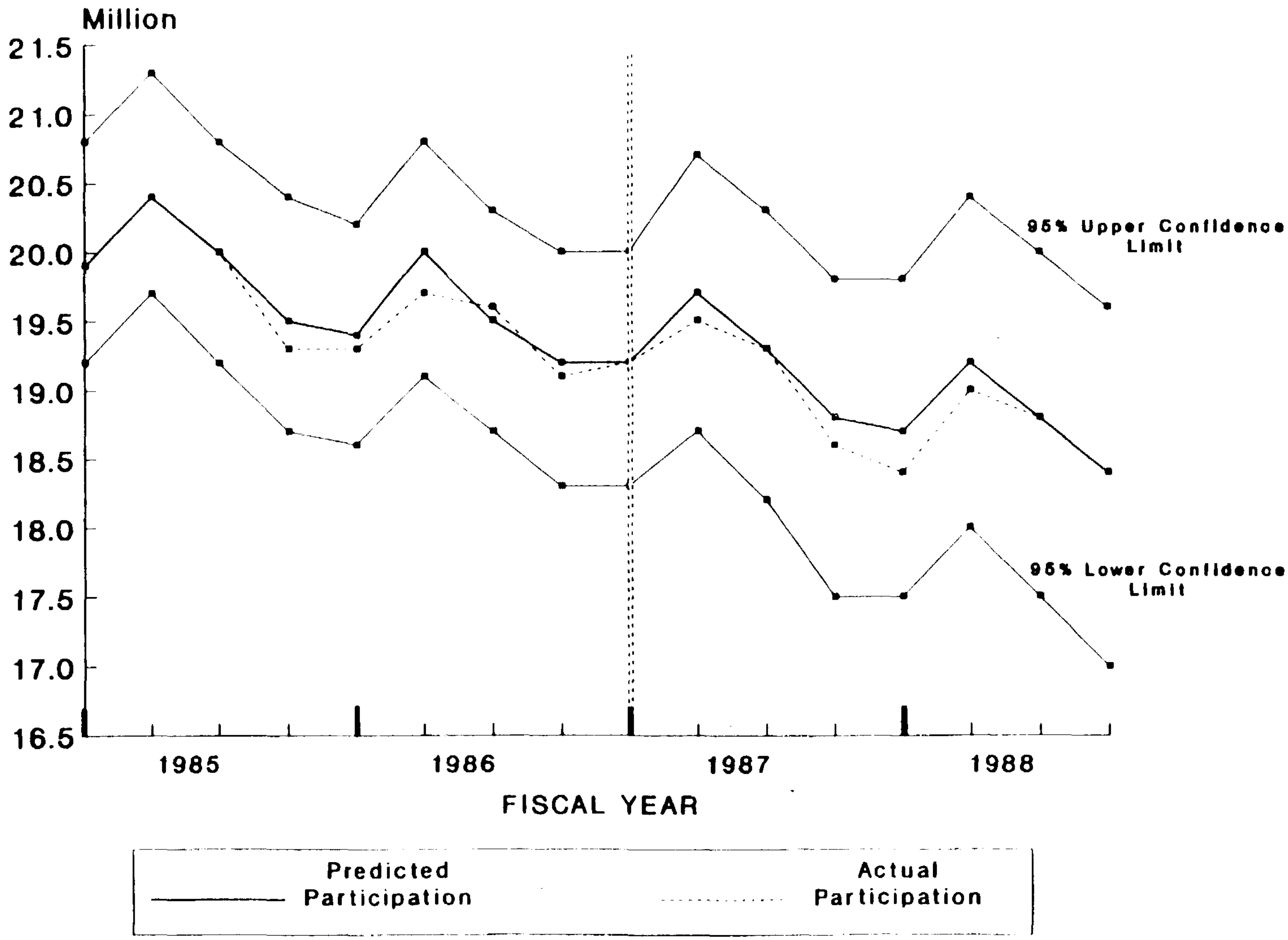


Table of Contents

FIGURE III.2

CONFIDENCE LIMITS FOR FSP PARTICIPATION MODEL P.8
OUT-OF-SAMPLE FORECAST: 1988-1989

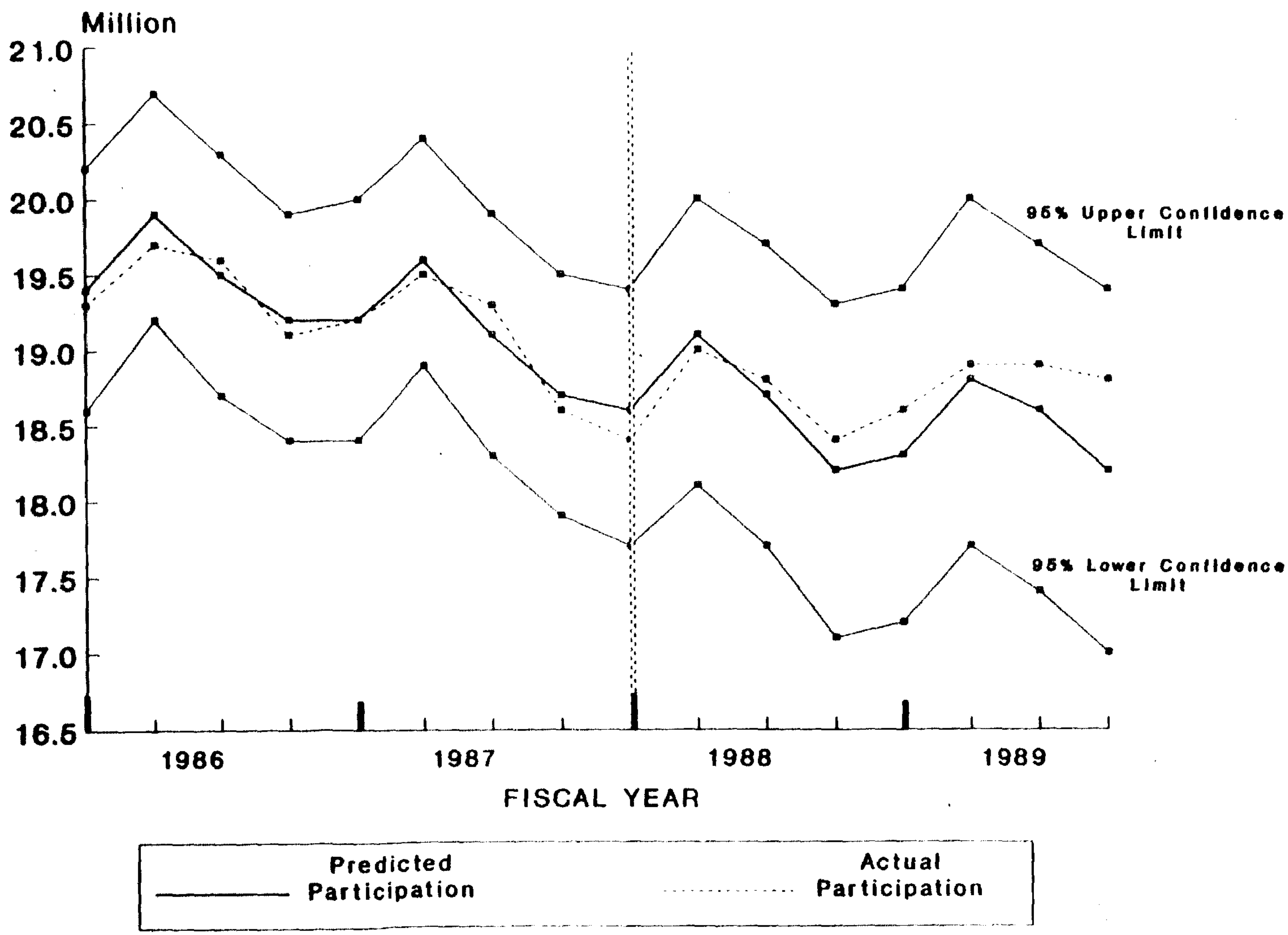


Table of Contents

C. FORECASTING MODELS OF AVERAGE FOOD STAMP BENEFITS

This section examines two methods for forecasting average food stamp benefits. The first method, which FNS used through FY 1989, is based on a formula derived from simulation methods to predict average benefits received. A modified formula is also examined, which predicts average benefits out-of-sample more precisely than does the FNS formula.

Regression models of average benefits are then examined, and their forecasting accuracy is compared with the forecasting accuracy of formula-based models.

1. Formula-Based Forecasting Models of Average Benefits

Through FY 1989, FNS forecast the average value of benefits received by program participants according to a formula that adjusted actual benefits received in 1981 for increases in the value of the maximum benefit allotment, the standard deduction, and the excess shelter and child care deductions for each year since 1981.[22] The formula decays the base value of average benefits in 1981 by roughly 3 percent in each subsequent year to allow for income growth, and uses quarterly decay factors within the fiscal year to account for the within-year pattern of decline exhibited by average benefits.[23] An adjustment of $0.35 was added in 1985, and an additional $0.50 was added in 1986 to close a gap between forecasted and actual average benefits. The FNS benefit forecasting model as of 1989 can be written as:

$$(B.1)\quad AB_{qt} = .85 + .97^{(t-1983)} d_q 40.18 + .245(MBA_t - 233) + .092(SD_t - 85) + .018(ESCC_t - 115)$$

where AB_{qt} is the average monthly benefit per participant in year t quarter q, d_q are the quarterly decay factors, MBA is the maximum benefit allotment, SD is the standard deduction, $ESCC$ is the

---

[22]The formula uses pass-through factors that represent the effect of changes in deductions and the maximum benefit allotment on average food stamp benefits. Due to the complexity of benefit eligibility determinations and the interactions among deductions, the pass-through factors are generated with a simulation methodology. FNS updates these estimated pass-though factors periodically. Recent updates indicate that the factors have been highly stable over time.

[23]The quarterly decay factors were 0.9925, 0.9875, 0.9750, and 0.9690, respectively.

Table of Contents

excess-shelter/child-care deduction, and the coefficients that multiply the deductions are the pass-through factors discussed in footnote 22.[24]

The formula had a two-year-ahead *rmse* of $0.17 per month. This *rmse* represents an error of 0.3 percent from the average benefit of $51.86 in 1989. However, because average benefits are multiplied by the total number of forecasted participants to arrive at a budgetary estimate for program benefits, relatively small errors in forecasting average benefits have significant budgetary implications. For example, assuming 20 million FSP participants per month, an average benefit forecast that was too low by $0.16 per month would generate a budgetary shortfall of $40.8 million a year.

Four modifications to the formula were made to improve its forecasting performance: (1) the value of average benefits in the previous fiscal year was substituted for the value of average benefits in 1981, (2) gross income deductions and maximum benefit allotment values in the previous fiscal year were substituted for 1981 program values, (3) the annual 3 percent decay factor and the $0.85 adjustment were deleted, and (4) the quarterly decay factors, d_q, were calculated as one minus the average percentage decline in benefits from one quarter to the next quarter over the 1983-1987 period.[25] The motivation for these changes was to benchmark forecasts of average benefits to the best available information on average benefits, which is the average benefit from the previous year. In formula terms, the modified model can be written as:

$$(B.2)\quad AB_{qt} = \quad d_q AB_{q-1,t} + .245*(TFP_t - TFP_{t-1}) + .092*(SD_t - SD_{t-1}) + .018*(ESCC_t - ESCC_{t-1}).$$

---

[24]For Model B.1, the numerical values in parentheses are the 1981 values of the Thrifty Food Plan, the standard deduction, and the excess-shelter/child-care deduction.

[25]The modified estimates for the quarterly decay factors were .996 for the fourth quarter of the previous fiscal year to the first quarter of the next fiscal year, .998 for the first to the second quarters, .989 for the second to the third quarters, and .996 for the third to the fourth quarters. These factors were calculated by subtracting the pass-through values of the maximum benefit allotment and income deductions for a fiscal year from the actual average benefits in each quarter, beginning in FY 1983 and ending in FY 1987. This procedure yielded a time series of average benefits adjusted for changes in deductions and maximum benefit allotments. With this series, the percentage decline in benefits from one quarter to the next quarter was calculated for each of the five years. Quarterly decay rates were defined as one minus the average percentage decline in average benefits from one quarter to the next during the five-year period.

The modified formula yielded accurate out-of-sample forecasts of average benefits, with a two-year-ahead *rmse* of \$0.10 and a two-year-ahead *mfe* of -\$0.03. The *rmse* for the modified formula represents a significant improvement over the *rmse* for the basic formula (\$0.10, compared with \$0.16).

2. Regression Models of Average Benefits

FNS average benefit forecasts since FY 1989 have been based on a regression model that includes program parameters and seasonal dummy variables as explanatory variables. However, the lower out-of-sample forecasting ability of the regression models suggests that the modified formula may be the better approach for forecasting average benefits.

Column 1 of Table III.9 reports the estimated coefficients for the FNS regression model currently being used by FNS to forecast average benefits. The model includes the maximum benefit allotment, the average monthly net income eligibility limit for FSP participants, a dummy variable for periods after FY 1987, and seasonal dummy variables. The model is based on data beginning in 1982.

The estimated coefficients for the FNS average benefit model were reasonable. A one-dollar increase in the maximum benefit allotment increased the average benefit by \$0.23, and an increase in the net income eligibility limit of \$100 reduced the average benefit by \$0.90. The R^2 statistic was .99, and the Durbin-Watson statistic indicated that the error term was not serially correlated.

Table III.10 shows the out-of-sample forecasting statistics for regression models of average benefits. The out-of-sample forecasting accuracy of the model was similar to the accuracy of the FNS average benefit formula discussed earlier, with a two-year-ahead *rmse* of \$0.17, compared with \$0.16 for the FNS formula model.

Two variations of the FNS average benefit regression model were estimated. The first variation added the standard deduction and the excess-shelter/child-care deduction to the FNS regression model. Column (2) of Table III.9 shows the results from this model. The estimated coefficients

Table of Contents

TABLE III.9

REGRESSION MODELS OF AVERAGE MONTHLY BENEFITS

	(1) Model B.3	(2) Model B.4	(3) Model B.5
Maximum allotment for a family of four	0.23 (0.01)	0.22 (0.01)	0.23 (0.01)
Net monthly household income eligibility limit for a family of four	-0.009 (0.002)	-0.011 (0.003)	--
FY 1988 dummy	-0.16 (0.22)	-0.23 (0.24)	--
Standard deduction	--	0.014 (0.020)	--
Per-capita disposable income (thousands of dollars)	--	--	-0.44 (0.15)
1st quarter dummy	0.83 (0.09)	0.83 (0.09)	0.54 (0.13)
2nd quarter dummy	0.75 (0.09)	0.75 (0.09)	0.58 (0.12)
3rd quarter dummy	0.25 (0.09)	0.25 (0.09)	0.17 (0.10)
Constant	-7.20 (1.29)	-7.73 (1.51)	-9.83 (1.49)
ρ	--	--	0.27 (0.12)
$\bar{R}^2$	0.99	0.99	0.99
Standard error of the regression	0.18	0.17	0.22
Durbin-Watson statistic	1.90	2.07	1.83
Sample period	FY 1982.1- FY 1989.4	FY 1982.1- FY 1989.4	FY 1982.1- FY 1989.4

NOTES: Standard errors are in parentheses. The dependent variable is the average monthly benefit received by FSP participants during a fiscal-year quarter. The data used for estimation are shown in Appendix D.

Table of Contents

TABLE III.10

MEASURES OF OUT-OF-SAMPLE FORECAST ERROR FOR MODELS OF AVERAGE MONTHLY BENEFITS (Dollars)

	(1)[a] Model B.3	(2) Model B.4	(3) Model B.5
Two Years Ahead			
mfe			
FY 1988	-0.16	-0.20	0.28
FY 1989	-0.01	-0.08	0.77
Average	-0.09	-0.14	0.51
rmse			
FY 1988	0.24	0.27	0.35
FY 1989	0.10	0.12	0.75
Average	0.17	0.20	0.55
One Year Ahead			
mfe			
FY 1988	-0.16	-0.21	0.33
FY 1989	0.15	0.15	0.50
Average	0.00	-0.03	0.41
rmse			
FY 1988	0.25	0.29	0.39
FY 1989	0.17	0.16	0.51
Average	0.21	0.23	0.45
Six Months Ahead			
mfe			
FY 1988.3-FY 1988.4	0.08	0.08	0.37
[illegible]	[illegible]	[illegible]	[illegible]

Table of Contents

differed considerably from the coefficients for the FNS model, but the differences were probably due to the high degree of correlation among the program parameter variables. For example, the simple correlation between the standard deduction and the excess-shelter/child-care deduction was .98. Under these circumstances, ordinary least squares yields unstable parameter estimates. The out-of-sample forecasting accuracy of this model was also poor, with a two-year-ahead *rmse* of $1.44, compared with $0.17 for the FNS regression model.

The second model variation included disposable income and the maximum benefit allotment, and excluded average net income eligibility, the standard deduction, and the excess-shelter/child-care deduction. In preliminary runs, this model exhibited a low Durbin-Watson statistic, and a correction for serial correlation was thus included in the model shown in Table III.9. The signs and magnitudes of the estimated coefficients were reasonable. A one-dollar increase in the maximum benefit allotment increased average benefits by $0.23 per participant. This value is similar to the value of the pass-through factor of $0.245 used in the average benefit formulas discussed in the previous section. Higher average disposable income reduced average benefits, with a $1,000 increase in disposable income reducing average benefits by $0.44. However, the out-of-sample forecasting accuracy of this model was poor, with a two-year-ahead *rmse* of $0.55, compared with $0.17 for the FNS regression model.

D. REGRESSION MODELS OF TOTAL PROGRAM BENEFITS

The strategy used thus far has been to treat total program benefits as the product of two components--the total number of participants and the average benefit per participant--and to forecast each of the components separately. An alternative approach is to use a regression model to forecast total program benefits directly. Natural choices for the variables to be included in this regression model are those that have thus far been shown to affect participation and average benefits.

Three models of total benefits were specified. The first model included the unemployment level, the lagged unemployment level, the maximum benefit allotment, the net income eligibility limit, and seasonal dummy variables. The second model included the variables in the first model, as well as the number of female-headed households with children younger than 18. The third model included the variables in the second model, as well as disposable income. On the basis of the Durbin-Watson statistics in preliminary runs, a serial correlation correction was included in all models.

Table III.11 reports the estimation results for three regression models of total program benefits. In the first model, the current and lagged unemployment level increased total benefits, though the statistical significance of the estimated coefficients was low. The maximum benefit allotment had a strong positive effect on total program benefits, with a one-dollar increase in the maximum benefit allotment adding $4.3 million to total monthly program benefits. The net income eligibility limit also had a positive effect on total program benefits, with a one-dollar increase in the eligibility limit adding $0.71 million to total monthly program benefits. Total program benefits show a strong seasonal pattern.

The second regression model for total program benefits included the number of female-headed households, which as shown earlier is a useful explanatory variable for FSP participation. The results in column (2) of Table III.11 indicate that the estimated coefficients are somewhat sensitive to the inclusion of the female-headed household variable. The magnitudes of the unemployment coefficients increased, and the magnitudes of the program parameters decreased. The estimated coefficient for female-headed households was positive and statistically significant, with an additional female-headed household adding $147 to total program benefits.

The third regression model for total program benefits added disposable income to the second model. The estimated coefficients were virtually the same as those of the second model, and the coefficient for disposable income itself was statistically insignificant.

Table III.12 compares the out-of-sample forecasting accuracy of the three total-benefit models. The two-year-ahead *rmse* was $38 million per month for Model T.1, $21 million per month for Model

Table of Contents

TABLE III.11

REGRESSION MODELS OF TOTAL PROGRAM BENEFITS

	(1) Model T.1	(2) Model T.2	(3) Model T.3
Number of unemployed workers (millions)	7.86 (7.86)	16.33 (8.18)	16.34 (8.52)
Number of unemployed workers lagged one quarter (millions)	9.85 (7.29)	7.56 (6.66)	7.56 (7.21)
Maximum allotments for a family of four	4.30 (0.59)	4.04 (0.56)	4.06 (0.58)
Net monthly household income eligibility limit for a family of four	0.71 (0.29)	0.46 (0.29)	0.46 (0.30)
Number of female-headed households with children under 18 (millions)	--	146.93 (65.87)	147.80 (67.75)
Per-capita disposable income (thousands of dollars)	--	--	-0.19 (14.70)
1st quarter dummy	3.24 (9.96)	12.57 (10.12)	12.52 (10.46)
2nd quarter dummy	28.09 (8.78)	26.78 (8.04)	26.74 (8.31)
3rd quarter dummy	11.75 (6.30)	15.41 (5.94)	15.40 (6.29)
Constant	-1,225.7 (405.9)	-2,281.18 (743.59)	-2,293.10 (944.6)
ρ	0.96 (0.03)	0.98 (0.02)	0.98 (0.02)
$\bar{R}^2$	0.96	0.96	0.96
Standard error of the regression	10.40	9.57	9.81
Durbin-Watson statistic	1.98	1.85	1.85
Sample period	FY 1982.1- FY 1989.4	FY 1982.1- FY 1989.4	FY 1982.1- FY 1989.4

NOTES: Standard errors are in parentheses. The dependent variable is the monthly average of total benefits (in millions) received by FSP participants during a fiscal-year quarter. The data used for estimation are shown in Appendix D.

TABLE III.12

MEASURES OF OUT-OF-SAMPLE FORECAST ERROR FOR REGRESSION MODELS OF TOTAL FSP BENEFITS
(Millions of Dollars per Month)

	(1) Model T.1	(2) Model T.2	(3) Model T.3
Two Years Ahead			
mfe			
FY 1988	-23.72	11.38	65.30
FY 1989	50.21	24.80	148.00
Average	13.25	18.09	106.65
rmse			
FY 1988	25.03	13.69	68.03
FY 1989	51.28	28.84	149.55
Average	38.15	21.26	108.79
One Year Ahead			
mfe			
FY 1988	32.96	3.25	83.00
FY 1989	-4.10	-4.60	-3.88
Average	14.43	-0.68	39.56
rmse			
FY 1988	33.98	10.29	85.39
FY 1989	8.94	7.87	7.37
Average	21.46	9.08	46.38
Six Months Ahead			
mfe			
FY 1988.3-FY 1988.4	12.65	10.80	30.15
FY 1989.3-FY 1989.4	16.15	11.05	11.68
Average	14.40	10.93	20.92
rmse			
FY 1988.3-FY 1988.4	13.84	11.73	32.39
FY 1989.3-FY 1989.4	17.50	13.31	14.01
Average	15.67	12.52	23.20

NOTE: Models were estimated through FY 1986 and FY 1987 to generate two-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; through FY 1987 and FY 1988 to generate one-year-ahead out-of-sample forecasts for FY 1988 and FY 1989, respectively; and through FY 1988.2 and FY 1989.2 to generate six-month-ahead out-of-sample forecasts for FY 1988.3-FY 1988.4 and FY 1989.3-FY 1989.4, respectively. The estimation results for the models used to generate out-of-sample forecasts are included as supporting tables in a separate volume.

T.2, and $109 million per month for Model T.3. One-year-ahead and six-month-ahead *rmse* values had the same pattern.

E. COMPARING ALTERNATIVE APPROACHES FOR FORECASTING TOTAL PROGRAM BENEFITS

Two approaches were considered for forecasting total program benefits. The first approach was to separate total benefits into two components--program participation and average benefits per participant--and to use regression models or formula-based models to forecast each component. The second approach was to estimate a regression model of total program benefits. The utility of the two approaches for forecasting FSP benefits can be compared by combining particular participation and average benefit models to generate forecasts in terms of total program benefits. The total benefit forecasts from combined participation and benefits models can then be compared with benefit forecasts from the total benefit regression models.

Table III.13 shows out-of sample participation forecasts for 1988 and 1989 for two regression models of participation, the FNS model (Model P.2) and the "basic" model (Model P.8).[26] Column 3 of Table III.13 shows actual participation in 1989. The forecasts of the two models are similar in 1988, except for the fourth quarter. For the last two quarters of 1989, the participation forecast of Model P.8 was greater by 600,000 than the forecast of Model P.2, and actual participation in those quarters exceeded even these forecasts.

Table III.14 shows out-of-sample average benefit forecasts for 1988 and 1989 from three models: (1) the original FNS formula, Model B.1, (2) the modified formula, Model B.2, and (3) the average benefit regression, Model B.3. Column 4 of Table III.14 shows actual average benefits in 1988 and

[26]Model P.8 generated somewhat less accurate forecasts than did Model P.4. However, Model P.8 was chosen for comparison purposes because its inclusion of unemployment levels as an explanatory variable is more consistent with food stamp participation than Model P.4's inclusion of unemployment rates.

Table of Contents

TABLE III.13

COMPARISON OF ALTERNATIVE OUT-OF-SAMPLE FORECASTS OF MONTHLY FSP PARTICIPATION IN FY 1988 AND FY 1989
(Millions of Participants)

	(1)[a] FNS Model P.2	(2)[a] Model P.8	(3) Actual Participation
FY 1988			
Quarter 1	18.7	18.7	18.4
Quarter 2	19.2	19.2	19.0
Quarter 3	18.7	18.8	18.8
Quarter 4	18.0	18.4	18.4
Average	18.6	18.8	18.7
mfe	.009	-.133	
rmse	.232	.191	
FY 1989			
Quarter 1	18.1	18.3	18.6
Quarter 2	18.6	18.8	18.9
Quarter 3	18.2	18.6	18.9
Quarter 4	17.6	18.2	18.8
Average	18.1	18.5	18.8
mfe	.640	.317	
rmse	.701	.352	
Two Years Ahead			
Average *mfe*	.325	.092	
Average *rmse*	.467	.242	

NOTE: The forecasts refer to the monthly average of FSP participants (in millions) during the fiscal-year quarter.

[a]Out-of-sample forecasts for FY 1988 and FY 1989 were generated from models estimated through FY 1986 and FY 1987, respectively, and based on the actual values of explanatory variables in FY 1988 and FY 1989.

Table of Contents

TABLE III.14

COMPARISON OF ALTERNATIVE OUT-OF-SAMPLE FORECASTS OF AVERAGE MONTHLY FSP BENEFITS IN FY 1988 AND FY 1989
(Dollars)

	(1) FNS Formula B.1	(2) Modified FNS Formula B.2	(3) Regression Model B.3[a]	(4) Actual Average Monthly Benefit
FY 1988				
Quarter 1	50.28	50.15	50.39	50.21
Quarter 2	50.11	50.08	50.29	49.99
Quarter 3	49.70	49.66	49.78	49.49
Quarter 4	49.50	49.51	49.47	49.62
Average	49.90	49.85	49.98	49.82
mfe	-0.07	-0.02	-0.16	
rmse	0.14	0.12	0.24	
FY 1989				
Quarter 1	52.09	52.23	52.26	52.21
Quarter 2	51.93	52.16	52.22	52.10
Quarter 3	51.33	51.74	51.70	51.59
Quarter 4	51.34	51.59	51.37	51.55
Average	51.72	51.93	51.89	51.87
mfe	0.19	-0.04	-0.01	
rmse	0.20	0.08	0.10	
Two Years Ahead				
Average *mfe*	0.06	-0.03	-0.09	
Average *rmse*	0.17	0.10	0.17	

NOTE: The forecasts refer to the monthly average benefits received by FSP participants in the fiscal-year quarter.

[a]Two-year-ahead out-of-sample forecasts for the regression model in column 3 were generated with estimates from Model B.3 in Table III.9, based on data through FY 1986 and FY 1987.

1989. The forecasts from all three models were generally close to actual average benefits. However, the modified formula tended to overforecast benefits, whereas the FNS formula tended to underforecast benefits.

Table III.15 shows out-of-sample total program benefit forecasts for 1988 and 1989 from three models: (1) the FNS participation model (Model P.2) combined with the FNS average benefit regression model (B.3), (2) the basic participation model (Model P.8) combined with the modified benefit formula (B.2), and (3) the total benefit regression model (Model T.2). Each of the three models underforecast total program benefits for 1989, especially in the last two quarters of the year. The total benefit regression model, Model T.2, underforecast benefits by $12 million per month in the third quarter of 1989 and by $32 million per month in the fourth quarter of 1989. The *rmse* values were close in 1988, but were quite different in 1989, equalling $53 million for the combined FNS model, $17 million for Model P.8 combined with the modified benefit formula, and $29 million for Model T.2.

The evidence suggests that Model P.8 (column 1 of Table III.6) combined with the modified average benefit formula may provide the most accurate forecasts of total program benefits among the set of alternative models. However, the total benefit regression model, Model T.2, performed relatively well considering its modest size, and should be considered an attractive alternative.

If a participation forecast is required independently of the total benefit forecast, the total benefit regression model could be used in combination with a participation model, such as Model P.8. A forecast of average benefits is implied by the ratio of forecasted total benefits to forecasted participation. For example, based on the 1989 forecasts of total benefits from the total benefit regression model, Model T.2, and participation from Model P.8, the implied 1989 quarterly average benefit forecasts would be $52.33, $51.97, $50.73, and $50.67. The last two values are somewhat low relative to the forecasts of the other models, but, overall, the implied average benefit forecasts are similar to the direct forecasts from the other average benefit models.

Table of Contents

TABLE III.15

COMPARISON OF ALTERNATIVE TWO-YEAR-AHEAD OUT-OF-SAMPLE FORECASTS OF MONTHLY FSP BENEFITS IN FY 1988 AND FY 1989
(Millions of Dollars)

	(1) FNS Model P.2, FNS Benefit Regression B.3	(2) Model P.8, Modified FNS Benefit Formula B.2	(3) Total Benefit Regression Model T.2	(4) Actual Total Program Benefits
FY 1988				
Quarter 1	940.4	938.8	922.7	926.1
Quarter 2	963.0	962.7	942.4	947.8
Quarter 3	932.3	934.9	917.3	931.2
Quarter 4	891.7	908.9	889.3	912.0
Average	931.9	936.3	917.9	929.3
mfe	-2.6	-7.1	11.4	
rmse	14.6	10.1	13.7	
FY 1989				
Quarter 1	943.9	957.6	957.7	969.7
Quarter 2	969.3	981.1	977.0	986.5
Quarter 3	941.6	960.7	943.6	976.4
Quarter 4	872.3	939.8	922.2	967.0
Average	931.8	959.8	950.1	974.9
mfe	43.1	15.1	24.8	
rmse	52.8	17.0	28.8	
Two Years Ahead				
Average *mfe*	20.3	4.0	18.1	
Average *rmse*	33.7	13.6	21.3	

NOTE: The forecasts refer to total FSP monthly benefits in the fiscal-year quarter.

Table of Contents

Three approaches for forecasting FSP benefits were not pursued here, but may be interesting avenues for future investigation. The first approach would be to estimate separate regression models of FSP participation for each state, to use the estimated models to forecast participation for each state, and to aggregate the state forecasts to arrive at a national forecast. Implementing this approach may be difficult if unemployment levels or rates are unavailable at the state level, but it would potentially yield more accurate forecasts if the estimated state-level regression models differed considerably from the estimated aggregate model.

The second approach would be to view the FSP participation model and the average benefits models as two equations whose random error terms may be correlated. For example, random factors may increase FSP participation at the same time that they increase average benefits. The two-equation model could be estimated simultaneously as a *seemingly unrelated* regression model, which, if the error terms of the two equations are correlated (Johnston, 1984), will generate more precise parameter estimates than separate estimation. However, in general, forecasting experience with these models to date has been limited.

The third approach would be to consider FSP participation and average benefits as contemporaneous functions of each other, leading to a simultaneous-equations model. In this case, sophisticated estimation techniques, such as two-stage least squares, would be required (Johnston, 1984).

F. SUMMARY

A variety of FSP participation and benefit models have been estimated and assessed in terms of their forecasting accuracy. The key findings are as follows:

- The participation models that were estimated generally yielded two-year-ahead participation forecasts that were accurate to within plus or minus 6 to 7 percent per month, or, equivalently, to within plus or minus roughly 1 million participants per month. Even if future average benefits were known with certainty, this level of forecasting accuracy implies that models may have forecast errors on the order of plus

Table of Contents

or minus $840 million annually. Larger errors may occur if the forecasts of macroeconomic quantities on which the participation forecasts are based are inaccurate.

- The forecasting performance of some participation models was marginally better than the performance of others. Forecasts generated by a participation model whose explanatory variables included the number of unemployed workers, variables for seasonality, and a correction for the correlation of random factors over time were the most accurate.

- The forecasting accuracy of the limited number of regression models of average benefits that were estimated was generally inferior to the accuracy of a formula approach for forecasting average benefits that relied on parameters estimated with a simulation methodology.

- A regression model of total program benefits provided forecasts whose accuracy was similar to the accuracy of forecasts from two-equation participation and average benefit models.

FSP participation and benefits are affected by numerous factors that cannot be captured fully in any one model. The structure of the Food Stamp Program and the structure of other related programs evolve continuously, as does the structure of the economy and public perceptions of the program. In this context, the best forecasting model at any particular point in time may become outmoded quickly, and frequent assessments of a model's forecasting performance are sensible. In particular, the alternative models explored in this chapter have been evaluated in part according to their ability to predict participation in 1989, a year in which historical patterns of participation may have shifted. Future assessments are necessary to determine whether the models that performed well here continue to do so during more stable periods.

Table of Contents

IV. A FRAMEWORK FOR FORECASTING FOOD STAMP PROGRAM PARTICIPATION AND BENEFITS

Having considered a number of candidate models for forecasting FSP participation and benefits, it is appropriate to discuss the framework of a forecasting process in which the models will play a role.

The purpose of a forecasting framework is (1) to provide a mechanism for tracking the quality of forecasts over time, (2) to provide a system for updating the empirical model to reflect new information and data, and (3) to provide a vehicle for replicating forecasts. A forecasting process that meets these three criteria can be described in the following six steps, which are discussed in turn:

1. Specifying and estimating the empirical model
2. Obtaining forecasts of the independent variables
3. Generating forecasts of the dependent variable (that is, FSP participation or benefits)
4. Specifying and estimating the effects of out-of-model factors
5. Reporting the results of steps 1-4
6. Evaluating the quality of forecasts

A. SPECIFYING AND ESTIMATING THE EMPIRICAL MODEL

The first step in the forecasting process is to specify an empirical model precisely, and then to estimate the specified model. The necessary specifications include (1) the independent variables to be included in the model, (2) the time period over which the model is to be estimated, (3) the functional form of the model, and (4) the statistical characteristics of the error term. It is important that the model specifications and any changes to the specifications be documented clearly, to facilitate assessing the extent to which changes improve the quality of forecasts.

In general, estimating the model presents no difficulties for the forecasting process. Models that are estimated with ordinary least squares should yield the same estimates with all publicly available

Table of Contents

software packages. However, the techniques used in some packages to estimate more sophisticated models (such as those with serially correlated errors) are different, and it is useful to note these features of the estimation techniques explicitly when the results of the estimation are reported (in step 5).

B. OBTAINING FORECASTS OF THE EXPLANATORY VARIABLES

Some types of models require forecasted values of the explanatory variables to calculate forecasts of the dependent variable. For example, some of the models examined in Chapter III use the unemployment rate as an explanatory variable for FSP participation, thus requiring forecasts of the unemployment rate. Forecasted values for other explanatory variables might also be necessary--for instance, for such demographic variables as the number of female-headed households. Forecasts of the unemployment rate can generally be obtained from the Office of Management and Budget (OMB), and forecasts of some demographic variables are available from the U.S. Bureau of the Census. The approximate date on which the forecasts were made should be noted, since forecasts for a future time period will differ according to when they were generated.

Of course, the quality of FSP participation and benefit forecasts depends on the quality of the other forecasts used in the process. For this reason, it is useful to monitor the quality of the forecasts of the explanatory variables. (Appendix A examines the quality of the OMB forecast of the unemployment rate, and indicates that the OMB forecasts were inaccurate in the early 1980s but have been more accurate in recent years).

C. GENERATING FORECASTS OF THE DEPENDENT VARIABLE

Substituting forecasted values of the independent variables into the empirical model to generate forecasts of the dependent variables, such as FSP participation or average food stamp benefits

Table of Contents

received (see Section III.A), is straightforward.[1] However, it is also useful to calculate a confidence interval of the forecast, to assess the reliability of the forecast itself (see Section III.B.3).

D. SPECIFYING AND ESTIMATING THE EFFECTS OF OUT-OF-MODEL FACTORS

It is generally possible to obtain a better forecast of the dependent variable if additional information is known about the factors that affect the dependent variable in the forecasting period. For example, if one knows that a program change will increase participation in the future, this information could be used to supplement the model forecast, which by design cannot capture this type of information. However, if model forecasts are modified in arbitrary ways, forecasts would be open to the claim that they are unsubstantiated and reflect only the judgment of the forecaster.

A compromise between purely model-based forecasts and arbitrary forecasts is possible if out-of-model adjustments to the model forecasts are accompanied by detailed justifications of the magnitudes used for the adjustments. For example, a prospective program change might be assumed to generate a 5 percent increase in future participation because the change is similar in nature to a previous program change that had generated a 5 percent participation increase. Though it is generally possible only to place a range of values on the effects of prospective program changes or changes in other factors outside the model, it would be useful to document the derivation of the range of values. A later examination of the documentation may reveal insights to help improve future out-of-model forecasts.

A clear distinction should be made between out-of-model adjustments to the model forecasts and adjustments made to compensate for previous model forecast errors. The estimated models discussed in Chapter III account for previous errors via the estimated serial correlation parameter. If out-of-model adjustments to the model forecasts are made to account for previous forecast errors, the errors will in a sense have been accounted for twice, which may generate larger future forecast errors.

---

[1]Generating forecasts of the dependent variable becomes more complex when serial correlation is present. The appropriate method for generating forecasts of the dependent variable for this case was discussed in Section A of Chapter III.

A similar problem would arise if adjustments were made to account for economic factors that may affect future participation but are not explicitly contained in the model. The difficulty with such adjustments is that economic factors excluded from the model may be correlated with economic factors that are included in the model. The model incorporates the effects of excluded economic factors via their correlation with the economic factors that are included. Adjusting forecasts to account for excluded economic factors may mean that the effects of the excluded economic factors will be double-counted, which may increase future forecast errors.

Out-of-model adjustments should reflect the effects of factors that are not captured by the model in any way. Program changes that are scheduled to be implemented in the future are examples of factors that are not captured by the model. However, for the preceding reasons, out-of-model adjustments for factors other than future program changes should be considered carefully, to determine whether some part of the adjustment is not already captured by the model.

E. REPORTING THE RESULTS OF STEPS 1 TO 4

The ultimate objective of the forecasting process is to provide reasonably precise forecasts of FSP participation and benefit receipt. If the results of steps 1-4 are reported in a consistent format over time, it will be possible to evaluate the performance of the forecasting process over time, to ascertain the precision of out-of-model adjustments from accumulated experience, and to verify previous forecasts if necessary. The reporting format could follow steps 1-4, with a preface summarizing the model forecasts and significant adjustments made to the model forecasts due to out-of-model changes. Attachment IV.1 provides a prototype reporting form that could be used for the forecasting process.

F. EVALUATING THE QUALITY OF FORECASTS

By adhering to the process represented by steps 1-5, accumulated forecasting experience will be compiled in a uniform manner that will generate useful input for assessing the forecasting model

Table of Contents

periodically. A straightforward way to assess the quality of the forecasts over time would be to record the actual levels of participation and benefits in future periods on the recordkeeping form that contained the forecasts for those periods. The same procedure could be followed for assessing the accuracy of the forecasts of the explanatory variables on which the forecasts of participation and benefits were based. A simple calculation would then reveal the extent to which forecasts of participation and benefits were inaccurate due to inaccuracies in the forecasts of the explanatory variables. Summary statistics of forecast errors could be produced and accumulated over time.

The purpose of structuring the forecasting process as it is presented in this chapter is to maximize what can be learned from previous forecasting experience. By recording and annotating changes to the forecasting models and the data, it will be possible to determine the sources of improvements in the quality of the forecasts. This information may help guide future evaluations of alternative forecasting models.

Table of Contents

ATTACHMENT IV.1

PROTOTYPE RECORDKEEPING FORM:

FORECASTS OF FOOD STAMP PROGRAM COSTS

DATE:______________________________

PERIOD OF FORECAST:________________

TYPE OF FORECAST

- INITIAL BUDGET _____
- MID-SESSION REVIEW _____
- WITHIN-YEAR REVIEW _____
- OTHER (SPECIFY)

____

A. MODEL OF FSP PARTICIPATION (If using more than one model to predict participation, document information for each model separately.)

1. Sample Period Used in the Estimation:________________________________

2. Frequency of Observations (Monthly or Quarterly):________________________

3. Dependent Variable:____

4. List of Independent Variables:_______________________________________
____
____
____
____
____
____
____

5. Seasonal Adjustment Method: (Check One)

   (a) None _______

   (b) Regression method _______

   (c) Ratio to moving average method
      - Multiplicative _______
      - Additive _______

Table of Contents

(d) Other (specify)

____

If (c) is checked, list seasonal adjustment factors. (This form assumes that quarterly data are used in the estimation. Specify monthly adjustment factors if monthly data are used.)

Quarter 1 ________
Quarter 2 ________
Quarter 3 ________
Quarter 4 ________

6. Error-Term Specification

- Classical ________
- Autoregressive (specify order) ________
- Moving average (specify order) ________
- Other (specify)

____

7. If any item in 6 is checked, specify the estimation technique:

Cochrane-Orcutt iterative least squares ________
Durbin two-stage method ________
Hildreth-Lu search procedure ________
Maximum likelihood ________
Other (Specify) ________

8. Econometric software package used:

SAS ________
TSP ________
Other (Specify) ________

9. Attach to this form the regression output and data used in the estimation.

10. List modifications made to the model since the last date of the forecast, and comment on the reason for the change in specification.

Replacing forecasts with actual values (specify variables)____________________
____
____

Changes in sample period______________________________________
____
____

Changes in program/legislative variables______________________________
____
____

Table of Contents

Changes in other regression variables____
____
____

Other changes (specify)____
____
____

B. FORECASTED VALUES

1. Forecasts of Independent Variables

(a)______________________________ Source of Forecast: OMB______________

Other (Specify):________________________

Date That Forecast Was Generated:__________________

Date That Actual Values Were Recorded:______________________

Forecasted Value	Actual Value	Actual Minus Forecasted Values of Independent Variables
______________	______________	______________
______________	______________	______________
______________	______________	______________
______________	______________	______________
______________	______________	______________
______________	______________	______________
______________	______________	______________
______________	______________	______________

Table of Contents

(b)______________________________ Source of Forecast: OMB____________

Other (Specify):____________________

Date That Forecast Was Generated:__________________ Date That Actual Values Were Recorded:____________________

Forecasted Value	Actual Value	Actual Minus Forecasted Values of Independent Variables
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________

IF MORE THAN TWO FORECASTED EXPLANATORY VARIABLES ARE USED, ATTACH FORECASTED VALUES, SOURCE, AND DATA ON ADDITIONAL SHEET.

2. Forecasts of the Dependent Variable

	Forecasted Value	Actual Value	Forecast Error (FE) (Actual Value Minus Forecasted Value)
<u>One Year Ahead</u>:			
Q1	____________	____________	____________
Q2	____________	____________	____________
Q3	____________	____________	____________
Q4	____________	____________	____________

Table of Contents

Two Years Ahead:

Q5	____________	____________	____________
Q6	____________	____________	____________
Q7	____________	____________	____________
Q8	____________	____________	____________

Out-of-sample forecast statistics:

One year ahead:	
Mean forecast error[a]	________
Root mean square error[b]	________
Two years ahead:	
Mean forecast error	________
Root mean square error	________

---

[a]The mean forecast error (mfe) is the average value of the forecast errors for the period under consideration. The one-year-ahead mfe is:

$$mfe = \sum_{i=Q_1}^{Q4} \frac{FE_i}{4}$$

and the two-year-ahead mfe is:

$$mfe = \sum_{i=Q5}^{Q8} \frac{FE_i}{4}$$

[b]The root mean square error (rmse) is computed by taking the square root of the average of the squared forecast errors. The one-year-ahead rmse is:

$$rmse = \sqrt{\sum_{i=Q1}^{Q4} \frac{FE_i^2}{4}}$$

and the two-year-ahead rmse is:

$$rmse = \sqrt{\sum_{i=Q5}^{Q8} \frac{FE_i^2}{4}}$$

C. AVERAGE BENEFIT FORECASTS

1. Check type of model used for benefit forecasts

   Regression model __________ Formula __________

2. IF A REGRESSION MODEL IS USED TO ESTIMATE BENEFITS, USE THE GUIDELINES SPECIFIED IN SECTIONS A AND B, AND REPORT MODEL ESTIMATES AND FORECASTS.

3. If using a formula to calculate average benefits, specify formula:____________________

4. List all modifications made to the formula since the last forecast date.

   Changes in formula parameters____________________

   Changes in program variables____________________

   Other changes (specify)____________________

5. Forecasts of Variables Used in the Formula to Predict Benefits.

   (a)____________________ Source of Forecast: OMB____________

   Other (Specify):____________________

   Date That Forecast Was Generated:________________ Date That Actual Values Were Recorded:________________

Forecasted Value	Actual Value	Actual Minus Forecasted Values of Independent Variables

Table of Contents

(b)__________________________

Source of Forecast: OMB____________

Other (Specify):____________________

Date That Forecast Was Generated:________________

Date That Actual Values Were Recorded:____________________

Forecasted Value	Actual Value	Actual Minus Forecasted Values of Independent Variables
________	________	________
________	________	________
________	________	________
________	________	________
________	________	________
________	________	________
________	________	________
________	________	________

IF USING MORE THAN TWO FORECASTED VARIABLES, ATTACH SOURCE, GENERATION DATA, AND FORECASTED VALUES ON ADDITIONAL SHEET.

6. Average Benefit Forecasts

	Forecasted Value	Actual Value	Forecast Error (FE) (Actual Value Minus Forecasted Value)
One Year Ahead:			
Q1	________	________	________
Q2	________	________	________
Q3	________	________	________
Q4	________	________	________
Two Years Ahead:			
Q5	________	________	________
Q6	________	________	________
Q7	________	________	________
Q8	________	________	________

Table of Contents

Out-of-sample forecast statistics:

One year ahead:
Mean forecast error ________
Root mean square error ________

Two years ahead:
Mean forecast error ________
Root mean square error ________

D. EFFECTS OF OUT-OF-MODEL FACTORS

1. List all out-of-model adjustments to participation forecasts, and comment briefly on the nature of the adjustments. Attach additional sheets if necessary.

   (a) Prospective program changes ________________________________

   (b) Expected changes in excluded economic factors ________________________________

   (c) Expected changes in excluded demographic factors ________________________________

   (d) Other changes (specify) ________________________________

2. Briefly comment on the magnitudes used for the adjustments, and report the expected out-of-model change in participation over the next eight quarters.

   Comments: ________________________________

Expected Out-of-Model
Changes in Participation

<u>One Year Ahead</u>

Q1 ________
Q2 ________
Q3 ________
Q4 ________

Table of Contents

Two Years Ahead

Q5 __________
Q6 __________
Q7 __________
Q8 __________

3. List all out-of-model adjustments to average benefit forecasts and comment briefly on the nature of adjustments.

(a) Prospective program change____

(b) Other changes (specify)____

4. Comment briefly on the magnitude used for the adjustments, and report the expected out-of-model changes in average benefits over the next eight quarters.

Comments:____

Expected Out-of-Model
Changes in Average Benefits

One Year Ahead

Q1 __________
Q2 __________
Q3 __________
Q4 __________

Two Years Ahead

Q5 __________
Q6 __________
Q7 __________
Q8 __________

Table of Contents

E. FORECASTS OF TOTAL PROGRAM COSTS

1. Type of Model Used for Total Cost Forecasts

   Product of forecasted participation and forecasted average benefits__________ Regression model _____________

2. IF A REGRESSION MODEL OF TOTAL BENEFITS IS USED, FOLLOW THE GUIDELINES SPECIFIED IN SECTIONS A AND B, AND REPORT MODEL ESTIMATES AND FORECASTS. ANY OUT-OF-MODEL ADJUSTMENTS MADE SHOULD BE REPORTED USING THE GUIDELINES IN SECTION D. SKIP STEP 3 AND WRITE TOTAL BENEFIT FORECASTS FROM THE MODEL IN COLUMN (7) OF STEP 4.

3. Compute total benefit forecasts as the product of forecasted participation and forecasted average benefits.

   (a) Write down forecasted participation from Step 2 in Section B and expected out-of-model changes from Step 2 in Section D in Columns 1 and 2, respectively.

	(1) Participation Forecasts from Model	+	(2) Out-of-Model Changes	=	(3) Total Participation Forecasts
One Year Ahead:					
Q1	________		________		________
Q2	________		________		________
Q3	________		________		________
Q4	________		________		________
Two Years Ahead:					
Q5	________		________		________
Q6	________		________		________
Q7	________		________		________
Q8	________		________		________

Table of Contents

(b) Write down forecasted average benefits from Step 5 in Section C and expected out-of-model changes in average benefits from Step 4 in Section D in Columns 4 and 5, respectively.

	(4) Average Benefit Forecasts	+	(5) Out-of-Model Changes	=	(6) Total Average Benefit Forecasts
One Year Ahead:					
Q1	____________		____________		____________
Q2	____________		____________		____________
Q3	____________		____________		____________
Q4	____________		____________		____________
Two Years Ahead:					
Q5	____________		____________		____________
Q6	____________		____________		____________
Q7	____________		____________		____________
Q8	____________		____________		____________

4. Total Benefit Forecasts

	(7) Total Benefit Forecasts Col. (3) times Col. (6)	(8) Actual Total Benefits	(9) Forecast Error (Actual Minus Forecasted Total Benefits)
One Year Ahead:			
Q1	____________	____________	____________
Q2	____________	____________	____________
Q3	____________	____________	____________
Q4	____________	____________	____________
Two Years Ahead:			
Q5	____________	____________	____________
Q6	____________	____________	____________
Q7	____________	____________	____________
Q8	____________	____________	____________

Out-of-sample forecast statistics:

One year ahead:	
Mean forecast error	__________
Root mean square error	__________
Two years ahead:	
Mean forecast error	__________
Root mean square error	__________

Table of Contents

REFERENCES

Burstein, Nancy R., and Mary G. Visher. "The Dynamics of Food Stamp Program Participation." Report prepared for the U.S. Department of Agriculture, Food and Nutrition Service. Cambridge, MA: Abt Associates, Inc., March 1989.

Carr, Timothy J., et al. "Turnover in the Food Stamp Program: A Preliminary Analysis." Report prepared for the U.S. Department of Agriculture, Food and Nutrition Service. Washington, D.C.: Mathematica Policy Research, Inc., July 1984.

Corson, Walter, and Sheena McConnell. "Recent Trends in Food Stamp Program Participation: A Preliminary Report to Congress." Report prepared for the U.S. Department of Agriculture, Food and Nutrition Service. Washington, D.C.: Mathematica Policy Research, Inc., July 1990.

Dickey, David A., and Wayne A. Fuller. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica*, vol. 49, no. 1, July 1981, pp. 1057-1072.

Harvey, Andrew. *The Econometric Analysis of Time Series*. Oxford: Philip Allan, 1981.

Hatanaka, Michio. "An Efficient Two-Step Estimator for the Dynamic Adjustment Model with Autoregressive Error." *Journal of Econometrics*, vol. 2, 1974, pp. 119-220.

Johnston, John. *Econometric Methods*, 3d edition. New York: McGraw-Hill, 1984.

McConnell, Sheena. "The Increase in Food Stamp Program Participation between 1989 and 1990: A Report to Congress." Report prepared for the U.S. Department of Agriculture, Food and Nutrition Service. Washington, D.C.: Mathematica Policy Research, Inc., August 1991.

Park, Soo-Bin. "On the Small-Sample Power of Durbin's *h* Test." *Journal of the American Statistical Association*, vol. 70, 1975, pp. 60-63.

Pindyck, Robert S., and Daniel L. Rubinfeld. *Econometric Models and Economic Forecasts*, 2d edition. New York: McGraw-Hill, 1981.

U.S. Bureau of the Census. *Statistical Abstract of the United States: 1990* (110th edition). Washington, D.C., 1990.

U.S. House of Representatives, Committee on Ways and Means. *The Green Book*. Washington D.C., 1991.

APPENDIX A

COMPARISON OF OMB FORECASTS OF UNEMPLOYMENT RATES WITH ACTUAL UNEMPLOYMENT RATES, 1982 TO 1989

Table of Contents

The quality of the FSP participation and benefits forecasts depends partly on the quality of the forecasts of other variables used to generate participation and benefit forecasts, especially the unemployment rate. Two-year-ahead OMB forecasts of the unemployment rate were examined; this time period corresponds to the forecast period for the FNS budgeting process. Because some of the forecast data were unavailable, the forecasts for fiscal year 1982 are the revised forecasts from February 1981, and the forecasts for fiscal year 1984 are the figures from the mid-session review in 1982.

As shown in Figure A.1, the actual (seasonally adjusted) unemployment rate deviated substantially from the OMB forecasts of the unemployment rate, especially in the early 1980s. The average magnitude of the forecast error was approximately .81, which represents just over 10 percent of the average unemployment rate of 7.4 percent over the period. The budget forecasts do not appear to be highly accurate predictions of the actual unemployment rate during the 1980s, though the forecast accuracy improves in recent years.

In addition, the forecast errors appear to be somewhat systematic. As shown in Figure A.1, the unemployment forecast tended to be below the actual unemployment rate when the actual rate was high, as it was in the early 1980s, and tended to be above the actual unemployment rate when the actual rate was low, as it was in the late 1980s. The positive relationship between the forecast error and the actual unemployment rate is demonstrated in the scatter diagram in Figure A.2. For high values of actual unemployment, the scatter points lie above the 45 degree line, while the scatter points lie below the line for low values of actual unemployment.

Table of Contents

FIGURE A.1

FORECASTED AND ACTUAL UNEMPLOYMENT RATES

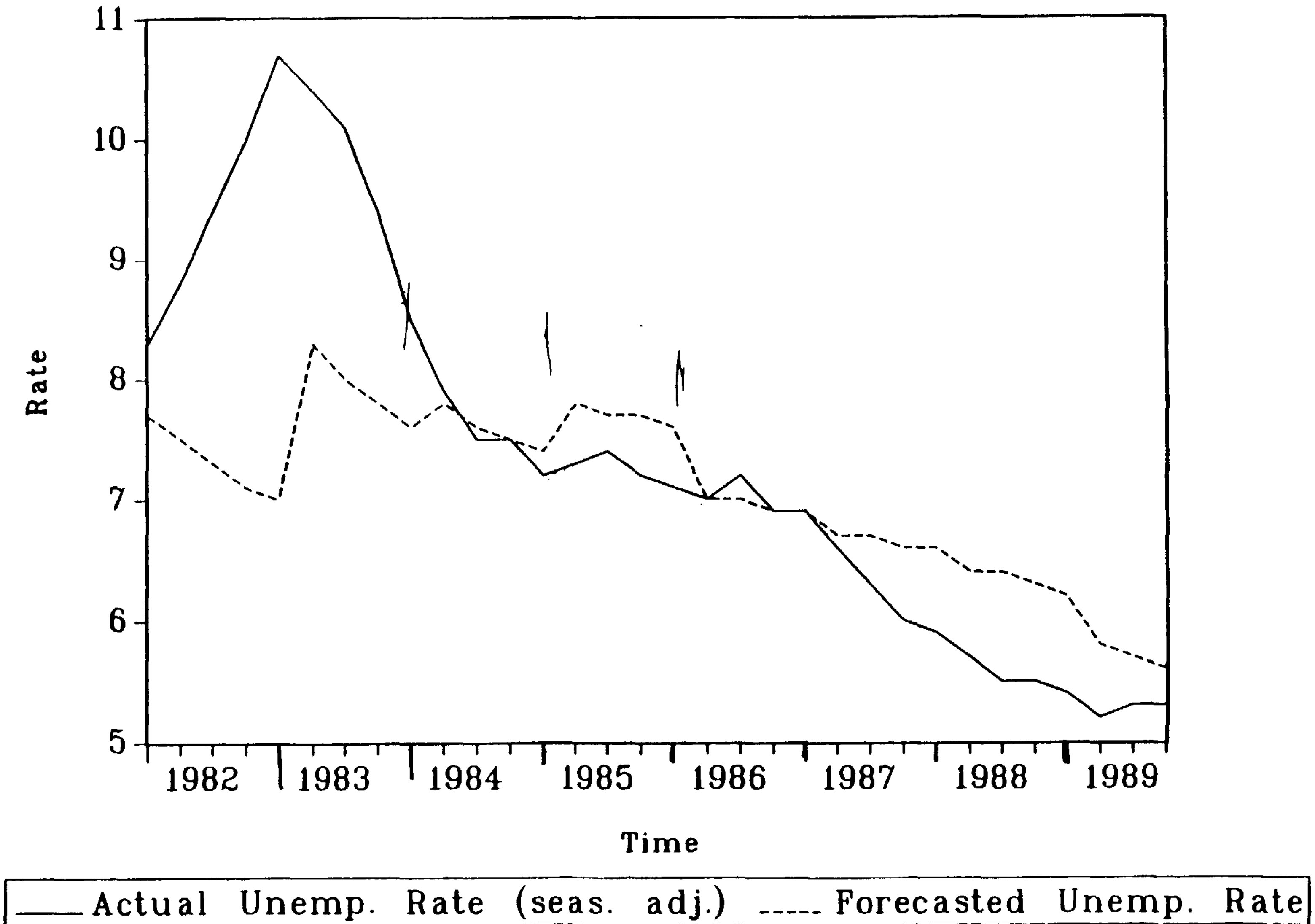
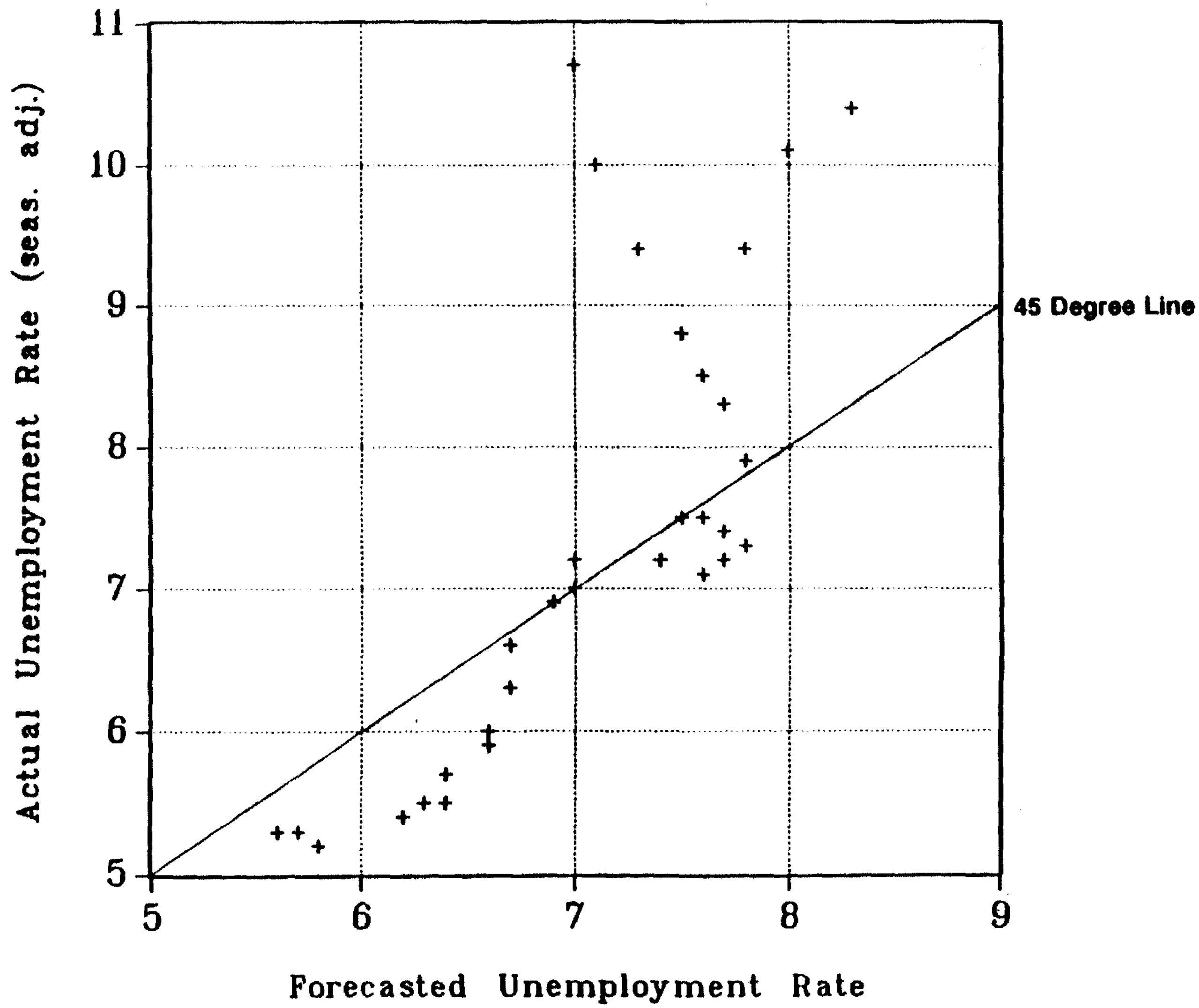


Table of Contents

FIGURE A.2

SCATTER DIAGRAM – FORECASTED AND ACTUAL UNEMPLOYMENT RATE



The effects of unemployment rate forecast errors on FSP participation forecast errors can be illustrated using a simple example. A simple regression model of FSP participation with the unemployment rate as a single explanatory variable can be written:

$$(A.1) \qquad Y_t = a + bX_t + \varepsilon_t$$

If it is assumed that (1) the true coefficients (a and b) are known, (2) the variance of the random error term ε is zero, and (3) an unbiased forecast of the explanatory variable X_t is available, with forecast variance σ_X^2, then the forecast variance of Y is given by:

$$(A.2) \qquad V_y = b^2 \sigma_X^2$$

For example, if we assume that the unemployment rate coefficient b has a value of 775, and the two-year-ahead unemployment rate forecast variance is .16, then according to equation (A.2), the forecast variance for FSP participation would be 360,375.[1] Further calculations indicate that in this case, the unemployment rate forecast error variance gives rise to a 95 percent confidence interval for monthly FSP participation two years ahead of plus or minus 620,000.

In theory, the confidence interval for the FSP participation forecast is larger because the coefficients of the model must be estimated, and because random factors affect participation. The FSP participation confidence interval estimated in Chapter III took into account these factors but did not include the effects of unemployment rate forecast error. If the three sources of variance are assumed to be independent, the 95 percent confidence interval for monthly FSP participation based on the sum of the variances would be roughly plus or minus 1.7 million participants. In other words, the confidence interval is larger by almost 55 percent when the forecast variance of the unemployment rate is included in the model.

[1]The two-year-ahead OMB unemployment rate forecast variance estimate of .16 was calculated over the period 1984 to 1989 for the data shown in Figure A.1.

APPENDIX B

ADDITIONAL REGRESSION MODELS OF FSP PARTICIPATION

Table of Contents

TABLE B 1

REGRESSION MODELS OF FSP PARTICIPATION THAT INCLUDE ADDITIONAL EXPLANATORY VARIABLES REFLECTING GENERAL ECONOMIC CONDITIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Number of unemployed workers (thousands)	0.49 (0.13)	0.23 (0.20)	0.38 (0.11)	0.38 (0.12)	0.35 (0.13)	0.37 (0.12)	0.39 (0.11)	0.41 (0.12)	0.39 (0.11)	0.40 (0.11)	0.35 (0.10)	0.43 (0.11)
Number of unemployed workers lagged one quarter (thousands)	0.39 (0.12)	0.33 (0.18)	0.22 (0.10)	0.23 (0.11)	0.25 (0.10)	0.25 (0.10)	0.22 (0.10)	0.21 (0.10)	0.21 (0.10)	0.19 (0.10)	0.22 (0.09)	0.18 (0.10)
Number of workers exhausting UI benefits (thousands)	5.7 (2.8)		--	--	--	--	--	--	--	--	--	--
Number of first UI payments (thousands)		0.25 (0.32)	--	--	--	--	--	--	--	--	--	--
Consumer price index	--	--	5.8 (5.1)	--	--	--	--	--	--	--	--	--
Per capita disposable income (thousands of dollars)	--	--	--	81.9 (87.3)	--	--	--	--	--	--	--	--
Number of workers employed in retail trade (thousands)	--	--	--	--	0.01 (0.12)	--	--	--	--	--	--	--
Number of workers employed in non-agricultural jobs (thousands)	--	--	--	--	--	0.01 (0.03)	--	--	--	--	--	--
Hourly wages	--	--	--	--	--	--	229.2 (208.2)	--	--	--	--	--
Weekly wages	--	--	--	--	--	--	--	8.39 (6.52)	--	--	--	--
Hourly wages in the personal services industry	--	--	--	--	--	--	--	--	401.0 (350.5)	--	--	--
Hourly wages in the retail trade industry	--	--	--	--	--	--	--	--	--	700.9 (482.2)	--	--
Weekly earnings in the personal services industry	--	--	--	--	--	--	--	--	--	--	68.5 (21.2)	--
Weekly earnings in the retail trade industry	--	--	--	--	--	--	--	--	--	--	--	28.2 (14.5)
Elimination of purchase requirement (EPR)	5,205.3 (458.4)	5,012.5 (475.9)	4,765.7 (500.3)	4,943.0 (478.3)	5,089.3 (485.6)	4,987.3 (482.4)	4,853.0 (489.5)	4,809.9 (493.2)	4,768.2 (504.2)	4,872.8 (506.1)	4,342.1 (494.6)	4,597.5 (499.8)
OBRA 1981 (REC2)	-887.7 (243.9)	-847.3 (244.3)	-854.8 (281.7)	-788.5 (291.1)	-685.5 (295.9)	-735.6 (284.4)	-854.3 (285.2)	-884.1 (283.2)	-845.1 (276.5)	-882.1 (262.2)	-728.5 (248.1)	-891.4 (252.5)
1st quarter dummy	220.7 (71.3)	169.6 (88.8)	229.9 (74.7)	220.7 (76.7)	206.0 (74.6)	205.7 (73.9)	218.5 (73.3)	232.9 (74.3)	232.5 (74.4)	226.8 (70.8)	147.3 (89.0)	304.1 (84.0)

Table of Contents

TABLE B.1 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
2nd quarter dummy	580.4	382.7	499.6	490.2	510.2	512.7	480.2	503.1	479.7	431.7	611.0	583.7
	(117.2)	(180.5)	(113.8)	(118.4)	(118.0)	(116.9)	(118.7)	(113.4)	(114.5)	(118.8)	(107.3)	(117.2)
3rd quarter dummy	287.8	241.5	268.9	258.8	241.1	237.7	285.4	282.1	272.2	273.8	200.3	370.8
	(90.9)	(90.9)	(90.7)	(92.9)	(93.3)	(92.1)	(91.2)	(94.5)	(90.1)	(87.8)	(85.3)	(102.1)
Constant	9,223.1	9,983.2	8,672.0	9,533.6	9,872.2	8,675.8	8,501.1	7,971.1	8,267.4	6,627.6	-6,843.1	5,793.6
	(758.5)	(717.2)	(1,689.7)	(1,218.4)	(2,312.1)	(3,514.7)	(1,773.4)	(2,012.5)	(2,114.1)	(2,835.9)	(7,329.1)	(2,641.0)
ρ	0.77	0.79	0.81	0.79	0.78	0.79	0.81	0.82	0.83	0.84	0.98	0.84
	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.06)	(0.02)	(0.06)
$\bar{R}^2$	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Standard error of the regression	223.1	232.6	229.3	232.4	234.2	233.8	230.1	228.5	228.9	224.0	215.4	220.6
Durbin-Watson statistic	1.5	1.3	1.5	1.4	1.4	1.4	1.5	1.5	1.5	1.5	1.3	1.6
Sample period	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4	FY1977.1-FY1989.4

NOTES: Standard errors are in parentheses. The dependent variable is the monthly average of FSP participants (in thousands) during a fiscal year quarter.

Table of Contents

TABLE B.2

REGRESSION MODELS OF FSP PARTICIPATION THAT INCLUDE ADDITIONAL DEMOGRAPHIC CHARACTERISTICS, FSP PROGRAM PARAMETERS, AND LEGISLATIVE CHANGE VARIABLES

	(1)	(2)	(3)	(4)	(5)	(6)
Number of unemployed workers (thousands)	0.37 (0.11)	0.34 (0.13)	0.30 (0.13)	0.35 (0.11)	0.34 (0.12)	0.36 (0.11)
Number of unemployed workers lagged one quarter (thousands)	0.21 (0.10	0.25 (0.11)	0.16 (0.12)	0.25 (0.10)	0.25 (0.10)	0.25 (0.10)
Number of female-headed households below poverty with children under 18 (thousands)	846.6 (640.7)	--	--	--	--	--
Number of SSI recipients (thousands)	--	-0.20 (1.00)	--	--	--	--
Maximum allotment for a family of four	--	--	5.0 (10.7)	--	--	--
Standard deduction	--	--	32.1 (43.0)	--	--	--
Excess shelter and child care deductions	--	--	-24.6 (21.5)	--	--	--
Food Stamp Act 1986 dummy	--	--	--	24.2 (238.7)	--	--
IRCA dummy	--	--	--	--	-76.4 (241.8)	--
Hunger Prevention Act dummy	--	--	--	--	--	193.9 (232.2)
Elimination of purchase requirement (EPR)	4,927.9 (472.6)	5,088.5 (467.5)	237.9 (3481.3)	5,083.1 (461.6)	5,117.8 (455.7)	5,052.5 (451.1)
OBRA 1981 (REC2)	-855.1 (268.9)	-657.9 (254.2)	-971.8 (268.9)	-678.5 (246.0)	-655.3 (247.4)	-692.8 (234.4)
1st quarter dummy	223.7 (72.8)	201.1 (79.2)	252.3 (77.6)	205.3 (74.6)	200.3 (77.5)	197.8 (74.4)
2nd quarter dummy	497.4 (113.4)	515.0 (120.3)	534.2 (141.6)	507.6 (119.2)	512.7 (117.6)	500.1 (117.0)
3rd quarter dummy	273.0 (90.6)	242.7 (93.3)	339.9 (100.8)	239.9 (91.7)	232.2 (93.7)	234.8 (91.1)

TABLE B.2 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	7,937.7 (1,991.7)	11,066.6 (4,979.0)	15,565.1 (4,290.4)	10,095.4 (717.4)	10,145.0 (707.0)	10,036.2 (660.9)
ρ	0.81 (0.09)	0.77 (0.10)	0.78 (0.08)	0.78 (0.10)	0.78 (0.09)	0.77 (0.09)
$\bar{R}^2$	0.99	0.99	0.97	0.99	0.99	0.92
Standard error of the regression	227.9	234.1	221.3	234.2	233.9	232.2
Durbin-Watson statistic	1.5	1.4	1.5	1.4	1.4	1.4
Sample period	FY1977.1- FY1989.4	FY1977.1- FY1989.4	FY1977.1- FY1989.4	FY1977.1- FY1989.4	FY1977.1- FY1989.4	FY1977.1- FY1989.4

NOTES: Standard errors are in parentheses. The dependent variable is the monthly average of FSP participants (in thousands) during a fiscal year quarter.

Table of Contents

APPENDIX C

TIME SERIES MODELS OF FOOD STAMP PROGRAM PARTICIPATION

This appendix analyzes the pattern of FSP participation over time based on a time-series model. The pattern of participation is first investigated to ensure that participation has the properties that are consistent with the assumptions that underlie a time-series model. A time-series model of participation is then specified as a function of previous values of participation. Two models of participation are estimated, and out-of-sample forecasts of participation are calculated from these models. The results indicate that time-series models generate less precise forecasts than do regression models.

A. STATIONARITY TESTS OF THE FSP PARTICIPATION SERIES

A time-series model of FSP participation is based on the assumption that participation is generated by a *stationary stochastic process*--that the value of participation at each point in time represents a random draw from a probability distribution whose parameters do not vary over time. If the stochastic process underlying participation is not stationary, it is usually possible to transform the series into one that is stationary, by taking first or second differences.

One method for determining whether a series is stationary is to examine the sample autocorrelation function of the series. The autocorrelation function measures the correlation between neighboring data points in a series. More specifically, if we denote the value of the variable in period t as P_t, then the autocorrelation function gives for each lag k the correlation between P_t and P_{t+k} (denoted as r_k). If a series is stationary, the autocorrelation function will approach zero as k, the number of lags, becomes large. The autocorrelation function will approach zero very slowly or not at all if the series is nonstationary.

A plot of the sample autocorrelations presented in Figure C.1 suggests that participation may be stationary, but the evidence is not conclusive. The positive autocorrelations in the early lags drop off quickly toward zero, but the autocorrelations are negative at fairly long lags.

FIGURE C.1

SAMPLE AUTOCORRELATIONS FOR LEVELS OF FOOD STAMP PARTICIPATION

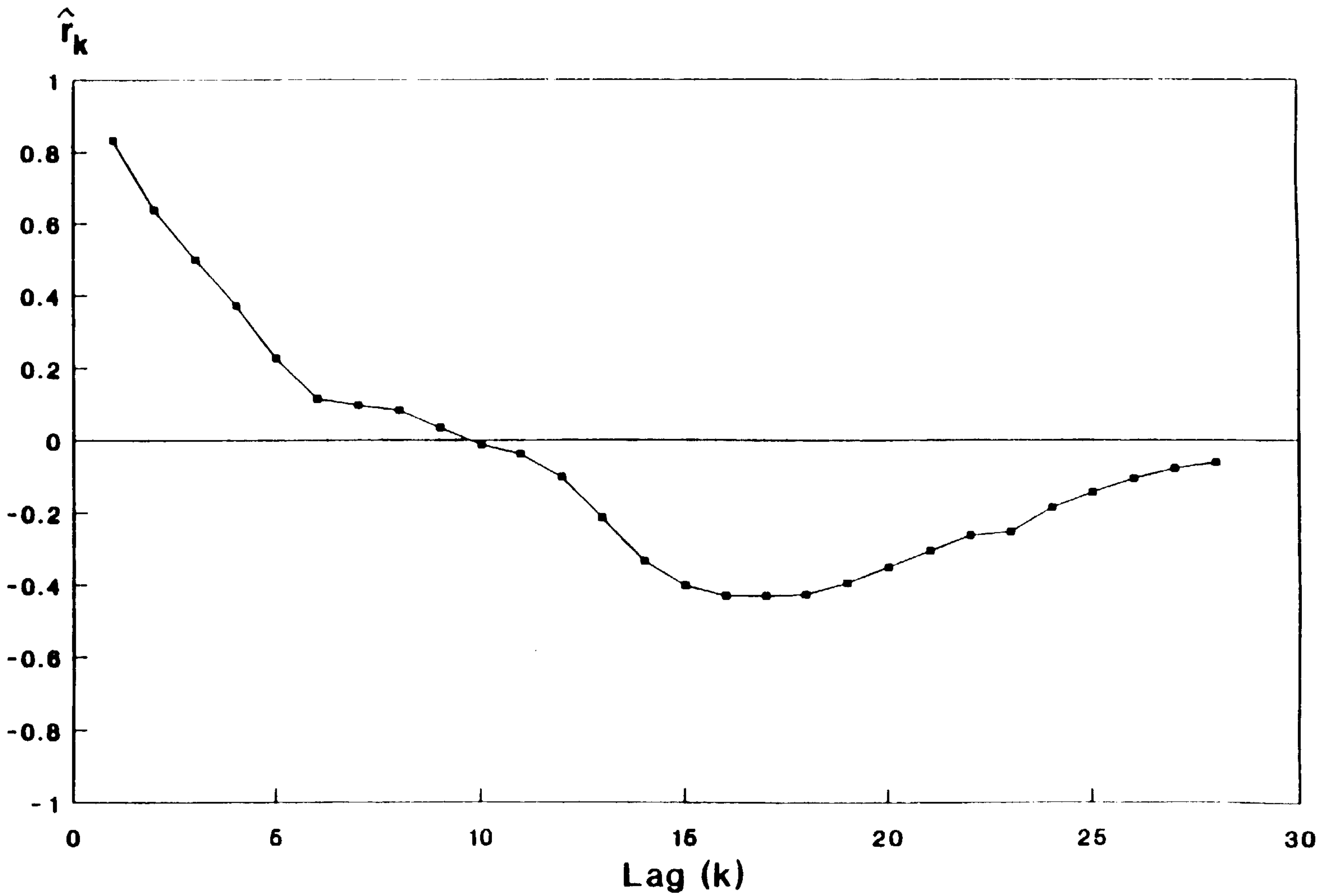


Table of Contents

Given the uncertainty about the stationarity of the participation series based on the sample autocorrelation function, Dickey-Fuller tests were conducted to determine whether the participation series is stationary (Dickey and Fuller, 1981). The Dickey-Fuller stationarity test is based on a regression of the form:

$$(1) \qquad P_t = a_0 + a_1 P_{t-1} + a_2(P_{t-1} - P_{t-2}) + \ldots + a_j(P_{t-j+1} - P_{t-j}) + u_t,$$

where P_t is the level of participation in time t, u_t is a random disturbance, and the a's are the parameters to be estimated. The Dickey-Fuller test is a test of the null hypothesis that a_1 is equal to one versus the hypothesis that a_1 is less than one. If the null hypothesis is rejected, then the conclusion is that a_1 is less than one and the series is stationary. If the null hypothesis that the estimate of a_1 is equal to one cannot be rejected, then the conclusion is that the series may be nonstationary.

Subtracting P_{t-1} from both sides of equation (1) yields:

$$(2) \qquad P_t - P_{t-1} = a_0 + (a_1 - 1)P_{t-1} + a_2(P_{t-1} - P_{t-2}) + \ldots + a_j(P_{t-j+1} - P_{t-j}) + u_t.$$

In this case, the test for stationarity is a test of the null hypothesis that the coefficient on the lagged participation level, $a_1 - 1$, is equal to zero. If the null hypothesis is rejected and the coefficient on lagged participation is negative, then we conclude that the series is stationary. The test statistic to be used for this version of the Dickey-Fuller test is the t-statistic for the coefficient on lagged participation--the coefficient divided by its standard error. The critical value for the t-statistic corresponding to the 95 percent level of confidence with 50 observations is -2.93.

Table C.1 contains the regression estimates for the Dickey-Fuller tests. Five different specifications were tested, using up to four lags of the first difference of participation on the right-hand-side of the equation. For all of the five specifications tested, the test statistic exceeded the

critical value, meaning that the null hypothesis of nonstationarity can be rejected at the 95 percent level of confidence in each case.

B. TIME-SERIES MODELS OF FSP PARTICIPATION

There are two basic types of time-series models. The first model is termed an *autoregressive process*, in which the current observation of a variable is generated by a weighted sum of past observations and a random disturbance in the current period. An autoregressive process of order p is written as:

$$(3) \qquad P_t = b_0 + b_1 P_{t-1} + b_2 P_{t-2} + \ldots + b_p P_{t-p} + e_t,$$

and is denoted as AR(p). The second basic model is termed a *moving average process*, in which the current observation of the variable is generated by a weighted sum of random disturbances from previous periods and a random disturbance in the current period. A moving average process of order q is written as:

$$(4) \qquad P_t = c_0 + c_1 e_{t-1} + c_2 e_{t-2} + \ldots + c_q e_{t-q} + e_t,$$

and is denoted as MA(q).

Stationary processes may be combinations of autoregressive and moving average elements. Combined processes are called *mixed autoregressive moving-average processes*, and are denoted ARMA(p,q). The objective of our time-series analysis was to determine combinations of these two types of processes that best characterize the FSP participation series.

Based on the plot of the sample autocorrelation function presented in Figure C.1 and on a preliminary investigation of alternative specifications, two specifications were used to describe the FSP participation series. These two equations, whose estimates are presented in Table C.2, specify that participation follows an ARMA(5,1) procedure, with controls for the seasonality of the series,

Table of Contents

TABLE C.2

TIME-SERIES MODELS OF FSP PARTICIPATION
(Standard Errors in Parentheses)

	Model	
Variable	(1)	(2)
Constant	19,500.0 (385)	20,622.6 (1,863)
P_{t-1}	0.84 (0.05)	1.57 (0.17)
P_{t-2}	--	-0.70 (0.16)
P_{t-4}	0.70 (0.10)	
P_{t-5}	-0.68 (0.08)	
e_{t-1}	0.69 (0.16)	-0.09 (0.23)
Seasonality Coefficient γ		0.78 (0.08)
Standard Error of the Regression	350.8	359
Durbin-Watson Statistic	1.94	1.91
R^2	0.97	0.97
Sample Period	1977.1-1989.4	1977.1-1989.4
Out-of-Sample Forecasts		
Mean Forecast Error		
1988	-1,267.3	-1,270.3
1989	248.0	-527.0
Average	-509.5	-898.5
Root Mean Square Error		
1988	1,286.5	1,301.2
1989	266.0	535.1
Average	776.3	918.1

or an ARMA(5,1) procedure, with alternative controls for seasonality. Seasonality in the first equation is modeled by including the fourth and fifth lags of the autoregressive process. The observation in the current quarter is related to the value four quarters earlier, due to seasonality. The fifth lag is also included in the equation due to the combination of seasonality and the AR(1) component of the series. In the second equation, seasonality is modeled alternatively by specifying the equation as:

$$(5) \qquad P_t = b_0 + b_1P_{t-1} + b_2P_{t-2} + c_1e_{t-1} + \gamma(P_{t-4}) + \gamma(b_1P_{t-5}) + \gamma(b_2P_{t-6}) + e_{t,}$$

where the fourth, fifth, and sixth lags of participation are included in the equation due to the seasonality of the series, which is measured by the autoregressive seasonal adjustment parameter.

The sample autocorrelation function for the residuals of the estimated equations provides a diagnostic check of the validity of the time-series models of FSP participation. If a model is specified correctly, the residuals should be uncorrelated with each other. The sample autocorrelation functions for the two models are shown in Figure C.2. The absence of any large spikes in the residual autocorrelations indicates that the estimated residuals are uncorrelated with each other in both models.

The hypothesis that the residual autocorrelations back to some maximum lag k were equal to zero was tested with the Q-statistic. The Q-statistic is equal to:

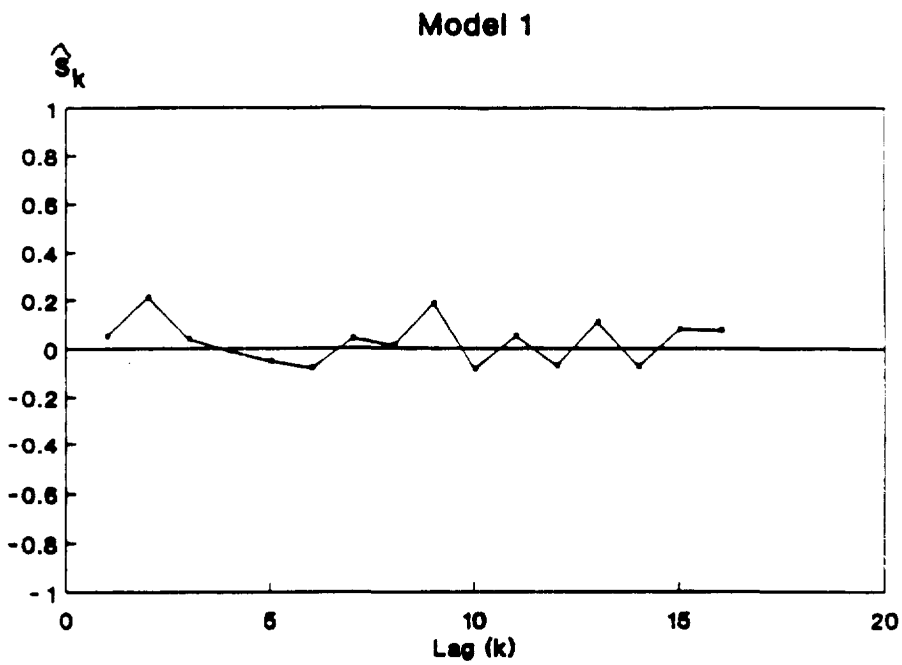
$$(6) \qquad Q = T\sum_{k=1}^{K} \hat{s}_k^2$$

where T is the number of observations in the time series, and s_k is the estimate of the kth residual autocorrelation. A large Q-statistic arises when the residual autocorrelations are large. The Q-statistic is distributed χ^2 (K-p-q) degrees of freedom. The Q-statistics for the models in Table C.2 are equal to 5.9 and 9.6. Both of these statistics are small and fall well below the 90 percent

Table of Contents

FIGURE C.2

SAMPLE AUTOCORRELATIONS OF RESIDUALS



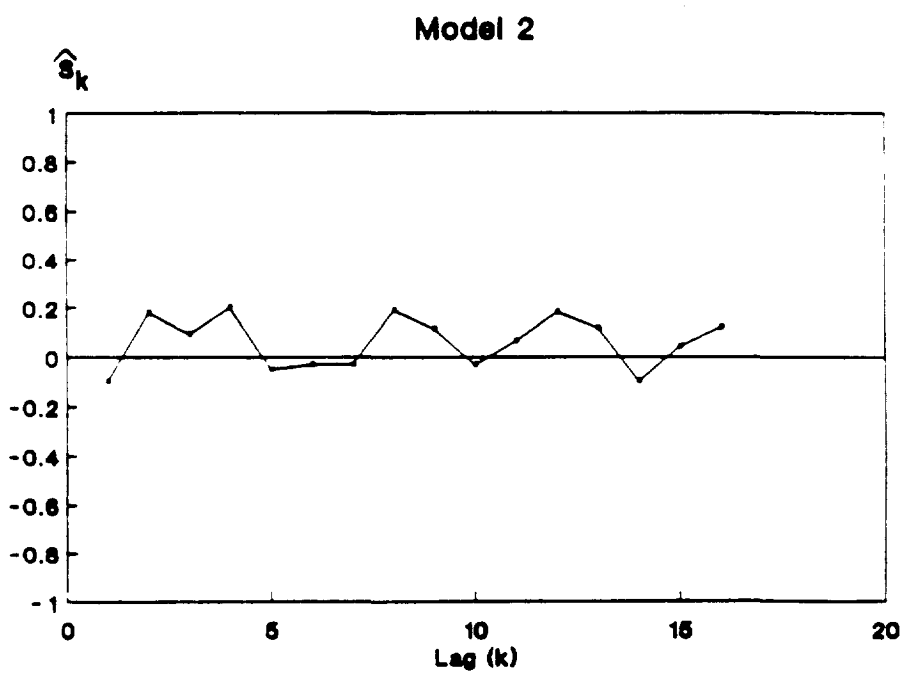


Table of Contents

critical value of the χ^2 distribution, which is equal to 21.0 with 14 degrees of freedom. The Q-statistics thus provide evidence that the models are correctly specified.

As a final check of the models, the sample autocorrelation functions for the predicted participation series for each of the models were compared with the sample autocorrelation function for the original series. The plots of the sample autocorrelation functions for the two predicted participation series and the original series are shown in Figure C.3. Both models generated predicted participation series whose sample autocorrelations look similar to the autocorrelations for the original series, providing further evidence that the stochastic process is specified correctly.

Two reasonable time-series representations of the stochastic process that generate the level of FSP participation have been found. However, as evident from the out-of-sample forecast statistics shown in Table C.2, neither model is able to forecast accurately out-of-sample. The *rmse* was 776 for model 1 and 918 for model 2, which are considerably larger than the *rmse* values obtained for the regression models examined in Chapter III. The implication is that it is feasible to use time-series methods to forecast participation, but using regression methods will generate more precise forecasts.

D. SUMMARY

Time-series analysis is the major alternative to regression analysis for forecasting the level of FSP participation over time. The result of the time-series analysis can be summarized briefly. Participation appears to be a stationary series--the stochastic process that generates the series is not variable over time. The time-series pattern of participation can be explained with a simple autoregressive moving-average model that controls for seasonality. However, the out-of-sample forecasting abilities of the time-series models are limited.

Table of Contents

FIGURE C.3

SAMPLE AUTOCORRELATIONS ACTUAL AND PREDICTED PARTICIPATION

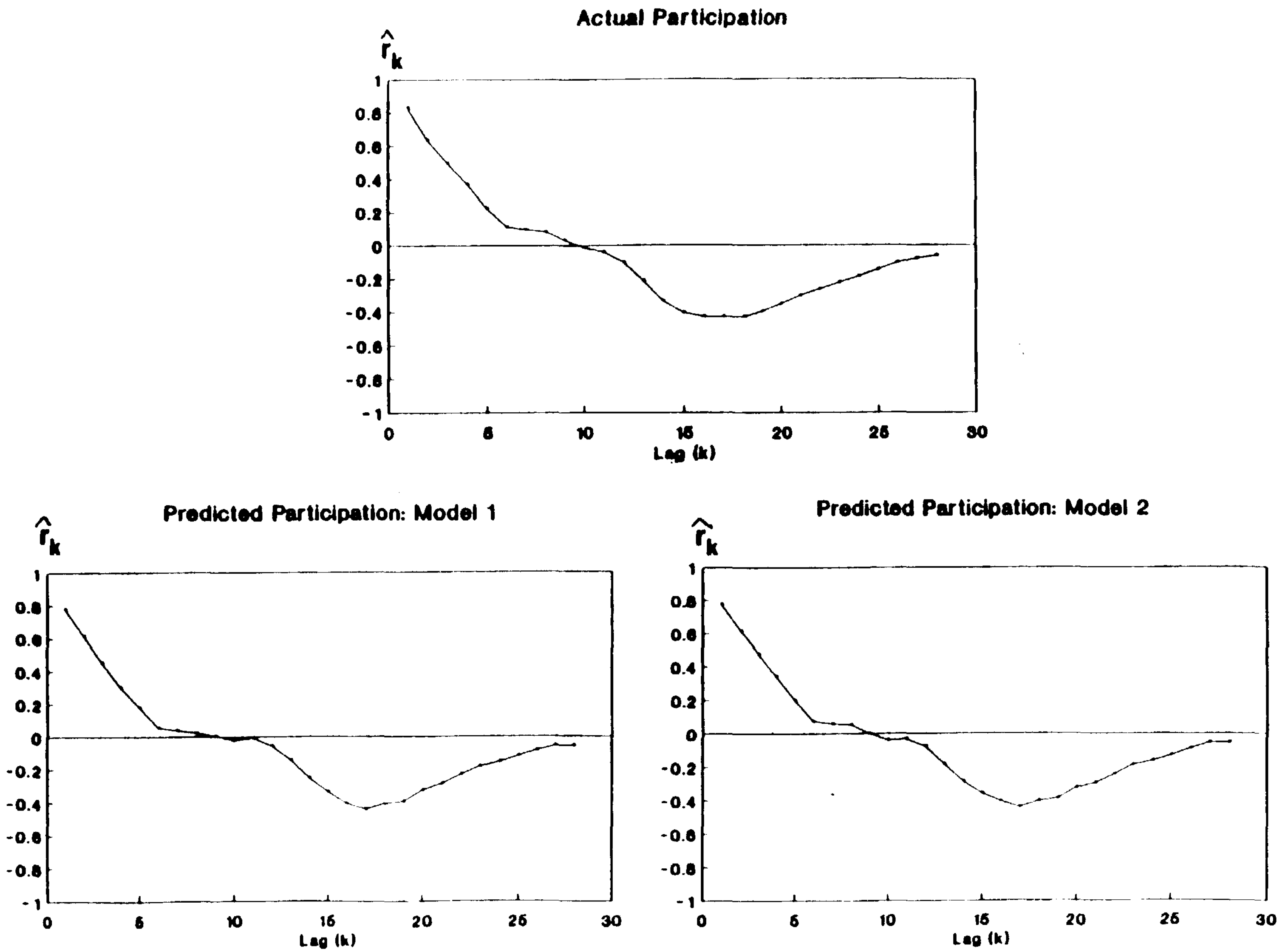


Table of Contents

APPENDIX D

DATA USED IN MODEL ESTIMATION

Table of Contents

QUARTERLY MODELS

AVNETELG	Net monthly household income eligibility limit for a family of four
BONUS	Average monthly benefit received by FSP participants during a fiscal year quarter
CPI	Consumer price index
CPIFAH	CPI for food at home, not seasonally adjusted
DISPINC	Per-capita disposable income (in thousands of dollars)
EPR	Elimination of purchase requirement dummy (equals 0 through FY 1979.1; fluctuates between 0 and 1, FY 1979.2 - FY 1980.1; equals 1 thereafter)
ESCC	Excess shelter and child care deductions
FEMHEAD	Number of female-headed households with children under 18 (in thousands)
FEMPOV	Number of female-headed households below poverty level with children under 18 (in thousands)
FIRSTPAY	Number of first UI payments (in thousands)
FOODACT	Food Stamp Act of 1986 dummy (equals 0 through FY 1986.4; equals 1 thereafter)
HACT	Fiscal year 1989.3 dummy (Hunger Prevention Act) (equals 0 through FY 1989.2; equals 1 thereafter)
HRLYLND	Hourly wages in the personal services industry
HOURLY	Hourly wages
HRLYRET	Hourly wages in the retail trade industry
IRCA	IRCA dummy (equals 0 through FY 1987.3; equals 1 thereafter)
NAEMP	Number of workers employed in nonagricultural jobs (in thousands)
NEWAFDC	Number of AFDC recipients (in thousands)
NEWPSEM	Number of workers employed in the personal services industry (in thousands)
NEWREC	OBRA 1981 dummy (referred to as REC2 in the text) (equals 0 through FY 1981.4; equals 1 thereafter)
NEWUIEX	Number of workers exhausting UI benefits (in thousands)

Table of Contents

NUSPART	Monthly average of FSP participants (in thousands) during a fiscal year quarter (not seasonally adjusted)
PADUM	Fiscal year 1988 dummy (equals 0 through FY 1987.4; equals 1 thereafter)
QRT1	1st quarter dummy (equals 1 if fiscal year quarter is 1; equals 0 otherwise)
QRT2	2nd quarter dummy (equals 1 if fiscal year quarter is 2; equals 0 otherwise)
QRT3	3rd quarter dummy (equals 1 if fiscal year quarter is 3; equals 0 otherwise)
REC	OBRA 1981 dummy (equals 0 through FY 1981.4; is between 0 and 1, FY 1982.1 - FY 1982.2; equals 1, FY 1982.3 - FY 1983.4; fluctuates between 0 and 1, FY 1984.1 - FY 1988.2; equals 0 thereafter)
RETEMP	Number of workers employed in retail trade (in thousands)
SAUSPART	Seasonally adjusted monthly average of FSP participants (in thousands) during a fiscal year quarter
SCPIFAH	CPI for food at home, seasonally adjusted
SD	Standard deduction
SSI	Number of SSI recipients (in thousands)
TFP	Maximum allotment for a family of four
TOTBEN	Monthly average of total benefits (in millions of dollars) received by FSP participants during a fiscal year quarter
UAJULEV	Number of unemployed workers (in thousands)
URADJ	Unemployment rate, seasonally adjusted
URATE	Unemployment rate, not seasonally adjusted
USPART	Monthly average of FSP participants (in thousands) during a fiscal year quarter (not seasonally adjusted)
WEEKLND	Weekly earnings in the personal services industry
WEEKLY	Weekly wages
WEEKRET	Weekly earnings in the retail trade industry

MONTHLY MODELS

MCPIFAH	CPI for food at home
MEPR	Elimination of purchase requirement dummy (equals 0 through Dec 1978; fluctuates between 0 and 1, Jan – Dec 1979; equals 1 thereafter)
MFSPART	Monthly average of FSP participants (in thousands)
MREC	OBRA 1981 dummy (referred to as REC2 in the text) (equals 0 through Sep 1981; equals 1 thereafter)
MURATE	Unemployment rate