Food Stamp Program Certification Costs and Errors, 1989-2005

Final Report

Contractor and Cooperator Report No. 45
June 2008

By Christopher Logan, Ryan Kling, and William Rhodes, Abt Associates Inc.

Abstract

Preventing and detecting certification errors in the Food Stamp Program (FSP) is a major policy concern. In 2005, the cost of overpayments was $1.29 billion, about 4.5 percent of the $28.6 billion in benefits issued. This report examines the State-level relationships between FSP certification error rates and certification expenditures, program policies, caseload characteristics, and economic conditions. The results show that, during the study period of 1989-2005, a 10-percent increase in certification "effort"—about $35 per participating household—would reduce an index of certification errors by 2 percent (0.3 percentage points out of a mean of 15.1 percent). The effect of certification effort was significantly smaller between 1997 and 2002, when States were implementing welfare reform. Key simplification policies authorized by the 2002 Farm Bill were estimated to jointly reduce the error index by 4.4 percentage points.

Data files and documentation for this report are available upon request from Kenneth Hanson at khanson@ers.usda.gov.

This study was conducted by Abt Associates Inc. under a cooperative research contract with USDA's Economic Research Service (ERS) Food and Nutrition Assistance Research Program (FANRP): contract number 59-5000-6-0078 (ERS project representative: Kenneth Hanson). The views expressed are those of the authors and not necessarily those of ERS or USDA.

Acknowledgments

The authors wish to thank Chris Flygare of Abt Associates Inc. for programming support, and the production staff of Eileen Fahey, Katheleen Linton, and Jan Nicholson. Jacob Alex Klerman provided thoughtful advice as the internal Project Quality Advisor. Joseph Sabia, formerly of Abt and presently of the University of Georgia, played an important role in developing the methods for the previous study and co-authored portions of the previous report that were adapted for use in this report. Kenneth Hanson delivered essential support and helpful comments and advice as the Department Officer's Representative. We benefited from the advice and feedback of Margaret Andrews of ERS, who also assisted us in obtaining the Food Stamp Rules database. We greatly appreciate the assistance of Kristen Dowling Hyatt and Bob Dalrymple of FNS in providing the administrative cost data. We thank Kristen Dowling Hyatt, Jenny Laster Genser, Dan Wilusz, Charles Okal, Art Foley, Linda App, and Lael Lubing of FNS, and John Kirlin of ERS, for their comments on the draft final report.

Contents

Acknowledgments .... **ii**

Executive Summary .... **v**
- Background v
- Research Methods v
- National Summary of FSP Administrative Costs vi
- Trends in FSP Caseloads, Certification-Related Effort, and Error Rates vii
- Results of Modeling Error Rates ix
- Explanation of Trends in Error Rates ix
- Causes of Changes in the Marginal Effect of Effort on Error xii
- Limitations of the Study xii

Chapter One: Introduction and Study Highlights **1**
- Context 1
- Study Objectives and Research Questions 3
- Contribution to the Literature 3
- Organization of the Report 4

Chapter Two: FSP Certification Errors and Quality Control **5**
- Types and Sources of FSP Certification Errors 5
- FSP Quality Control 6
- Use of QC Data for this Report 8
- Trends in Error Rates and Caseloads 10

Chapter Three: FSP Administrative Costs and Certification Effort **13**
- Administrative Activities in the Food Stamp Program 13
- National Summary of FSP Administrative Costs 15
- Variation Across States in FSP Administrative Costs per Household 22
- Welfare Worker Pay and FSP Certification Effort 22

Chapter Four: Literature and Conceptual Model of FSP Error Rates **27**
- Studies of Error Rates and Error Reduction Strategies 27
- Conceptual Model of Effort and Error Rates 31
- Hypotheses About the Relationships of Certification Effort and Policies to the FSP Error Index 33

Chapter Five: Model Estimation and Results **39**
- Data 39
- Estimation of Models of the Food Stamp Error Index 46
- Summary of Results 58

Chapter Six: Conclusions and Discussion **65**
- Trends in FSP Caseloads, Certification-Related Effort, and Error Rates 65

Explanation of Trends in Error Rates 66
Causes of Changes in the Marginal Effect of Effort on Error 67
Limitations of the Study 67
Suggestions for Future Research 68

Citations 71

Appendix A: Sources and Methods for Analysis of FSP Administrative Costs A-1

Appendix B: Additional Explanation of Multivariate Analysis B-1
Estimation of Variable Contributions to the Marginal Effect of Effort on Error B-3

Appendix C: Alternate Models of Error Rates C-1
Alternative Measures of Error C-1
Alternative Specifications of Effort Variables C-4
Alternative Models with Additional Time-Varying Relationships C-6
Alternative Models using the Base Specification C-9

Appendix D: Explanation of Alternative Models D-1
Prais-Winsten FGLS Model D-1
Partial Adjustment Model D-2
Arellano-Bond Model D-3

Executive Summary

This report presents a study of the relationship of State certification error rates in the Food Stamp Program (FSP) to State certification expenditures, program policies, caseload characteristics, and economic conditions, in the period from 1989 to 2005. During this period, the number of FSP participants rose from 18.8 million (in 1989) to 27.4 million (in 1994), fell to 17.2 million (in 2000), and rose again to 25.7 million (in 2005). The study analyzed recent trends in FSP error rates and expenditures, and then examined the causes of recent declines in error rates through multivariate analysis of State panel data. From 1993 to 2005, the proportion of FSP benefits representing overpayments fell from 8.27 percent to 4.53 percent. The study focused on the relationship of error rates to State expenditures on certification and recent options for program simplification.

Background

The FSP provides assistance to low-income Americans for the purchase of food at authorized stores, farmers' markets, and other locations. The Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA) administers the program in partnership with the 50 States and the District of Columbia.

Preventing and detecting benefit issuance errors is a major concern of FNS and the States. The total cost of FSP benefits in 2005 was $28.57 billion, and the cost of overpayments was $1.29 billion. FNS pays the entire cost of the benefits and exercises considerable oversight to assure the integrity of certification (the process of determining eligibility and benefits). FNS also pays about half of the administrative cost at the State and local levels, and thus has a substantial interest in the prudent expenditure of these funds.

Two laws made major changes to the FSP during the study period. One was the Personal Responsibility and Work Opportunity Act (PRWORA), enacted in 1996. This act created the Temporary Aid for Needy Families (TANF) program to replace the Aid to Families with Dependent Children (AFDC) program, which provided cash assistance to 37 percent of FSP households in 1996. TANF implementation led to major changes in the operation of FSP/TANF agencies. The second law was the Farm Security and Rural Investment Act (the 2002 Farm Bill), which authorized several options to simplify FSP certification, including simplified reporting, transitional food stamp benefits for households leaving TANF, and simplified definition of income to conform with TANF rules. Both simplified reporting and transitional benefits allow participants and agencies to ignore changes in household income and expenses during the certification period, within certain limits. The Farm Bill also changed FSP rules so that State liabilities for excessive error rates were based on two years' performance, not just one.

Research Methods

This study had two main objectives: to describe FSP administrative costs and error rates in 1989-2005, and to explain the relationship of error rates to administrative costs, State policy choices, and other factors. In particular, two of the authors had previously estimated the relationship of certification-related effort expended by local and State agencies to error rates for the 1989-2001 period. The earlier study ended during a time of major change in the FSP. This study extended the time period through 2005 and explored the stability of the estimated relationship after that period of change. Another focus was the relationship of certification error rates to new State FSP policy options for program simplification that were widely implemented in 2001-2005.

The researchers compiled and analyzed FNS data on FSP administrative costs, certification errors, and case characteristics. Additional sources included a recently-compiled Urban Institute database of State food stamp policy options used in 1996-2004, prior studies of FSP error rates, and State-level statistics from the Current Population Survey and the Annual Survey of Government Employment.

The study used a measure of the quantity of administrative resources expended on certification-related activities, controlling for the effect of wage variation on administrative costs. This measure, called "effort", was computed by dividing the certification-related cost per FSP case by the average public welfare worker wage. The measure of effort was, therefore, the ratio of full-time equivalent staff to cases with a multiplier for nonpersonnel costs.[1] The underlying hypothesis was that when workers spend more time per case, they make fewer errors. Ideally, the measure of effort would be based on a standardized unit, such as a full-time worker with specified education and experience, but the necessary data were not available.

An error index was computed that combined the percentages of FSP ***cases*** with overpayments (including ineligible households receiving benefits), underpayments, and improper denial or termination of benefits (negative action errors). The inclusion of negative action errors made the error index more inclusive than the official payment error rate, which is the percent of food stamp ***benefits*** that are overpayments or underpayments.

National Summary of FSP Administrative Costs

Over the period from 1989 to 2005, the average real annual cost of FSP administration for the U.S was $4.30 billion, including Federal and non-Federal funds. The total real annual cost of FSP administration for the U.S rose from $3.18 billion to $4.99 billion, an increase of 57%. (All cost estimates in this report are in 2005 dollars, so inflation was not a factor in this increase.) At the national level, the total FSP administrative cost increased every year until 2002 (even after adjusting for inflation). The total administrative cost leveled off in 2002-2004, and then fell in 2005.

Figure ES-1 shows the percentage distribution of the national total FSP administrative cost for the period (summed over all 17 years) among FSP administrative activities. Certification was by far the largest category, representing three-fifths (59.1%) of the total. This percentage was essentially the same over the 17 years in the study. The average total annual cost for certification was $2.54 billion. The next largest categories were unspecified other (8.5%), issuance (7.2%), employment and training (E&T) services (7.1%), and automated data processing (ADP) operations (5.9%). ("Unspecified other" is the category for costs not reported elsewhere, and so the scope of these costs may have varied from State to State or over time.) The smallest categories were miscellaneous, fraud control, food stamp nutrition education (FSNE), and ADP development.

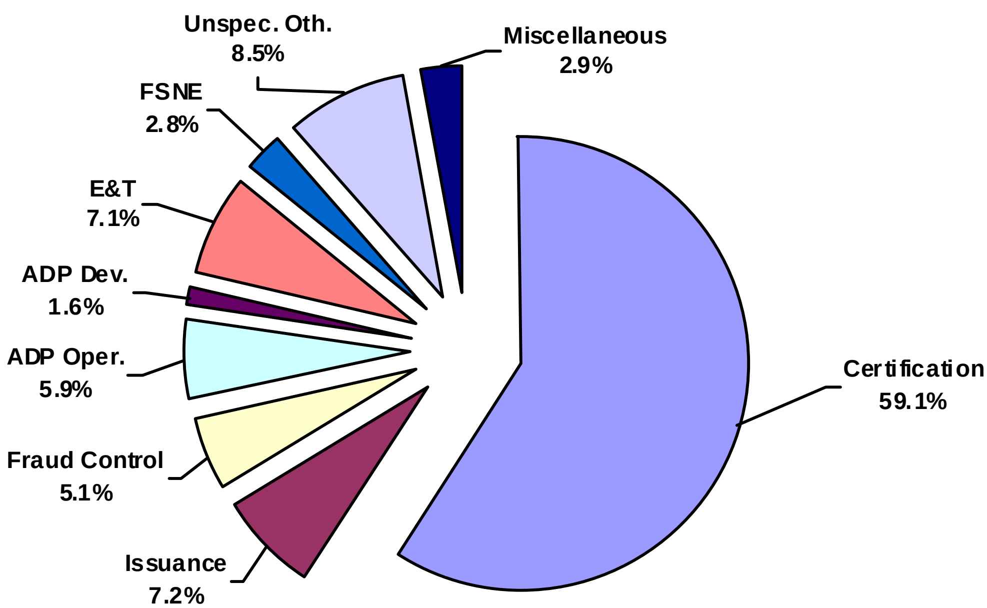
---

[1] Certification-related costs are computed in a variety of ways, but they can be reasonably simplified as the product of the average pay per FTE, the number of worker FTEs, and an "overhead" multiplier for nonpersonnel costs, such as facilities and supplies. Thus, if the cost is divided by the pay per FTE, the result is the number of FTEs times the multiplier.

The average annual total administrative cost per FSP household for the U.S. from 1989 to 2005 was $479.16. (As a point of reference, the average annual benefit outlay in 2005 was $2,453 per household.) The average annual certification cost was $283.09 per FSP household. The average "unspecified other" cost was $40.63 per FSP household; the average issuance cost was $34.12 per FSP household. For FSNE, the overall average was $13.36 per FSP household, and the average for the years with non-zero totals (1994-2005) was $18.92. The averages for the remaining components ranged from $7.58 per FSP household (for ADP development) to $33.90 per FSP household (for E&T).

Figure ES-1

Percentage of Total FSP Administrative Cost, 1989-2005



Key to abbreviations: "Unspec. Oth."=Unspecified Other. FSNE=Food Stamp Nutrition Education. E&T=Employment and Training. ADP Dev.=Automated Data Processing (ADP) Development.
ADP Oper.=ADP Operations.

Trends in FSP Caseloads, Certification-Related Effort, and Error Rates

This study identified several national trends in the FSP from 1989 to 2005:

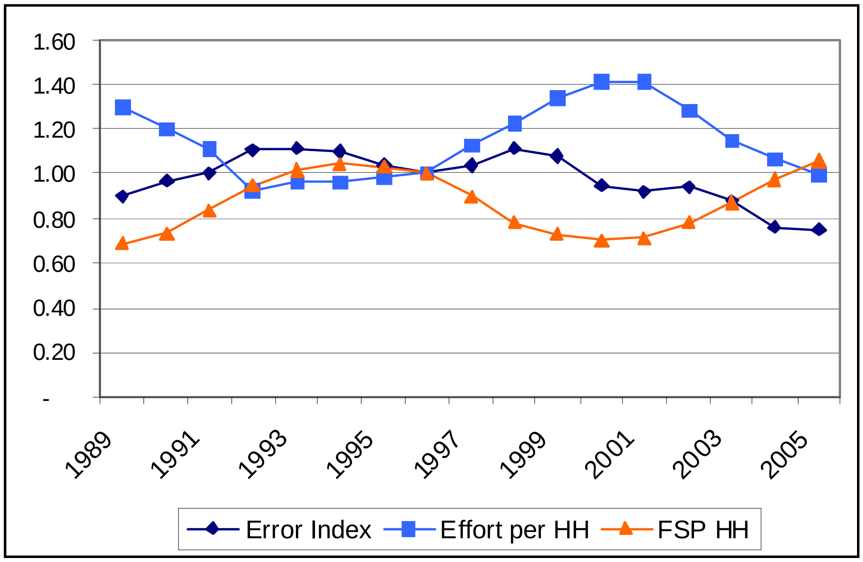
- In the early 1990's, error rates rose while the FSP caseload rose and the certification-related effort per household fell.
- Error rates fell in the mid-1990's as the trends in the FSP caseload and effort reversed.
- In 1997-1998, as States implemented welfare reform, the FSP caseload continued to decrease and effort continued to increase, but error rates rose.

- From 1998 to 2000, error rates again fell, while the decline in the caseload and the increase in effort continued.
- From 2001 to 2005, error rates continued to fall as the caseload increased and effort declined.

Figure ES-2 compares the trends in the average error index (as computed for the study), the average certification-related effort per household (as computed for the study), and the average monthly number of FSP households (i.e., the FSP caseload) from 1989 to 2005. Each variable has been divided by the value for 1996, so the value for 1996 equals 1. For example, the maximum index value for effort per household, 1.41, represents a level of effort 1.41 times the level in 1996.

Figure ES-2

Trends in Error Index, Effort per Household, and Number of FSP Households, 1989-2005 (Annual Averages Divided by 1996 Average)



The trends in these statistics from 1989 to 2000 suggest a positive relationship between error rates and the size of the FSP caseload. The authors' previous study suggested an explanation for this relationship. Average spending per FSP household fell when the number of households increased; the authors explained this pattern as evidence that when caseloads rise, agencies do not proportionately increase staffing levels, and the existing staff serves more cases. The previous study also indicated that a reduction in the ratio of staff to cases—proxied by the effort measure—would increase the error rate, all else equal. On the other hand, falling caseloads free up resources that can be used to reduce errors.

The bivariate trends in 1996-1998 and 2000-2005 seem to contradict this explanation, or at least suggest that other factors had more influence on error rates than caseload trends. The analysis for this report sought to explain these seemingly contradictory patterns by modeling the relationship of FSP error rates to staffing levels, FSP policies, and other factors.

Results of Modeling Error Rates

Like the authors' previous study, this study found a significant and robust negative effect of certification-related effort on error, and that the marginal effect of effort on the error index became smaller (in absolute value) after 1996; i.e., each dollar of spending yielded a smaller drop in the error rate. Thus, it is important to take spending or effort into account when modeling error rates. Over the entire study period, the results imply that an increase of 10 percent in effort would reduce the error index by 0.3 percentage points (relative to a mean of 15.1 percent), at an annual cost of just over $35 per household. The average annual cost of a 1 percentage point reduction in error was $66 per household in 1989-1996; due to the changes in the marginal effect of effort, this cost increased to $296 per household for the 1997-2002 period, then fell to $101 per household for the 2003-2005 period (using the overall average cost per unit of effort).

The present study also confirmed the conclusion of the authors' previous study and others that increased use of short certification periods (1 to 3 months) in the 1990's significantly reduced error rates. For the years with the greatest increase in use of short certification periods (1992 to 2000), the change was 14.5 percentage points, and the estimated net effect on the error index was -1.5 percentage points. The study also confirmed that error rates increased with the percentage of FSP households with earnings. The variables in the regression model and their estimated effects on the error index are shown in table ES-1. The base model included State fixed effects and State time trends, to control for omitted variables that might be correlated with the specified variables. Numerous alternate models were tested, and the findings were highly robust across alternate specifications and models.

Explanation of Trends in Error Rates

Referring back to figure ES-2, the trends in certification-related effort per case played an important role in the trends of the average FSP error index between 1989 and 2005. Based on the model of the error index, we conclude that changes in the average error index were partly or largely due to changes in the average effort per case, in the marginal effect of effort on error, or both. The explanation of how effort influenced error rates is different for the periods leading up to and after the enactment of the TANF program in 1996. The model of error rates also identifies factors other than effort that had important influences on the trends in the error index, particularly after 1996, as discussed below.

The role of effort was greatest in 1989-1996, when the marginal effect of effort on error was greatest in absolute value. During this period, the average level of effort dropped from 12 units per 1000 FSP households to 9 units, a reduction of 25 percent. (In figure ES-2, this change is represented by the drop in the indexed value of effort from 1.3 to 1.0). Based on the average marginal effect for the period, this reduction in effort would be expected to increase the error index by 1.4 percentage points, slightly less than the actual total increase from all factors (positive and negative) of 1.5 points. Another factor contributing to the increase in error in this period was the increase in the percentage of FSP households with earnings. Meanwhile, an increase in use of short certifications and a reduction in the percent of FSP households receiving AFDC pushed the error index in the opposite direction.

Table ES-1

Analysis Variables and Their Estimated Effects on the Error Index: Base Fixed Effects Model

Variable	Definition	Estimated Parameters (with Std. Errors)
ERROR[a]	Error index computed as the weighted sum of positive and negative case error rates (as defined in Chapter Two)	
EFFORT[a]	Certification-related cost per FSP household, normalized by the State wage for a full-time public welfare worker	-7.979*** (2.737)
EFFORT_SQ	Square of EFFORT (used to test for diminishing or increasing returns to effort)	63.309 (82.268)
PCT_TANF[a]	Percent of food stamp households receiving AFDC or TANF	-0.037 (0.065)
TANFIMP	Proportion of months in the year operating TANF program; ranges from 0 (all states, 1989-1996) to 1 (all States in 1999-2005)	-0.019* (0.010)
AFDCWAIV	AFDC reform waiver (percent of months in year)	0.011** (0.006)
NOTI_TANF_EFFORT	Interaction between (1-TANFIMP), PCT_TANF, and EFFORT	2.640 (4.515)
TANFIMP_TANF_EFFORT	Interaction between TANFIMP, PCT_TANF, and EFFORT	17.811*** (3.989)
CM13[a]	Percent of food stamp households with 1-3 month certification periods	-0.184*** (0.042)
CM13_EFFORT	Interaction of CM13 and EFFORT	7.990*** (2.832)
CM1299ER	Percent of food stamp households with earnings having certification period of 12+ months	0.001 (0.010)
PCTEBT[a]	Percent of food stamp households that receive electronic benefits	-0.003 (0.006)
PCT_QUARTREP	State has quarterly reporting for earners (percent of months in year)	-0.012* (0.007)
PCT_SIMPREP	State has simplified reporting for earners (percent of months in year)	-0.036*** (0.007)
PCT_TRANBEN	State has transitional FS benefit for TANF leavers (percent of months in year)	-0.027*** (0.010)
SIMPINCOME	State uses simplified definition of income	-0.009 (0.008)
EARNINC[a]	Percent of food stamp households with earned income in case record	0.225*** (0.060)
SSINC[a]	Percent of food stamp households with OASDI or SSI benefits	-0.009 (0.052)
SINGLEPAR[a]	Percent of food stamp households with children headed by a single adult	-0.010 (0.043)
FYUN0[a]	Unemployment rate	0.123 (0.171)
PCT_CHCASELOAD	Percent change in average number of participating households from previous year	0.016 (0.010)

Table ES-1

Analysis Variables and Their Estimated Effects on the Error Index: Base Fixed Effects Model

Variable	Definition	Estimated Parameters (with Std. Errors)
PCTBLACK	Percent of State population black	0.115 (0.296)
PCTHISP	Percent of State population Hispanic	-0.382 (0.490)
PCTMET	Percent of State population metropolitan	-0.022 (0.047)

[a] Variables used in the previous study (Logan, Rhodes, and Sabia, 2006).

For 1997-1998 and 2001-2005, the trend in the error index was the opposite of what would be expected based on the trend in effort alone. Other factors in the model explain why this happened. Nevertheless, the error index would have been substantially higher in 1997-1998 and lower in 2001-2005 in the absence of the trend in effort.

The model of error indicates that the spike in the error index in 1997-1998 occurred largely because the implementation of TANF reduced the effect of FSP effort on error. This change is highlighted by the estimated cost per case to reduce the error index by 1 percentage point, which jumped from under $66 in 1989-1996 to almost $296 in 1997-2002. The effect of TANF implementation varied, depending on the percentage of FSP households receiving TANF, so the larger the State's TANF program was (relative to the FSP), the larger the reduction in the effect of effort on error. The variable contributing this effect in table ES-1 is TANFIMP_TANF_EFFORT, the interaction of effort with the percentage of FSP households receiving TANF (PCT_TANF) and the indicator for whether a State operated the TANF program during the year (TANFIMP). The TANF percentage fell by more than half from 1997 to 2005, diminishing the effect of this factor on the effort-error relationship. The cost per case to reduce the error index by 1 percentage point was $101 in 2003-2005 (versus $296 in 1997-2002). An increase in the percentage of FSP households with earnings also contributed to the 1997-1998 spike in the error index.

Along with the decline in the TANF percentage, policy variables played an important role in driving down the error index from 1998 to 2005. The key policies changed over time. In 1998-2000, the use of short certification periods grew from 14.7 percent to 18.8 percent of FSP cases. This change offset the diminished marginal effect of effort on error due to the TANF interaction effect. The introduction of quarterly reporting may have also contributed to the reduction in error during these years.

The decline in the error index in 2001 through 2005, despite declining certification effort, was mainly due to the introduction of simplified reporting, simplified definition of income, and transitional benefits for TANF leavers. (Simplified definition of income did not have a significant effect as an individual policy, but it was part of the package of "simplification" policies and is therefore included here.) We estimated that the actual adoption of these policy options reduced the error index by 2.5 percentage points in 2003, 4.0 percentage points in 2004, and 4.4 percentage points in 2005, based on comparison with a simulation of the error index with no States using these options in each year. The effect was smaller in 2001-2002, before all of the options were available to all States. If all States had adopted these options in 2005, the error index would have been 2.8 percentage points lower than it actually was. Thus, these three policy options jointly explain the reduction in the error index, which would have increased if they had not been available. The simulations for this study suggest that wider adoption of simplified reporting and other simplification policies could further reduce the error index, assuming that barriers to wider adoption (such as reprogramming data systems) can be overcome.

Causes of Changes in the Marginal Effect of Effort on Error

The authors' previous study attributed the change in the marginal effect of certification-related effort on error after 1996 to PRWORA in general and found a separate, positive effect of the percentage of FSP households receiving TANF on the level of error. With more data and a better specification, our results in this study suggest a different explanation--an interaction of TANF implementation, the TANF percentage, and effort (i.e., the TANFIMP_TANF_EFFORT variable). This interaction was the largest contributor to the differences in marginal effect between periods. The additional years of data (2002-2005) and the additional policy variables allowed us to improve our analysis of this important change. The present study also uncovered the fact that short certification periods reduce the net effect of effort, underscoring that the benefits of short certification periods are better understood in the context of their costs.

The new results helped us focus on the effects of TANF implementation in understanding why we observe the smaller marginal effect of effort on error after 1996. While the previous study explored the possible benefits of cost-sharing between cash assistance and the FSP, the present study suggests that it is challenging to maintain FSP accuracy while operating TANF and the FSP, with their different rules and priorities, in the same offices. There is no evidence that this challenge existed when States operated the AFDC program under standard rules, but presence of AFDC reform waivers was associated with higher error rates. We find some evidence that the effect of the interaction of effort with the TANF percentage was transient, but we cannot rule out the more parsimonious conclusion that the effect for a given TANF percentage was the same from 1997 to 2005, and that it is evidence of a long-term change.

It is possible that changes in State procedures for allocation of shared costs across programs contributed to the observed reduction in the effect of effort on error, but the evidence from the study casts doubt on this interpretation. We did not find a significant interaction of the percentage of FSP households with AFDC with effort in the pre-TANF period, contrary to what would be expected if cost allocation practices under AFDC consistently shifted common costs to AFDC that would otherwise have been shared by the FSP. Furthermore, we did not find evidence that States with higher percentages of food stamp households receiving AFDC were more likely to have substantial changes in the share of common costs allocated to the FSP after 1996.

Limitations of the Study

One important limitation of the study is that the measure of certification-related effort is a proxy for the ideal measure of actual staff time devoted to certification. As noted above, the underlying hypothesis was that when workers spend more time per case, they make fewer errors. The effort measure varies not only because of differences in actual certification worker time per case but also because of differences in staffing mix and overhead. In addition, the definition of certification-related costs includes some costs that may not be related to certification, because of variations in how States report costs. Further, certification costs may be spent on improvements in other dimensions besides accuracy, such as timeliness of application processing and facilitating access to the FSP. There is still uncertainty about how variations in cost allocation methods, both over time and among States, affect reporting of FSP administrative costs. Thus, more detailed studies of FSP administrative costs may be needed to validate that the effort measure—as computed from the available data—is a good proxy of actual certification staff time.

Another limitation of the study is the lack of success in using automated data processing (ADP) costs as a measure of a State's degree of automation when modeling error rates. These costs had no identifiable effect on their own, and results did not change when they were combined with certification-related costs in the effort measure. It is possible, of course, that automation does not have a detectable effect on error once effort is taken into account. The alternative explanation is that ADP costs are a poor proxy for the capabilities of a State's ADP system to reduce errors. Better understanding of the variation in automation among States might permit development of a proxy measure that could be used more successfully in modeling error rates. An alternative would be a more in-depth approach to relate specific changes in automation to the likelihood of error in activities affected by the automation.

We note that an important but unaddressed feature of the 2002 Farm Bill was the reform of quality control sanctions, so that State liabilities for excessive error rates were based on two years' performance, not just one. With only three years of data after the enactment of this provision, we did not attempt to model its impact.

Last but not least, the study's results are inconclusive on a key question: whether the reduced effect of effort on error after 1996 was transient or more long-term. Additional years of data might help, but it is also possible that, as error rates continue to drop to historically low levels, the power to detect changes in the effect of effort may diminish.

Chapter One: Introduction and Study Highlights

This report presents a study of the relationships of State certification error rates in the Food Stamp Program (FSP) to State certification expenditures, program policies, caseload characteristics, and economic conditions, in the period from 1989 to 2005. The study analyzed recent trends in these variables and then examined the causes of recent declines in error rates. In particular, the study explored the relationship of error rates to State expenditures on certification and recent options for program simplification through multivariate analysis of state panel data.

In this chapter, we discuss the context of the Food Stamp Program, the purpose of the study, its contribution to the literature on error rates, and the organization of the report.

Context

The FSP provides assistance to low-income Americans for the purchase of food at authorized stores, farmers' markets, and other locations. The Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA) administers the program in partnership with the 50 States and the District of Columbia.[2] FNS establishes FSP policy, oversees State FSP administration, and directly manages the participation of retailers and financial institutions. State agencies establish procedures, operate data processing systems, and contract for services provided by other government agencies or private organizations. Local food stamp offices (usually operated by State, county, or municipal agencies) process applications and provide other program-related services to food stamp applicants and recipients.

The administration of the FSP is a major expense to FNS and the States. In Federal Fiscal Year (FFY) 2005, the cost of State and local FSP administration was $4.99 billion (according to estimates computed for this study). FNS pays about half of this cost and thus has a substantial interest in the prudent expenditure of these funds. The total cost of FSP benefits was $28.57 billion (FNS, 2007a), so the total cost of the FSP was $32.56 billion, of which administrative costs represented 14.9 percent. During FFY2005, about 11.2 million households participated in the FSP in the average month (FNS, 2007b), and the annual administrative cost was $447 per household.

Preventing and detecting benefit issuance errors is a major concern of FNS and the States. FNS pays the entire cost of the benefits and, therefore, exercises considerable oversight to assure the accuracy of certification, the process of determining eligibility and benefits. This oversight includes technical assistance, the establishment of plans of corrective action and, in some cases, financial sanctions and bonuses for States. During the period from 1993 to 2005, the proportion of FSP benefits representing overpayments fell from 8.27 percent to 4.53 percent, and the annual cost of overpayments fell from $1.82 billion to $1.29 billion (without accounting for the increase in average benefits due to inflation) (FNS, 1994 and 2007c).

---

2 The FSP also operates in the Virgin Islands and Guam. Due to the unique operating conditions in these territories, it was expected that relationship of error rates to administrative costs and other factors might be atypical and create analytic problems. Given the very small size of the FSP in these territories, they were not included in this study.

As shown in table 1-1, the period covered by the study includes three distinct phases in the history of the FSP, with different trends in FSP participation and other key events that shaped FSP operations.

Table 1-1

Highlights of the Three Periods Included in the Study

Start of Period	Number of FSP Participants (millions)[a]	End of Period	Number of FSP Participants (millions)[a]	Key Changes in the FSP and Its Environment
1989	18.8	1996	25.5	FSP participation peaked at 27.4 million in 1994. State reforms to AFDC through Federal waiver authority. PRWORA enacted in 1996, replacing AFDC with TANF.
1997	22.9	2000	17.2	FS/TANF agencies made major changes in operations to implement TANF. [b] Number of TANF participants fell by 42 percent. Increased emphasis on error reduction in FSP. [c]
2001	17.3	2005	25.7	Recession squeezes State budgets. Farm Bill enacted in 2002, provides options to simplify FSP administration, reduce errors and increase access.

Sources: [a] FNS, 2007a; [b] Holcomb and Martinson, 2002; [c] Rosenbaum, 2000; GAO, 2001.

1989-1996: FSP participation reached a record level and major welfare reform was enacted. FSP participation rose from 18.8 million persons in 1989 to 27.4 million in 1994, the highest level achieved in the history of the program (FNS, 2007a). Participation fell modestly (about 2 percent per year) in 1995-1996. Starting in 1993, States began reforming the Aid to Families with Dependent Children (AFDC) program through waiver authority granted by the Federal government. At the end of this period, the Personal Responsibility and Work Opportunity Act (PRWORA) was enacted. This act made major changes to the FSP and created the Temporary Aid for Needy Families (TANF) program to replace the Aid to Families with Dependent Children (AFDC) program. In 1996, about 37 percent of FSP households received AFDC, and the administration of the two programs was closely linked in most States (Logan, Sabia and Rhodes, 2006).

1997-2000: FSP participation declined and welfare reform was implemented. From 1997 to 2000, the number of participants in the FSP fell, reaching a level 10 million below the 1994 peak. TANF implementation led to major changes in the operation of FSP/TANF agencies (Holcomb and Martinson, 2002). The number of TANF participants fell by 42 percent, from 10.9 to 6.3 million (HHS, 2007). Meanwhile, there was increased emphasis on error reduction in the FSP, as described in Chapter Two (Rosenbaum, 2000; GAO, 2001).

2001-2005: FSP participation increased and the 2002 Farm Bill was implemented. From 2001 to 2005, the FSP grew again to 25.7 million persons. State budgets were squeezed by a recession (Rosenbaum, 2003). As a result of FNS regulatory changes and the 2002 Farm Bill, States implemented new options to simplify FSP administration, reduce certification errors, and improve access to the program. (The 2002 Farm Bill was formally titled the “Farm Security and Rural

Investment Act".) The Farm Bill also changed FSP rules so that State liabilities for sanctions were based on two years' performance, not just one.

In the following chapters, changes in FSP policy and related trends in FSP error rates and administrative costs during the study period are described.

Study Objectives and Research Questions

This study described FSP certification error rates and administrative costs from 1989 to 2005, and then used this information and other data in multivariate models to explain variation in certification error rates among States and over time. Specifically, the study addressed the following research questions.

1. What was the annual cost of FSP administration during this period, and what were its components?
2. What were the trends in costs and staff effort per FSP household for certification, and in overall costs per FSP household from 2002-2005? Were the costs and effort per FSP household similar to or different from costs in 1989-2001?
3. How did the cost components contribute to the overall trends?
4. What was the relationship of certification error rates to the effort spent by State and local personnel on certification-related activities, caseload characteristics, program policies, and economic conditions during 1989-2005? Did the observed relationships in 1989-2001 persist in 2002-2005?
5. What was the relationship of certification error rates to new State FSP policy options after controlling for other factors known to be related to error rates?

The first three questions are descriptive, while the last two are causal.

Contribution to the Literature

This study extended previous research on the same topics for the period from 1989 to 2001 (Logan, Rhodes, and Sabia, 2006) and other studies of variation in FSP error rates. The previous study found that the level of effort expended by State and local agencies on certification had a negative relationship to a weighted index of all types of certification errors in the FSP, after controlling for caseload characteristics, policies, and economic conditions. While previous research had sought to explain variation in error rates, it had not taken into account the effect of the level of effort (Puma and Hoaglin, 1987; Mills, 1991; Kabbani and Wilde, 2003; Mills, LaLiberty, and Rodger, 2004). Logan, Rhodes, and Sabia (2006) found that the estimated effect of certification effort on error was smaller after 1996. Several possible explanations were offered, including: disruption of operations during the implementation of PRWORA, changes in cost allocation rules and practices, and unknown changes in State policies and procedures. With only five years' data after the enactment of PRWORA, the previous study could not differentiate temporary (implementation) effects from more long-term (structural) effects.

The collection and analysis of data for 2002-2005 allowed the present study to test whether the change in the estimated effect of effort on error was temporary or long-term. Furthermore, data on new policy options in the 2002 Farm Bill were used to estimate those options' effects on error rates.

A key question for the study was the effect of one of these options—simplified reporting. With simplified reporting, the State can assign a certification period up to 12 months, with reporting required at least every six months and otherwise only if the household's gross income exceeds the FSP eligibility limit. Within these limits, participants and agencies can ignore changes in household income and expenses during the certification period.

The effect of simplified reporting on error rates was estimated by a recent simulation study (Fink and Carlson, 2005). Using Food Stamp Quality Control System data for Fiscal Year 2000, the authors' simulations produced estimates that simplified reporting as used in 2004 reduced payment error rates (percent of benefits issued in error) by 1.2 to 1.5 percentage points, from the estimated rate of 8.5 percent in the absence of simplified reporting. The impact estimate considers only the proportion of errors that would be disregarded under simplified reporting. Simplified reporting could have an indirect effect, by freeing up certification effort that could be used to prevent or detect other types of errors. Unlike the simulation approach used by Fink and Carlson, the present study used econometric time-series modeling to estimate the effect of several policy options, staff time spent on certification, and caseload characteristics on error rates.

Organization of the Report

Chapter Two of this report describes the types of certification errors in the FSP, the quality control data that are used to estimate error rates, and the trends in certification errors during the study period. This chapter thus provides background information on the dependent variable in the regression models. Chapter Three identifies FSP administrative activities, provides descriptive information on FSP administrative costs from 1989 to 2005, explains the measure of certification effort computed for the study, and presents descriptive statistics on this key explanatory variable in subsequent analyses. Chapter Four discusses the literature on FSP error rates and the conceptual model of the relationship of error rates to certification effort, policies, caseload characteristics, and other factors. Chapter Five presents the methods and results of the modeling, and Chapter Six discusses the conclusions. Appendix A provides details on the source, definition, and computation of administrative cost measures. Appendix B provides supporting information on the econometric methods for the base model. Appendix C presents the results of alternative econometric models, and Appendix D provides supporting information on the methods for these models.

Chapter Two: FSP Certification Errors and Quality Control

As background to the multivariate analysis of the factors that may affect FSP certification error rates, this chapter provides descriptive information on these error rates. We first identify the types of certification errors that occur in the FSP. Next, we describe the quality control (QC) process that produced the public-use QC data used for this study, and we discuss the key changes in policies regarding FSP error rates during the study period (1989-2005). We also present a descriptive analysis of trends in FSP certification errors during the study period.

Types and Sources of FSP Certification Errors

FSP rules and performance measures identify four types of errors in the determination of household eligibility and the calculation of benefits:

- Payments to ineligible households
- Overpayments to eligible households
- Underpayments to eligible households
- Negative action errors, i.e. the denial or termination of benefits to eligible households.

In this report, as in FSP terminology, “certification” includes the initial certification of households, recertification of households previously approved for benefits, and processing of periodic reports or interim changes. An extensive array of information must be determined when a household applies, and information subject to change must be updated when required by FSP rules (at recertification or more frequently, depending on the information and its impact on eligibility and benefits). “Household errors” occur when information is misreported or not reported when required. “Agency errors” occur when the agency fails to use the information correctly by miscalculating amounts or applying rules incorrectly. The following are the general categories of information used in FSP certification and the potential sources of error for each category.[3]

- Household composition: the certification worker must determine which individuals make up the FSP household according to FSP rules. The worker may fail to include an individual who should be considered part of the FSP household, or fail to exclude an individual who should not be counted.
- Income: the certification worker must determine the total household income from all countable sources. Errors may arise due to unreported sources of income, misreporting of income, misapplication of rules determining whether income is counted, or incorrect calculation of income.

[3] The potential sources of error are listed to provide background for the reader, and to motivate later discussion of case characteristics that may be related to the incidence of error. Listing a potential source of error is not intended to imply that it is common. The relative importance of these potential sources was not addressed by this study

- Deductions from income: the certification worker must determine the correct amount of deductions from household income to establish the net income. Allowable deductions include a portion of earnings, excess shelter costs, dependent care expenses, child support payments, and medical expenses. Errors arise because of incorrect or changing deductions, and because of errors in determining a household's deductions.
- Assets: FSP rules restrict the value of liquid assets and vehicles that eligible households may own. Sources of error include unreported assets, incorrect determination of asset values, and incorrect application of rules regarding the treatment of assets in certification.
- Other eligibility factors: to be eligible, an adult must meet applicable work requirements, which depend on age and responsibility for dependents. FSP rules deny benefits to certain types of non-citizens and convicted felons, and to individuals who have been disqualified from the FSP for program violations. Errors occur because of misreported information, unreported changes in status, or mistakes in applying this information.
- Benefit computation: once a household has been determined eligible, the worker must compute the household's monthly food stamp benefit, based on the applicable household size and net income after deductions. This process is often automated, but workers may make errors in applying rules.

Underpayments, overpayments, and payments to ineligible households can occur in any month that a household is active as an FSP case. Negative action errors can only occur when a worker takes an action that denies benefits to a household—either denying an application for benefits or terminating an active case.

One important source of error is the volatility of income, employment status, and deductible expenses. Another is the complexity of FSP rules and the differences in eligibility rules among assistance programs (GAO, 2001).

In this report, we distinguish between **positive** and **negative** errors, from the perspective of the impact on benefit payments. Food stamp eligibility workers make **positive** errors when they approve benefits for ineligible households, or when they approve benefits that are greater than what households are eligible to receive. Workers make **negative** errors when they deny benefits to eligible households or approve benefits that are less than what households are eligible to receive. Both positive and negative errors diminish the integrity of the FSP, which depends on the payment of correct amounts to eligible households.

FSP Quality Control

Under FSP rules, States must maintain a quality control (QC) process. Each State must review a sample of active cases and a sample of cases subject to negative actions, in order to determine the annual rates of the four types of errors. FSP rules specify the sample sizes and the procedures for these reviews. Nationwide, about 50,000 active cases and 40,000 negative action cases are selected for QC reviews each year. For active cases, the QC reviews include examination of electronic and hard-copy case records, household interviews, and collateral contacts (employers, landlords etc.). These reviews identify errors made by the FSP agency and incorrect reporting of eligibility information by FSP households. FNS conducts QC re-reviews of a subsample (one third) of state QC reviews to verify that the QC procedures were followed and the information was used correctly. For the negative action case sample, the State must review the case record, but no other information collection is required. Until FY2000, FNS did not re-review the state negative action QC reviews.

The threshold for a countable error changed during the study period. From FY1989 through FY1999, an error was counted if the monthly benefit was more than $5 too high or low. This threshold (the "tolerance level") was increased to $25 in the FY2000 QC reviews. Therefore, fewer cases were deemed to be in error.

There are two ways of computing error rates from the QC data. FNS monitoring has focused on **payment error rates**, i.e., the ratio of the dollar value of underpayments or overpayments to the total amount of authorized benefits. QC data can also be used to compute **case error rates**, i.e., the percent of FSP cases with specified types of errors. Case error rates are available or can be computed for ineligible cases and negative action errors, as well as for eligible cases with underpayments and overpayments. Payment error rates are not available for negative action errors.[4]

States are liable for sanctions (i.e., financial penalties) if they have excessive payment error rates, after adjustments by FNS. In FY1989 through 1997, a State incurred a liability if its combined payment error rate (underpayments plus overpayments) exceeded the average for all States, with the liability increasing on a sliding scale based on the amount over the national average. Thus, each State had to outperform roughly half of the other States to avoid QC sanctions (CBPP, 2001). As discussed below, States often were not required to pay the full amount of their liability for payment errors.

From FY1998 through FY2002, sanctions were determined after the error rates were adjusted for States with above-average percentages of FSP households with earnings or immigrants, and for States with above-average increases in one or both of these percentages. (The base year was 1992 for increases to FY1998 and FY1999; the base year was 1996 for FY2000 and FY2001. The national average error rates for determining performance were not adjusted.) For FY1998-1999, errors under $25 were ignored in computing the adjusted error rates for the purpose of establishing sanctions. Prior to 2002, there were no financial incentives to reduce negative action errors.

The 2002 Farm Bill (P.L. 107-171) made important changes to the financial incentives for the QC system, effective for FY2003 and later years. Under this law, States are penalized only for "persistently high error rates", defined as a combined payment error rate of more than 105 percent of the national average (with 95 percent confidence) for two consecutive years. States can earn performance bonuses by achieving the lowest error rates, the most improved error rates, high performance in the timeliness of eligibility determinations, or high performance in program accessibility. These bonuses replaced the enhanced funding that previous legislation authorized for States with combined error rates of 5.9 percent or less. At the same time, the Farm Bill introduced several policy options to simplify FSP administration and thereby reduce the potential causes for errors (as discussed in Chapter Four) (ERS, 2007).

States must pay the amount of their QC sanctions to FNS or establish agreements with FNS to reinvest a portion of the liability. Under such an agreement, the State must use reinvested funds to improve their processes for preventing and detecting error through worker training or other methods. A portion of the sanction remains "at risk", to be paid if the State does not achieve its target for error

---

4 In a negative action review, the case record may not contain sufficient information to determine the correct benefit.

reduction. Prior to enactment of the 2002 Farm Bill, FNS used its authority to waive a portion of State liabilities if States met specified targets for error reduction.

From the perspective of FSP households' resources, overpayments and certification of ineligible households have a positive impact: the affected households receive more benefits than they would be entitled to. (These benefits may be recovered by the FSP agency if the error is later detected.) Therefore, these two types of errors are treated as **positive errors** in this and subsequent chapters, and the sum of the case overpayment rate and the ineligible case rate is the **positive error rate**. On the other hand, underpayments and incorrect denial or termination of benefits have a negative impact on eligible households, so the sum of these case error rates is defined as the **negative error rate.**

Use of QC Data for this Report

Error rates analyzed in this report are derived from two sources of data. The first source is the QC public use microdata files for 1989 through 2005. These data comprise the active case QC sample, which is representative of all active FSP cases at the state and national levels. Each case record in these files indicates whether the QC review identified an underpayment or overpayment and the amount of the error. Prior to 2003, the case records identified cases that were determined ineligible in the QC reviews; these cases were excluded in the data for 2003-2005.

As noted above, from FY1989 through 1999, an underpayment or overpayment of at least $5 per month was identified as an error, while the threshold for identifying errors was $25 per month for FY2000 through FY2005. To make the definition of errors consistent across the whole time period, we used the QC data to recompute error rates using a consistent $25 per month definition in all years.[5] **Thus, trends in error rates and analysis results presented in this report do not include the effect of increasing the threshold for countable errors.** It is important to note, however, that State policy and management decisions regarding error reduction prior to FY2000 were based on the lower threshold of error that was in effect at the time and the error rates that were computed on this basis. It might be objected that States expended more certification-related effort before FY2000 than they would have if the $25 error threshold had been in place. On the other hand, sanctions were based on payment error rates, and errors over $25 had more influence on these rates than those between $5 and $25. Thus, there is some possibility of a distortion of the relationship of error rates to certification-related effort, but this possibility is of less concern than the problem that would arise by using an inconsistent measure of error.[6]

Case error rates—i.e., the ratios of cases with specific types of errors to all active cases—were computed for overpayments, underpayments, and ineligible cases, using the FSP QC public use microdata and the published case error rates. Published case overpayment rates from FNS annual QC reports were used for 2003-2005, because the microdata for these years do not permit the computation of error rates inclusive of ineligible households.[7] (The published overpayment rate is the proportion

---

5 It was not possible to use the $5 threshold for FY2000-2005 because errors of less than $25 were not recorded in the data.

6 The modeling issues raised by change in the error threshold are discussed in Chapter 5.

7 Ineligible households were included in the QC sample in 2003-2005, and reviews of these cases were used in computing the reported error rates. These households were excluded from the public use file starting in 2003. The documentation does not explain the reason for this change.

of FSP cases that are ineligible or received countable overpayments.) Published case overpayment rates for 2000-2002 were compared with error rates computed from the QC data; both sets of rates used the $25 error threshold. This comparison indicated an average difference in the rates of about 1.3 percentage points, possibly due to changes in the sampling weights or data cleaning in the preparation of the public-use data. Therefore, the case error rates for 1989-1999 were computed as follows. First, the QC microdata were used to compute the following percentages for each State in each year:

P_{io5} = percent of cases ineligible or with overpayments more than $5
P_{io25} = percent of cases ineligible or with overpayments more than $25

The positive error rate for each State in each year (P_{err}) was then computed for 1989-1999 as:

P_{err} = (published case overpayment rate + published ineligible case rate)*(P_{io25}/ P_{io5})

An equivalent adjustment was made to the published case underpayment rate. These adjustments to the published error rates for 1989-1999 made them comparable to later years, in terms of the threshold for countable payment errors.

The FNS annual QC reports were used as the source for information on negative action errors. Although microdata are not available on these errors, FNS publishes summary data on negative action QC reviews. FNS computes the negative action error rate as a percentage of **negative actions**. Reported negative action error rates were used.[8]

Both the negative action error rate and the active case error rates needed to be recomputed so that they had the same denominator, including all cases subject to sampling for quality control review. To facilitate explanation of these computations, table 2-1 defines the categories of active and negative action cases and assigns each a letter. The computations are described below the table.

- Before adjustment, the active case error rates included the following:
    - Overpayment case error rate=B/E
    - Ineligible case error rate=C/E
    - Underpayment case error rate=D/E.
- The active case error rates were adjusted by the ratio of active cases (E) to the sum of active cases and cases subject to negative actions (E+I).
- The published negative action error rate (H/I) was multiplied by the ratio of negative action cases (I) to the sum of active cases and cases subject to negative actions (E+I).

These adjustments assured that the error rates had the same denominator and were fully comparable.

As a result of this adjustment, these rates do not correspond to case error rates published elsewhere. For errors involving active cases, the difference is quite small, because the ratio of negative actions to active cases is very small.

---

8 Beginning in FY2000, FNS validated negative action error rates through re-reviews of the State samples. While the validated rates are preferable when available, for consistency across years, we use the reported rates.

Table 2-1

Components of Case Error Rates

Status	Active (Paid) Cases	Negative Action Cases
Correct	A	G
Overpayment error	B	(not applicable)
Ineligible case error	C	(not applicable)
Underpayment error	D	(not applicable)
Negative action error	(not applicable)	H
Total	(A+B+C+D)=E	(G+H)=I

It is important to note that the QC error rates in this report represent the proportion of cases with **reported** errors. It is conceivable that a QC review might fail to detect an error, or find a case to be in error that is in fact correctly paid. The QC review system is quite mature, however, and FNS and the States follow an extensive system of procedures—including standardized sample sizes and sampling methods—to assure that data are collected consistently and accurately (regulations at 7CFR Part 275).

Trends in Error Rates and Caseloads

Figure 2-1 compares the national trends in the average positive and negative case error rates (as computed for this study), the average combined case error rate, and the average monthly number of FSP households (i.e., the FSP caseload) from 1989 to 2005.[9] This figure shows different patterns for two periods: 1989 through 2000, and 2001 through 2005.

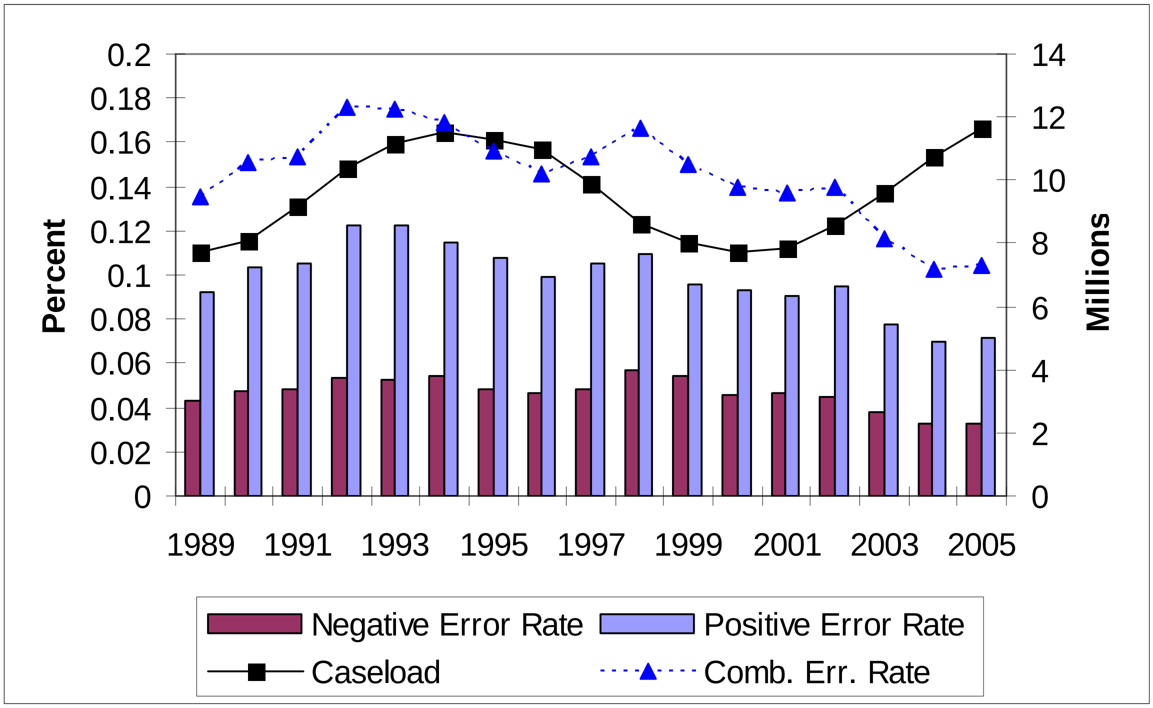
From 1989 through 2000, the error rates tended to increase when the caseload increased and fall when the caseload fell, although this pattern did not hold in 1997 and 1998.

- The national average positive error rate and the caseload increased from 1989 to 1993 and generally declined thereafter. The caseload peaked at 11.5 million in 1994.
- The positive error rate deviated from its general trend by increasing from 1996 to 1998, while the FSP caseload dropped by 2.4 million. These were the first two years of implementing PRWORA.
- The national average negative error rate had a similar but less pronounced pattern of rising and falling with the FSP caseload from 1989 to 1996, a more pronounced increase (in relative terms) from 1996 to 1998, and a less pronounced drop from 1998 to 2000.
- The average combined case error rate increased by 4.0 percentage points from 1989 to 1992, fell by 3.0 percentage points from 1992 to 1996, rose again 2.1 points from 1996 to 1998, and fell by 2.7 points from 1998 to 2000.

[9] The FSP caseload measure includes active cases and cases subject to negative action (denials and terminations), so it does not correspond exactly to published statistics.

Figure 2-1

FSP Caseload (Millions of Participating Households) and Case Error Rates, 1989-2005



Notes: Caseload includes active cases and cases subject to negative action. Negative error rate is percentage of caseload with underpayments or negative action errors. Positive error rate is percentage of caseload with overpayments or ineligible for benefits. Error rates were computed as a percentage of active FSP cases plus negative actions. A constant error threshold of $25 per month was used in estimating error rates.

On the other hand, the caseload and the combined error rate moved in opposite directions from 2000 through 2005.

- While the FSP caseload rose from 2000-2005, the positive error rate continued its previous declining trend, although it deviated from this trend with an up-tick in 2002.
- After a slight up-tick in 2001, the average negative error rate declined in 2002 through 2005.
- Overall, the average combined case error rate dropped 3.5 percentage points from 2000 to 2005, while the caseload increased by 3.9 million households.

The trends in these statistics from 1989 to 2000 suggest a positive relationship between error rates and the size of the FSP caseload. Average spending per FSP household fell when the number of households increased, as discussed in the next chapter. The authors' previous study suggested that this pattern is evidence that when caseloads rise, agencies do not proportionately increase staffing levels. Therefore, rising caseloads stretch the existing staff of FSP agencies over more cases. The previous study also indicated that a reduction in the ratio of staff to cases would increase the error rate, all else equal. On the other hand, falling caseloads free up resources that can be used to reduce errors.

The bivariate trends in 1996-1998 and 2000-2005 seem to contradict this explanation, or at least suggest that other factors had more influence on error rates than caseload trends. The analysis for this report sought to explain these seemingly contradictory patterns by modeling the relationship of FSP error rates to staffing levels, FSP policies, and other factors.

Chapter Three: FSP Administrative Costs and Certification Effort

This study used data on FSP administrative cost outlays for 1989-2005 provided by FNS from the agency's National Data Bank. FNS compiled these data from State cost reports, which include both State and local administrative costs. The sources of the FSP administrative cost data, the cost reporting categories, and the methods used to analyze them are described in Appendix A. The costs represented in this report do not include expenditures for FNS' own activities to administer the FSP.[10]

This chapter begins with a description of the FSP administrative activities for which costs are identified. Next, we summarize the trends and composition of FSP administrative costs from a national perspective. The estimates in this section were computed by adding up State totals. States with large numbers of participants make up a greater proportion of the national cost totals than smaller States. Therefore, the larger States have more influence on the national data. Summary measures of cost variation among States are also presented. The final section explains the computation of the measure of certification effort for the study and provides descriptive statistics on trends in this measure.

Administrative Activities in the Food Stamp Program

The major administrative activities of State and local FSP offices are:

- Certification
- Issuance of benefits
- Automated data processing system development and operations
- Fraud control and fair hearings
- Employment and training
- Nutrition education.

In addition, there are a number of minor administrative activities, some of which are mandatory and others, such as outreach, that are voluntary. State and local agencies use their own personnel, purchased services, supplies, data processing and communications systems, office equipment, and owned and leased facilities to administer the FSP.[11]

---

[10] In addition to oversight of State and local operations, FNS authorizes and oversees the participation of retailers and financial institutions in the FSP. In FY2005, the total Federal share of FSP administrative costs (including Federal, State and local administration) was $2,509.5 million (FNS, 2007a). The total Federal share of State and local administrative costs was $2,324.3 million. Thus, FNS spent about $185.2 million on FSP administrative expenses at FNS headquarters, regional offices and field offices.

[11] Interviews with FNS experts conducted in 2007 for another study and the authors' prior research indicate that local office personnel costs are the largest single component of FSP administrative costs, and personnel costs are also the largest component of costs at the State level. There are, however, no systematic national data on the composition of FSP administrative costs by object of expenditure (personnel, supplies, etc.).

Certification

Certification activities include the processing of initial applications for Food Stamp benefits, periodic recertification of Food Stamp households, and other actions to obtain, verify, and apply information on households' FSP eligibility and entitlement to benefits. These activities are generally performed in local FSP offices by State or local government employees, although some States use centralized call centers for some activities that do not require face-to-face contact (GAO, 2007). Local FSP offices usually perform certification tasks for State-administered cash assistance programs (such as Temporary Assistance for Needy Families or TANF and general assistance) and often for medical assistance, low income energy assistance, and other means-tested programs for low-income residents.

Issuance of Benefits

During the study period, States generally used one of two systems to issue FSP benefits. At the outset of the period, all States issued paper food stamp coupons to recipients, with varying methods and administrative structures (including State, local and contractor operations). In 1993, Maryland was the first State that entirely replaced the coupon issuance system with an electronic benefit transfer (EBT) system, which are similar to debit card systems. By FY2001, 80 percent of FSP households received benefits via EBT. The nationwide implementation of EBT was completed in 2004. Nearly all States contract with private firms to operate their EBT systems, but some issuance-related functions (such as card issuance) may be performed at local FSP offices (FNS, 2007d). During the study period, about one percent of FSP households received benefits in cash as a supplement to Social Security Income (SSI) or State aid to elderly or disabled persons.[12]

Automated Data Processing System Development and Operations

Each State operates a computer system to support FSP operations, usually as part of an integrated eligibility system serving the Temporary Assistance to Needy Families (TANF) and other related programs (such as State general assistance, Medicaid, or the Low Income Home Energy Assistance Program). Some State Food Stamp Agencies operate their own computer centers, while others contract out this function or use a center that services multiple State agencies.[13] Automated data processing (ADP) costs include development of new or upgraded systems, operation and maintenance of computer hardware and software, data communications networks, and local office equipment. Prior to 1995, FNS provided enhanced funding (up to 75 percent Federal) for approved ADP development costs.

Fraud Control and Fair Hearings

State and local FSP agencies investigate evidence of fraud by recipients through review of case records, interviews with recipients, and third-party sources of information. Fraud investigations most often involve misrepresentation of eligibility or dual participation, but other types of investigations deal with false reports of lost benefits and trafficking in benefits. When the FSP agency finds

---

[12] These "elderly cash-out" projects are limited in scope and specifically authorized by FNS. Due to their small size, these projects would not be expected to have a material effect on national averages and trends in issuance costs. California SSI households receive a State cash payment instead of food stamp benefits and are not counted as FSP households.

[13] Source: Personal communication with FNS grants management officials, October 2007.

evidence of fraud, it may initiate recovery of funds, termination of benefits, or prosecution. FSP agencies also initiate recovery of funds when recipients receive overpayments due to unintentional error by the recipient or the agency.

FSP agencies conduct fair hearings upon request by applicants or recipients who are subjected to adverse action (benefit reduction or termination). Such adverse actions are not necessarily the result of error or fraud.

Employment and Training

Each State provides FSP employment and training (E&T) services for FSP recipients who are subject to job search and work requirements. E&T services are provided through local FSP offices and other agencies. Allowable E&T components include supervised job search, job search training, work experience or workfare, vocational training, work-related education, and self-employment training.

Nutrition Education

States have the option of providing nutrition education to persons who are eligible for Food Stamp benefits, with the goal of promoting healthy food and lifestyle choices. Most States have agreements with the State Cooperative Extension Service, State universities, or other agencies to provide food stamp nutrition education (FSNE). FSNE grew from 7 States in FFY 1992 to 49 States (counting the District of Columbia as a State) in FFY 2001; all 50 States and the District of Columbia provided FSNE in FY2005.[14]

Other State and Local Program Administration

State and local FSP agencies carry out a number of other administrative functions, including quality control, management evaluation, outreach, demonstration projects, and oversight of program operations. As noted earlier, under the mandatory quality control (QC) system, each State must review a sample of active FSP cases and a sample of closed or denied cases to determine whether the determination of eligibility and benefits is correct. Each State must also conduct management evaluation reviews of local office operations. States may choose to conduct approved outreach activities with their own personnel or in partnership with community and faith-based organizations.

National Summary of FSP Administrative Costs

The average grand total real annual cost of FSP administration for the U.S between 1989 and 2005 was $4.30 billion, including Federal and non-Federal funds. (All cost estimates in this report are in 2005 dollars, so inflation was not a factor in this increase. The price deflator for the Gross Domestic Product was used as the index of inflation.)[15] Of this total, the average total annual cost for certification was $2.54 billion. Thus, the certification cost was 59.1 percent of all administrative costs.

---

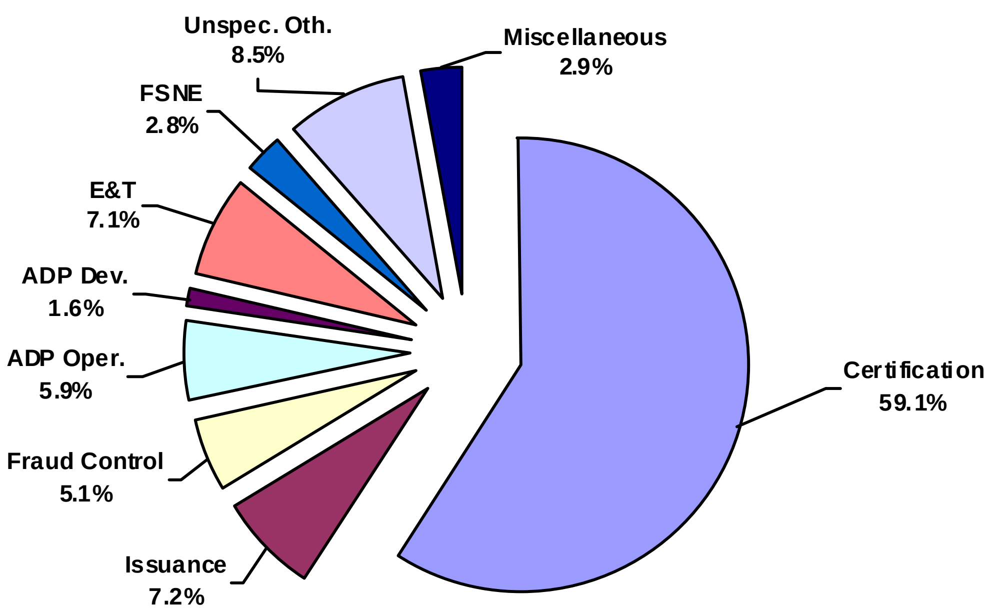
[14] The Virgin Islands also provided FSNE in FY2005; Guam did not.

[15] For this national summary, costs were adjusted for the national average rate of inflation. As discussed later in the chapter, the measure of certification-related effort takes into account State differences in pay rates for personnel, which represent the primary factor of production in FSP administration. See Appendix A for discussion of the choice of price deflator.

Figure 3-1 shows the percentage distribution of the national total FSP administrative cost for the period (summed over all 17 years) among the analysis categories. The cost categories correspond to the FSP administrative activities described in the preceding section. Further information on the cost categories is provided in Appendix A. As discussed later in this chapter, there is evidence that assignment of costs to reporting categories has varied among States and over time. The amount of variation is not known but does not appear to have a major impact on national averages, with the possible exception of the "unspecified other" category.[16]

Figure 3-1

Percentage of Total FSP Administrative Cost, 1989-2005



Key to abbreviations: "Unspec. Oth."=Unspecified Other. FSNE=Food Stamp Nutrition Education. E&T=Employment and Training. ADP Dev.=Automated Data Processing (ADP) Development.
ADP Oper.=ADP Operations.

16 "Unspecified other" costs are, by definition, the FSP costs that are not reported in any other category. States vary in their use of the defined cost categories versus the "unspecified other" category. See Appendix A for more information.

Certification was by far the largest category, representing three-fifths (59.1%) of the total.[17] The next largest categories were unspecified other (8.5%), issuance (7.2%), E&T (7.1%), and ADP operations (5.9%). The smallest categories were miscellaneous, fraud control, FSNE, and ADP development. The small share for FSNE for the period as a whole (2.8%) was in part due to this category being added in 1994. In 2005, FSNE costs were 4.7% of the national total FSP administrative cost, while certification costs were 60.4% and unspecified other costs were 4.8% of the national total.

As shown in table 3-1, the grand total real annual cost of FSP administration for the U.S rose from $3.18 billion in 1989 to $4.99 billion in 2005, an increase of 57%. These are total expenditures. Some of their variation appears to be due to variation in the total caseload (i.e., we would expect total

Table 3-1

Annual Total FSP Administrative Cost in Billions of 2005 Dollars, Total and Components for U.S., 1989-2005

Fiscal Year	Total Cost	Certifi-cation	Issuance	Fraud	ADP opera-tions	ADP devel-opment	E&T	Miscel-laneous	FSNE	Unsp. Other
1989	3.18	1.84	0.22	0.12	0.17	0.04	0.23	0.10	0.00	0.45
1990	3.25	1.91	0.23	0.13	0.17	0.03	0.28	0.09	0.00	0.42
1991	3.39	2.03	0.24	0.13	0.18	0.05	0.26	0.09	0.00	0.40
1992	3.62	2.16	0.29	0.14	0.21	0.04	0.25	0.09	0.00	0.43
1993	3.73	2.22	0.30	0.16	0.21	0.04	0.25	0.10	0.00	0.45
1994	3.92	2.31	0.34	0.22	0.21	0.06	0.24	0.10	0.01	0.44
1995	4.13	2.41	0.35	0.23	0.23	0.09	0.29	0.11	0.02	0.39
1996	4.22	2.48	0.36	0.24	0.24	0.08	0.26	0.11	0.05	0.40
1997	4.27	2.49	0.33	0.26	0.24	0.07	0.28	0.11	0.08	0.41
1998	4.35	2.52	0.33	0.26	0.28	0.08	0.28	0.13	0.10	0.36
1999	4.58	2.68	0.32	0.27	0.31	0.08	0.30	0.13	0.14	0.35
2000	4.74	2.82	0.30	0.26	0.32	0.06	0.31	0.14	0.17	0.35
2001	4.88	2.92	0.30	0.27	0.31	0.07	0.33	0.15	0.18	0.35
2002	5.26	3.13	0.32	0.27	0.31	0.10	0.40	0.17	0.29	0.25
2003	5.26	3.10	0.34	0.27	0.29	0.11	0.40	0.16	0.35	0.25
2004	5.27	3.11	0.36	0.24	0.28	0.09	0.40	0.16	0.39	0.23
2005	4.99	3.01	0.32	0.22	0.32	0.06	0.42	0.17	0.24	0.24
Mean	4.30	2.54	0.31	0.22	0.25	0.07	0.30	0.12	0.12	0.36
Dollar Change ('89 v. '05)	1.81	1.17	0.10	0.10	0.15	0.02	0.19	0.07	0.22	-0.21
Percentage Change ('89 v. '05)	57.0%	63.3%	43.3%	86.0%	84.3%	41.6%	80.6%	71.8%	1964.4%	-46.4%

[a] Mean for FSNE is for all years. FSNE was included in "Unspecified Other" in 1992-1993. Mean FSNE cost for 1994-2005 was $0.17 billion. Dollar and percentage change for FSNE are from 1994 to 2005.

[17] As explained in Appendix A, FNS deducts from Federal share of certification costs an "offset" amount that is deemed to be included in the State's TANF grant because of the way that the grant was set. This offset reduces the Federal share and increases the non-Federal share, but it does not directly affect the total cost. It is possible, however, that some States reduced their total certification spending so that the adjusted non-Federal share would match the available funds. Certification costs in this study are equal to the total of Federal and State expenditures for certification as reported. Further details on the computation of certification costs are provided in Appendix A.

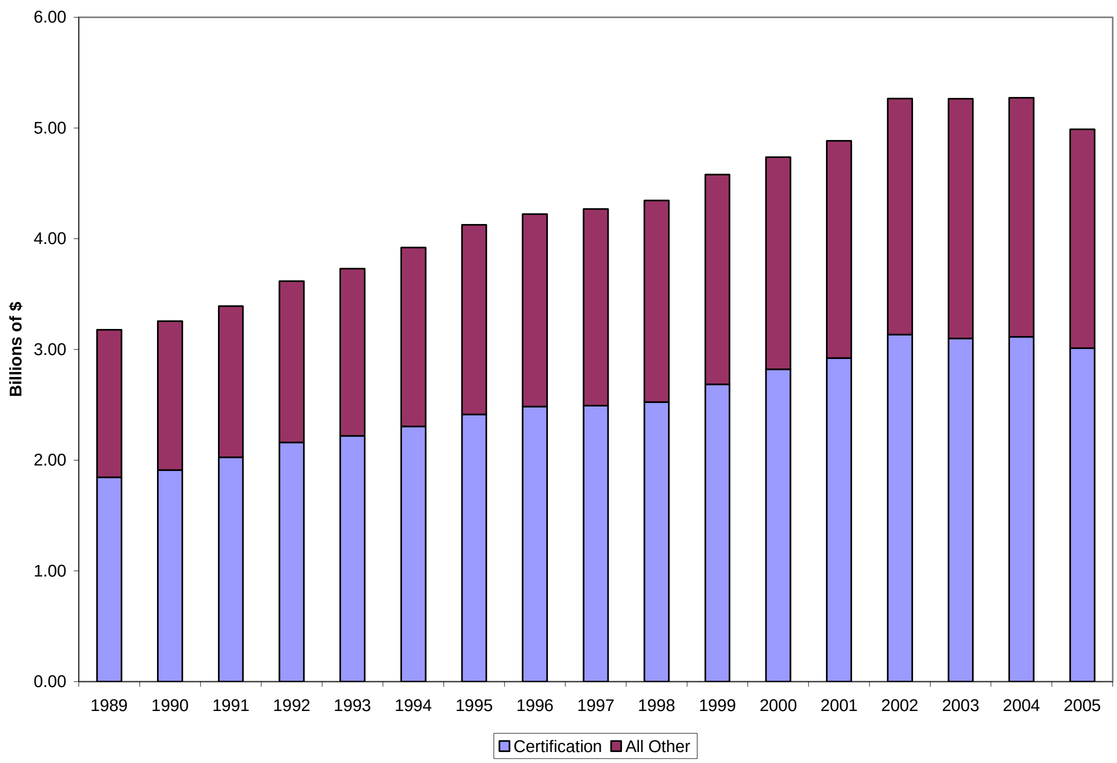
expenditures to rise and fall with the caseload). We defer until the next sub-section discussion of per-case expenditures and trends in that measure.

At the national level, the grand total FSP administrative cost increased every year until 2002 (even after adjusting for inflation). This cost leveled off in 2002-2004, and then fell in 2005. The trends in total certification costs and all other costs combined were the same, as illustrated in figure 3-2. Certification cost was between 58.0 and 60.4 percent of the total over the period. Of the individual cost categories shown in table 3-1, all but one (the "unspecified other" total) increased between 1989 and 2005.

Different trends were observed in the total costs of several categories. Total benefit issuance costs rose steadily until their peak of $0.36 billion in 1996, declined over the next five years, and then rose again on 2002-2004. The decline corresponded with the period of widespread conversion to EBT; the rise occurred as participation grew. Total fraud control costs increased steadily until 1997, when these costs essentially reached a plateau where they stayed until falling in 2004-2005. Total automated data processing (ADP) development costs fluctuated over the period, peaking in 1995 (the first full fiscal year after the federal share was reduced to 50 percent) and 2003.

Figure 3-2

Annual Totals of Certification and All Other FSP Administrative Costs for U.S. in Billions of 2005 Dollars, 1989-2005



Total ADP operations costs rose intermittently until they peaked in 2000; thereafter, they declined until 2004 and rose again in 2005 to match the 2000 peak. Total employment and training (E&T) costs fluctuated up and down but showed an overall increasing trend with a notable jump from 2001 to 2002. The total costs for the group of activities in the "miscellaneous" category showed a general upward trend from 1994 to 2002, then stayed about the same.[18] Food Stamp Nutrition Education (FSNE) costs were not reported before 1994, but they showed the most dramatic percentage rate of increase: 1964% over eleven years.[19] The dollar increase in FSNE was much less dramatic, due to the small size of this program component. Among costs reported for all seventeen years (from 1989 to 2005), the largest percentage increases were in fraud control (86%) and ADP operations (84%), while the smallest increases were in issuance (43%) and ADP development (42%). Only the "unspecified other" total declined over the period (by 24% from 1989 to 2001).[20]

Administrative Costs per FSP Household: Averages and Trends

As shown in table 3-2, the average annual total administrative cost per FSP household for the U.S. from 1989 to 2005 was $479.16.[21] The average annual certification cost was $283.09 per FSP household. The average unspecified other cost was $40.63 per FSP household; the average issuance cost was $34.12 per FSP household. For FSNE, the overall average was $13.36 per FSP household, and the average for the years with non-zero totals (1994-2005) was $18.92. The averages for the remaining components ranged from $7.58 per FSP household (for ADP development) to $33.90 per FSP household (for E&T).

Figure 3-3 compares the national trends in total administrative cost per FSP household, certification cost per FSP household, and participating FSP households. Earlier we noted that one would expect total expenditures to rise and fall with the caseload. Furthermore, we noted that in general actual total expenditures (as shown in table 3-2) followed the expected pattern. Nevertheless, figure 3-3 suggests another factor at work. The number of FSP households rose from 7.5 million in 1989 to 11.5 million in 1994, then declined to 7.7 million in 2000; it then rose to 11.6 million in 2005.[22] For most of the period, as the number of FSP households rose, the total *administrative cost per FSP household* declined—from $411 per FSP household in 1989 to $316 per FSP household in 1993. Conversely, as the number of FSP households fell from 1994 to 2000, the total cost per FSP household increased. The total cost per FSP household fell again during the surge in participation from 2000 to 2005. A nearly identical trend occurred in the national certification cost per FSP household.

---

[18] The "miscellaneous" costs include quality control, fair hearings, System for Alien Verification of Eligibility (SAVE), outreach, management evaluation, and demonstration/evaluation projects.

[19] FY1992 was the first year when States operated approved plans for FSNE. Prior to FY1994, FSNE costs were included in the "unspecified other" category. The Federal share of the FY1992 FSNE cost was $661,000 (in 1992 dollars); only 7 States had approved FSNE plans for that year. (Source: http://www.nal.usda.gov/foodstamp/pdf/FNSFactSheet03.pdf.)

[20] Possible explanations for the decline in total unspecified other costs include more careful allocation of direct or indirect costs that should be reported elsewhere; actual reduction in spending; or a combination of these factors.

[21] The averages for the entire period in table 3-2 are averages of the annual national averages, not the total cost divided by the total number of case-years.

[22] Beyond the study period, the number of participating FSP households increased to 11.7 million in FY2006.

Table 3-2

Average Annual FSP Administrative Cost per Household in 2005 Dollars, Total and Components for U.S., 1989-2005

Fiscal Year	Total Cost	Cert	Issuance	Fraud	ADP op	ADP dev	E&T	Misc	FSNE	Unsp Oth
1989	440.79	255.84	31.10	16.22	24.15	5.48	32.06	13.39		62.54
1990	420.13	246.59	29.17	16.54	22.02	4.05	35.64	12.10		54.01
1991	382.47	228.32	27.30	14.73	19.99	6.13	29.84	10.61		45.56
1992	359.96	215.01	28.44	14.28	20.91	4.45	24.99	9.24		42.64
1993	346.03	205.84	28.12	15.23	19.41	3.94	23.03	9.12		41.34
1994	353.83	208.12	30.40	19.49	18.80	5.56	21.89	9.22	1.03	39.31
1995	379.67	222.03	31.81	21.29	21.35	7.84	26.91	10.32	2.20	35.93
1996	400.75	235.67	33.86	22.93	23.19	7.28	24.80	10.02	4.73	38.25
1997	453.08	264.52	35.39	28.04	25.59	7.16	29.36	11.29	8.00	43.72
1998	527.81	306.54	39.56	31.97	34.07	9.84	33.49	15.70	12.72	43.92
1999	598.01	350.53	41.95	35.06	39.88	10.62	38.95	17.50	17.70	45.81
2000	647.00	385.29	41.46	35.20	43.64	8.75	42.06	19.37	23.07	48.14
2001	656.96	393.13	40.63	36.01	41.60	9.04	44.78	20.12	24.58	47.06
2002	643.39	383.04	39.63	32.91	38.44	12.73	48.32	21.32	36.01	30.99
2003	575.73	338.98	37.58	29.21	31.50	11.81	43.72	17.56	38.01	27.37
2004	513.54	303.31	34.82	23.53	27.42	9.20	39.04	15.85	37.92	22.46
2005	446.58	269.68	28.76	19.47	28.73	5.01	37.37	14.85	21.07	21.63
Average[a]	479.16	283.09	34.12	24.24	28.28	7.58	33.90	13.98	13.36[b]	40.63

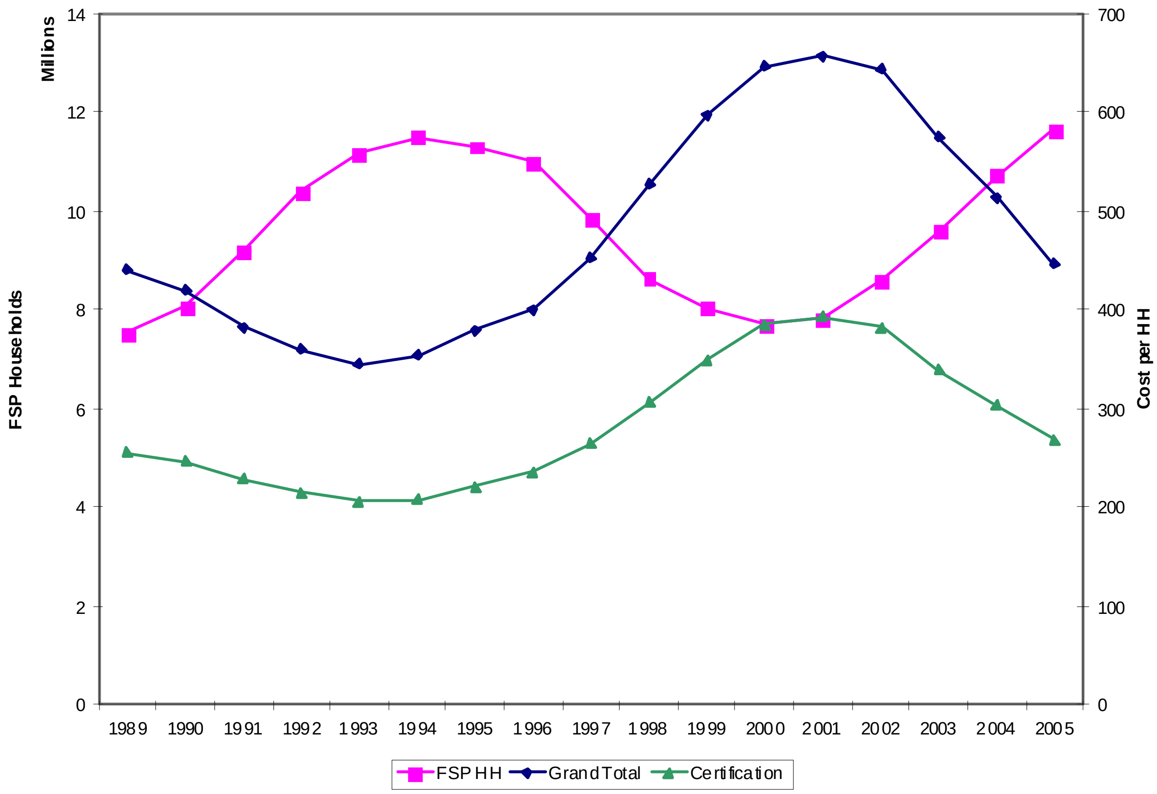
Source: FNS National Data Bank. Inflation adjustment and cost per household computed for this study.

[a] Average of annual average costs per household.

[b] FSNE was not authorized until 1992 and not reported separately until 1994. The annual average FSNE cost for 1994-2005 was $18.92 per household.

Figure 3-3

Average Annual Total FSP Administrative and Certification Costs per Household (in 2005 Dollars) and FSP Caseload of Participating Households for U.S., 1989-2005



Source: FNS National Data Bank. Inflation adjustment and cost per household computed for this study.

These trends are largely consistent with what one would observe if total FSP administrative costs changed less than proportionately in response to increases or decreases in the FSP caseload. The State share of spending for welfare programs (including the FSP) is appropriated before the fiscal year begins, and staffing levels are determined primarily by appropriated funds. When the FSP caseload increases, the State agency may not have funds available within its appropriation to increase staff. When the FSP caseload decreases, the agency may choose not to decrease staff, in order to retain experienced workers and to assure adequate staff in the event of another caseload increase.

Real total FSP costs and certification costs increased almost every year in the study period (as shown in table 3-1). As a result, the total and certification costs per FSP household in the years after 1993 were greater than they had been in prior years with comparable participation levels. The upward trend in the cost per household began in 1994, as FSP participation was peaking. The cost per FSP household in 1989 was matched in 1997, when there were 2.3 million more participating households (and thus a lower cost per household would be expected). The total cost at the peak of participation (11.5 million households) in 1994 was $354 per household; when this peak was reached again in 2005, the total cost was $447 per household. Thus, the chart can be interpreted as reflecting a combination of two trends: a cyclical pattern and a non-cyclical long-term rising trend. Under the cyclical pattern, the costs per FSP household tended to increase as FSP participation decreased, and vice versa. Meanwhile, there was a general trend of increasing costs per FSP household, starting in 1994, independent of the number of the number of households. The significance of this pattern for the study is that one would expect to be able to separate the effects of caseload levels and spending per household on error rates, if these factors have such effects.

Variation Across States in FSP Administrative Costs per Household

There was considerable variation in FSP administrative costs across the States. As shown in table 3-3, the average total cost per FSP household from 1989 to 2005 for individual States ranged from $122 to $1,681, with a median of $444—less than the national average of $479. Half of the States had average costs per FSP household between $345 (the 25th percentile) and $577 (the 75th percentile). The difference between the median and mean, and the size of the range above the 75th percentile, indicate that the cost distribution was skewed, with the top 25 percent of States having a disproportionate influence on the national average. Like the total, the component costs per FSP household had wide overall ranges (between minimum and maximum) and skewed distributions, with the majority of States (the 25th to 75th percentiles) in a narrower but still substantial range, as shown in table 3-3.[23]

Welfare Worker Pay and FSP Certification Effort

One would expect that workers with smaller caseloads would be less likely to make errors in determining FSP eligibility and benefits. Ideally, it would be useful to compare certification workloads across States, using a measure such as the ratio of full-time food stamp workers to food stamp households, rather than the cost per household. There is no extant database with this information. It is possible, however, to construct a proxy measure of the level of effort per FSP household using available data on administrative costs and public welfare worker pay rates. (The

Table 3-3

Annual FSP Administrative Cost per Household in 2005 Dollars, Total and Components, Summary of State Averages

Statistic	Total Cost	Cert	Issuance	Fraud	ADP op	ADP dev	E&T	Misc	FSNE [a]	Unsp Oth
Minimum	122.12	0.65	1.10	0.00	0.00	-0.24	0.15	0.84	0.00	-47.70
25th Percentile	345.43	180.88	22.75	7.92	13.25	0.00	10.38	9.18	1.79	0.18
Median	444.37	257.35	30.92	14.02	26.58	0.08	19.70	14.82	14.08	15.39
75th Percentile	576.76	337.20	46.13	25.04	51.19	5.51	39.40	23.24	32.14	59.20
Maximum	1681.39	955.00	303.14	160.83	336.16	268.23	328.15	98.22	308.41	238.68
US Average	479.16	283.09	34.12	24.24	28.28	7.58	33.90	13.98	18.92	40.63

Source: FNS National Data Bank. Inflation adjustment and cost per household computed for this study.

[a] FSNE statistics are for years with nonzero national totals (1994-2005).

[23] Negative costs appear in table 3-3 as minimum values for ADP development and unspecified other. These appear to result from reporting anomalies and are not widespread.

source of these rates is discussed below.) We constructed an effort measure for certification-related activities, taking the certification-related cost per FSP household and dividing the cost by the state wage for a full-time public welfare worker.

Definition of Certification-Related Cost

Certification-related cost was defined to include all FSP administrative costs except ADP, issuance, FSNE, and E&T. Thus, the effort measure was computed as in (3-1) below:

$$EFFORT = \left[\frac{ATC - AADPC - AIC - AFSNEC - AETC}{HH}\right]\left[\frac{1}{W_{FTE}}\right] \qquad (3\text{-}1)$$

where ATC is the annual total FSP cost, AADPC is the annual automated data processing (ADP) cost, AIC is the annual issuance cost, AFSNEC is the annual Food Stamp Nutrition Education cost, AETC is the annual employment and training cost, HH is the number of food stamp households (computed by averaging monthly data) and W_{FTE} is the annual public welfare worker wage rate per full-time equivalent (FTE) employee. All variables are specific to a year within a State, and costs and pay rates are in 2005 dollars. The first quantity in equation (3-1) represents the certification-related cost per FSP household. Dividing the certification-related cost per FSP household by the public welfare wage rate normalized the effort measure to control for differences in pay rates.

The definition of certification-related costs used in this study (and the authors' previous study) comprises the FSP administrative cost categories that we expected to have a possible impact on FSP certification errors. These categories are the following:

- the "certification" category as identified in FSP reports, which primarily represents the cost of local office eligibility workers, support staff, and their supervisors.
- the "fraud control" category, representing staff and other costs to investigate possible fraud and establish and collect claims for overpayments.
- the "miscellaneous" category, which consists primarily of costs for functions that contribute directly or indirectly to the performance of local office eligibility workers: quality control, management evaluation, verification of alien eligibility, and fair hearings.[24]
- the "unspecified other" category, at least part of which was likely to be related to certification and error-reduction (as discussed earlier in this chapter).

Analysis of the cost data and discussions with FNS indicated that States vary in the definitions that they use for "certification", "fraud control", "unspecified other", and the reporting categories grouped as "miscellaneous". Given that establishing and maintaining eligibility information is the primary function of State and local FSP staff, we treated all of these reporting categories as "certification-related" costs to obtain a consistent measure.

---

[24] Outreach and research and demonstration projects were also included in this category, but they were relatively minor components.

Automated data processing (ADP) costs were excluded from the certification-related cost measure, because a substantial portion of ADP costs were for equipment, software, and purchased services. Thus, the public welfare wage was not a suitable measure to convert ADP costs to units of resources. Spending on ADP systems might, however, affect the level of error or the effectiveness of eligibility workers. The relevant measure could be current-year spending, cumulative investment, or a combination of the two. Alternative methods of including ADP costs in models of error rates were explored, as discussed in Chapter Five.

The other FSP cost categories discussed in this chapter were excluded from the definition of certification-related costs, because they were not expected to affect QC errors. Benefit issuance is a separate process with its own measures of accuracy; EBT was implemented in part to reduce errors and fraud in this process. Food Stamp Employment and Training (E&T) and Food Stamp Nutrition Education (FSNE) services address program goals other than determining eligibility and providing benefits.

Public Welfare Pay Rate for Converting Cost to Effort

In principle, certification-related effort should be estimated by dividing the certification-related cost per FSP household by the average pay rate for FSP workers in each State and year. While there is no database of these rates, there is a plausible proxy for use in computing EFFORT. The Annual Survey of Government Employment (ASGE) annually collects payroll data for a large number of occupations, including "public welfare workers". As defined by the Census Bureau, this class of workers includes workers who administer the FSP, AFDC/TANF, medical assistance, and other forms of public aid or services typically targeted to low-income populations.[25] Many FSP eligibility workers also administer AFDC/TANF and medical assistance, and pay scales for specialized FSP workers are likely to be similar to those for multi-program workers and specialists in other programs. Thus, it is reasonable to assume that the overall average pay rate for public welfare workers is an acceptable proxy for the average for those who administer the FSP. The ASGE includes all States and a sample of local governments.[26]

The ASGE data were used to compute the average monthly cost per full-time-equivalent worker (FTE) for 1989-2005.[27] The pay rates were converted to 2005 dollars using the same methodology as was used for the costs.

It is important to note that the public welfare pay measure is a weighted average over all types of State and local public welfare workers, so it reflects the actual labor mix employed by public welfare agencies and their specific wage-setting practices, as well as the labor market from which these workers are hired. Public welfare agencies, including FSP agencies, have some flexibility to offset high wages in the labor market by hiring less-skilled workers. In addition, the scope of the services provided by public welfare workers varies among States and over time. Thus, the average pay rate

---

[25] See www.census.gov/govs/www/classfunc79.html for full definition and examples.

[26] The FSP is administered by county governments with State supervision in 10 States; State employees administer the program elsewhere. (Information from personal communication with FNS headquarters, October 2007.)

[27] The number of hours per month representing an FTE is defined by each State.

reflects a heterogeneous mix of workers and jobs.[28] Ideally, the measure of effort would be based on a standardized unit, such as a full-time worker with specified education and experience. This would require a pay rate for such a standardized caseworker job classification, but such a measure was not available in a State-level time series. We note, however, that persistent differences in skill mix across States may be among the variables captured by the fixed state effects in the econometric models for this study.

Summary Statistics for Certification-Related Cost, Worker Pay, and Effort

Table 3-4 shows summary statistics for the certification-related effort measure and the two measures used to compute it. All of the measures are based on data converted to real (2005) dollars. The national (weighted) average annual certification-related cost was $353.62 per household (computed from State data for 1989-2005). The unweighted median was $326.93 per household; thus, the mean was somewhat but not substantially influenced by the States in the upper end of the distribution. The range was from $91.23 to $1,084.34 per household. Half of the States had costs between $257.59 per household (25th percentile) and $420.52 per household (75th percentile).

The national average annual pay rate per FTE for public welfare workers in 1989-2005 was $35,134, as shown in table 3-4. The median of the State averages was $33,453 per year, and the range was from $18,321 to $56,867 per year. Half of the States had pay rates between $29,961 per year (25th percentile) and $38,982 per year (75th percentile).

Summary statistics for the certification-related effort measure are also presented in table 3-4. The national average was 0.01, or one unit of certification-related effort per 100 FSP households. This is not the same as estimating that the average worker served 100 FSP households, for two reasons.

Table 3-4

Annual Average Certification-Related Cost Per Household, Public Welfare Pay per Full-Time Equivalent and Certification-Related Effort per FSP Household, Summary of State Data, 1989-2005

	Certification-Related Cost Per Household	Annual Avg. Pub. Welf. Pay per FTE	Certification-Related Effort per FSP Household
Weighted Mean	$353.62	$35,134	0.010
Minimum	91.23	18,321	0.004
25th Percentile	257.59	29,961	0.008
Median	326.93	33,453	0.010
75th Percentile	420.52	38,982	0.013
Maximum	1084.34	56,867	0.026
N	867	866	866

Sources: FNS National Data Bank, Annual Survey of Government Employment.

[28] One could instead use the average pay for a standard type of worker to compute the "standard worker units per FSP household". In this measure, a worker that makes 25 percent more than the standard worker would count as 1.25 worker units. This approach assumes that worker productivity is proportional to pay, which we view as a rather strong assumption even in the aggregate.

First, the effort measure was a proxy for the quantity of administrative resources per food stamp household. The cost measure included allocated overhead costs (such as facilities, supplies, office equipment, and ancillary services). Thus, the effort measure cannot be interpreted as labor alone, but rather labor with a multiplier for overhead costs.[29] Second, FSP workers often determine eligibility for other programs as well, such as TANF and Medicaid. Thus, if an agency devoted 1 full-time equivalent of staff time per 150 FSP households and the average worker spent half of his/her time on the FSP, the average worker would serve 75 FSP households.

As with the cost per FSP household measures, there was a wide range in the effort measure among States: from .004 to .026 per household (or from 4 to 26 per 1000 households).[30] Much of this range was below the 25th percentile or above the 75th percentile. Thus, there was substantial variation among States in the amount of administrative resources expended per FSP household.

As discussed in the next chapter, this variation was expected to contribute to the variation in FSP error rates. In the conceptual model used by this study, the level of certification-related effort is a key determinant of the accuracy of FSP certification.

---

[29] For example, assume that overhead costs are allocated to the FSP by adding a fixed amount (O) per full-time equivalent worker. Thus, $CERTCOST=FTE*(W_{FTE}+O)$, and $EFFORT=FTE*(1+O/W_{FTE})$. Actual cost allocation procedures may be more complex. The authors' experience with analyzing the composition of FSP administrative costs suggests that labor costs are by far the largest component of these costs, so this simplification should not greatly affect the results.

[30] The effort measure is presented as used in the multivariate analysis, which was the primary focus of the study. The effort per 1000 households or some variant might be preferable for descriptions of variation across States or over time.

Chapter Four: Literature and Conceptual Model of FSP Error Rates

This chapter reviews the literature on error rates in the FSP, including descriptive studies of trends and error reduction strategies, as well as empirical studies of the determinants of error rates. The chapter then presents the conceptual model of the relationship of certification effort to error rates.

Studies of Error Rates and Error Reduction Strategies

For the purposes of this literature review, the existing literature is divided into two phases: studies of error rates prior to the 2002 Farm Bill, and more recent studies. As discussed in Chapter Two, the 2002 Farm Bill substantially changed the system of rules and incentives that shaped State approaches to error reduction. Thus, studies in the first phase examined a different program environment than studies in the second phase. This section discusses each phase of studies in turn.

Studies of Error Rates Prior to the 2002 Farm Bill

Five studies considered the determinants of FSP error rates before the 2002 Farm Bill. These papers vary in their time periods, methods and results. The two earliest studies used cross-sectional approaches and examined payment error rates (percentage of benefit dollars issued in error). Puma and Hoaglin (1987) analyzed two years of quality control (QC) data (FY1984-1985), with additional population variables from 1980 Census data. They found that the incidence and amount of overpayments were related to household size, sources of reported income, presence of reported assets, number of deductions, and the population density of local office area. Mills (1991) analyzed the relationship between underpayments and overpayments from 1980 to 1990, and found a positive correlation in cross-sectional data and in year-to-year variation. This finding suggested that at least some error-reduction practices had effects on both types of errors, rather than trading some errors for others.

Kabbani and Wilde (2003) conducted a cross-sectional time-series analysis of the total payment error rate (percentage of benefits) from 1990 to 2000. The purpose of this econometric study was to examine the trade-offs between reducing error rates and reducing FSP participation when States shortened certification periods. Parallel models of participation and error rates were developed, incorporating variables used in prior studies of participation rates. The authors consistently found that the proportion of FSP households with short certification periods was associated with their error measure, and this association was stronger than for any other variable in the models. They also found some caseload characteristics that appeared to influence error rates during the period, including the racial and ethnic composition, the percent elderly, and the percent in working households. Higher error rates were associated with the presence of Democratic Governors or legislatures, after controlling for the possible effects of unemployment rates and poverty levels; there was some evidence that poverty levels were positively associated with error rates. FSP outreach expenditures were included in the models, and one version included a proxy variable for monthly reporting, but neither of these variables was significantly associated with error rates. They also did not find a significant effect of AFDC waivers and TANF implementation.

Mills, Laliberty, and Rodger (2004) constructed Markov probability models using QC data for 1997-2001 in a simulation study that analyzed the composition of the total (combined) FSP case error rate and the effects of shortened certification periods. They found that the largest improvement in payment accuracy was in ongoing cases between certifications. The accuracy of cases in the month of initial certification also improved, and the percentage of cases in this status increased, amplifying the effect. They found, in addition, that improvements in recertification procedures appeared to increase the accuracy of expiring cases subject to recertification. Their state-level analysis provided an informative framework for comparing the accuracy of initial certification, interim action, and recertification across States. Simulations of the effect of shortening certification periods for households with earnings suggested a smaller effect on the total case error rate but a similar effect on the participation rate for these households, in comparison to the regression-based empirical results of Kabbani and Wilde (2003). An important factor in the effect on participation was that shortened recertification periods led to more frequent closure rates for correct cases, as well as for cases with errors. The authors suggested that the larger effect on error rates estimated by Kabbani and Wilde may have been due to one or more omitted variables correlated with the short certification variable, but they also acknowledged that further examination and testing of the Markov simulation models was needed.

In the previous empirical study by two of the authors of the present study, Logan, Rhodes and Sabia (2006) analyzed the relationship of error rates to certification effort, policies, and other State characteristics for the period from 1989 to 2001. As in the present study, FSP administrative costs and welfare pay rates were used to construct a measure of certification effort, and a weighted sum of positive and negative error rates (the error index) was the dependent variable. The effort measure had a large and statistically significant effect on the error index, providing evidence that the increase resources expended on the certification process, which began in 1994, contributed to the observed decline in the error index. Other changes in the FSP occurring at the same time as—and possibly related to—the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA) evidently contributed to the fall in the error index. The results implied that, in the post-PRWORA period, States on average had to spend more effort on certification-related activities than in previous years to achieve a given error rate. The study did not find a clear reason for this change, but the authors suggested three possible explanations: temporary changes due to PRWORA implementation, long-term changes related to PRWORA, or changes in cost allocation and reporting that shifted costs to the FSP that formerly were borne by the AFDC program. The models predicted that if a State's FSP certification budget is fixed and the number of FSP households increases, the effort per FSP household will fall and the error rates will rise, all other things equal. The study also confirmed that error rates fell with increased use of short certification periods and rose with the proportion of FSP households with earnings.

The results of these studies were used to identify economic, demographic, and policy variables for the present study. Three of the studies—including the authors' previous study incorporating effort as a control variable— confirmed the importance of short certification periods to error rates, although there were different estimates of the size of the effect. Three of the studies used the combined **payment error** rate (percentage of dollars paid in error) as the dependent variable, while two used the combined **case error** rate (percentage of cases with errors). The authors' previous study was the only one that took into account negative action errors as well as active case errors. More importantly, the authors' previous study identified certification-related effort as a key determinant of error rates, while it also posed the unanswered question of why this relationship changed with the implementation of PRWORA—a key question for the present study.

Studies of Error Rates and State Policies Since the 2002 Farm Bill

Three recent studies have documented and examined a decline in FSP error rates from 1998 to 2004, including the first two years under the 2002 Farm Bill. These studies all focused on payment error rates. Two were descriptive and one used simulations; none used econometric methods. In addition, several studies have examined error reduction practices adopted by States before and after the 2002 Farm Bill.

Studies of Trends in Error Rates

One analysis of error rates (CBPP, 2005) found that the official national overpayment and underpayment rates declined from 1998 to 2004 to the lowest recorded level, and that 48 States had lower overpayment rates in 2004 than in 1998. The CBPP study noted that the decline in payment error rates began during the economic boom of the late 1990's but continued in 2001-2004 while unemployment and FSP participation rose, trends which historically have been associated with increased error rates.

The CBPP study cited anecdotal evidence and expert opinion that FSP policy changes had simplified the determination of eligibility and benefits, contributing to the reduction in error rates. One key change was the option of simplified reporting, which was first authorized for households with earnings in 2000 and expanded to all households by the 2002 Farm Bill (as described in Chapter One). The other key Farm Bill provision cited by CBPP was the "simplified income definition" option for States to determine income and assets using the same definitions and exclusions as in the Temporary Assistance for Needy Families (TANF) and Medicaid, both of which allow States flexibility in defining eligibility standards. By conforming income definitions and exclusions across programs, CBPP argued, States reduced the amount of errors arising from confusion about these rules.

A report by the Government Accountability Office (GAO, 2005) examined the declining FSP payment error rates, the actions taken by States to reduce errors, and the sources of errors. Interviews with State officials identified a variety of methods to reduce error: monitoring at the county level, analysis of causes, specialized change reporting units, simplified reporting, and other FSP options to simplify administration. GAO concluded that no single factor or practice was the dominant reason for the decline in error rates. Analysis of QC data indicated to GAO that the main causes of errors are: not using updated information; not applying FSP rules correctly; clients not reporting changes; and reporting of incomplete or incorrect information.

A simulation study by the FNS (Fink and Carlson, 2005) focused specifically on the effects of simplified reporting on payment errors. The researchers computed adjusted error rates for 2000 (before any State had simplified reporting) by excluding errors that would have been ignored under simplified reporting. Simplified reporting as implemented by 42 States in 2004 was estimated to reduce the combined payment error rate (overpayments plus underpayments as a percentage of benefits issued) by 1.2 to 1.5 points. If all States had used simplified reporting to the greatest extent allowed by FNS regulations, the estimated reduction was as much as 2.2 points. (The actual combined payment error rate declined by about 2.3 points from 2000 to 2004.) At the State level, the estimated impact of simplified reporting varied, depending on the proportion of households subject to simplified reporting, whether the State had a waiver to act on all reported changes, and the proportion of errors attributable to unreported changes in income (absent simplified reporting). This is the only study that has systematically modeled factors contributing to the recent decline in FSP error rates.

An important limitation of the studies of error rates is that they controlled for few facets of State FSP administration other than short certification periods or simplified reporting. There are, however, several studies presenting views of program officials that simplification of FSP certification rules can reduce workloads for caseworkers, as well as reducing the incidence of errors. One GAO study cited State officials who advocated simplification of FSP rules, such as those governing income calculations, as a way to reduce errors and costs (GAO, 2001). Another GAO study (GAO, 2004) found that States' decisions about which of the 2002 Farm Bill FSP simplification options to adopt were influenced by officials' expectations about whether the options would reduce administrative burdens and errors. Local officials had mixed views regarding whether the options adopted by their States had actually reduced burdens on workers and errors; barriers to these outcomes included problems with computer systems and conflicting rules from other programs, such as Medicaid. These studies have two implications. First, simplification may enable States to reduce errors, or to reduce the level of effort needed to achieve the desired level of error, or both. Second, simplification policies may not have the desired effect on errors if there are barriers to implementation, or if other programs' rules for the same households are different.

Studies of State Policies

Recent studies have increased the availability of comparative information on State policies that may affect error rates, and one study in particular created and analyzed a comprehensive longitudinal database. GAO and FNS have published several reports documenting State policies that might have important effects on error rates (GAO, 2002; FNS, 2003). A new longitudinal database of State Food Stamp policies was compiled for a recent study of FSP participation from 1996 through 2003 (Ratcliffe, McKernan and Finegold, 2008). The database included measures of State FSP policy options regarding certification periods, reporting requirements, categorical eligibility, transitional benefits, immigrant eligibility, and vehicle exemptions. The authors found strong evidence that receipt of food stamp benefits by low-income persons was increased by longer certification periods, expanded categorical eligibility for food stamps, and expanded vehicle exemption policies. There was weaker evidence of impacts for simplified reporting, implementation of electronic benefits transfer, restoration of eligibility for legal immigrants, and outreach spending. The State Food Stamp policy database provides an opportunity to examine the effects of these policies on other aspects of FSP performance, particularly error rates.

Thus, the recent literature highlights a number of policy changes that may have affected FSP error rates in recent years, including options authorized or expanded by the 2002 Farm Bill. The present study used these studies to identify policy variables to be included in models of error rates and, in particular, used the database created by Ratcliffe, McKernan and Finegold (see Finegold, Margrabe, and Ratcliffe, 2007 for documentation). While previous studies point to ways that States use their resources to reduce FSP errors, only the authors' previous study addressed the impact of the quantity of resources—i.e., the level of effort relative to the caseload. The present study extends the authors' previous research on effort and error into the period after the 2002 Farm Bill (through 2005) and takes into account the key recent policy changes, particularly the shift from short certification periods to simplified reporting.

Conceptual Model of Effort and Error Rates

This study examined the relationship between States' level of effort on food stamp certification and their error rates. In its simplest form, the conceptual model of the effort-error relationship is:

$$\text{ERROR} = \alpha + \beta \ \text{EFFORT} + \varepsilon \qquad (4\text{-}1)$$

States expend effort to assure that their certification process has an acceptable error rate, both to meet their own standards of integrity and to avoid USDA penalties. The coefficient β is expected to be negative; i.e., higher EFFORT (as a normalized measure of certification-related spending) leads to lower error rates. This study sought to estimate how much an increase in EFFORT lowered ERROR (i.e., what was the value of β). (Throughout this chapter, variable names are capitalized, to distinguish specific variables from more general concepts.) The following discussion defines the terms in the simplified model and then elaborates the model to include other variables.

Definition of ERROR and EFFORT

Before elaborating on the model, the variables ERROR and EFFORT must be defined. This study used the definitions developed for the previous study and discussed in Chapters Two and Three.

The ERROR variable is the weighted sum of the positive and negative case error rates, which were defined in Chapter Two. Thus, this measure combines the incidence of eligible households with overpayments, ineligible households receiving benefits, eligible households with underpayments, and eligible households denied benefits (negative action errors).[31] Sanctions in the QC system are based on payment error rates (percentage of benefits issued incorrectly), but there is no way to include negative action errors in a payment error rate. States with excessive negative action error rates are required to establish corrective action plans, and high performance in the area of negative action error rates is one of the categories for bonuses authorized by the 2002 Farm Bill (7CFR 275.24(b)(2)). Because of the need to take into account all measured types of errors, case error rates were used. As in FNS policy, the rationale for combining positive and negative error rates is that errors in either direction diminish the integrity of the program, and program incentives encourage the reduction of both positive and negative errors. Combining the error rates, rather than modeling them separately, is preferable because the available data on EFFORT do not separate resources spent on reducing different types of errors.[32] As discussed in Chapter Two, national trends in positive and negative error rates over the study period were very similar, providing further support for the approach of modeling the sum of these rates.

The use of a weighted sum takes into account the possibility that EFFORT may not have equal effects on positive and negative errors. In fact, the authors' previous study estimated that an incremental unit of EFFORT produced a larger reduction in the positive error rate than in the negative error rate. A

---

31 Some negative action errors are procedural, and in such cases the household might have been found ineligible if the case had been processed correctly.

32 As discussed in Chapter Five, estimates for alternate models of payment error rates were very consistent with the main results using case error rates. Separate models of positive and negative error rates also yielded results consistent with the main model.

grid search method was used to estimate the weights for positive and negative errors, as discussed in Chapter Five.

EFFORT is a measure of the quantity of administrative resources per case. As discussed in Chapter Three, the ideal measure might be the ratio of caseworkers to cases, but the available data did not provide this information. Therefore, this study used a proxy measure of EFFORT based on the certification-related cost per FSP household, normalized for State pay levels for public welfare workers, as defined in Chapter Three.

Expanded Model of Effort and Error

The primary goal of the study was to identify the causal relationship between EFFORT and ERROR; i.e., the effect of EFFORT on ERROR for a given State in a given period, holding all else constant. Standard regression theory suggests that to do that, the model should include regressors that control as much as possible for "all else". The model in equation (4-1) is too simple to meet this requirement, because it implies that the relation between EFFORT and ERROR is constant across States and time. Instead, we expect that some characteristics of States and time periods shift the level of error (i.e., α in 4-1), for a given level of effort; and some characteristics shift the effect of EFFORT on ERROR (i.e., β in 4-1).

To address these issues we expand the simple model to:

$$\text{ERROR} = (\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_N X_n) + (\beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_M Z_m)\text{EFFORT} + \varepsilon \qquad (4\text{–}2)$$

Thus, the questions for research become: what factors shift the level of ERROR independently of the level of EFFORT (i.e., the X's), and what factors change the effect of another unit of EFFORT on ERROR (i.e., the Z's)? (A factor could, of course, be both an X and a Z, having both an independent effect and an interaction with EFFORT.) We are particularly interested in the FSP policies that may affect the level of ERROR, the response of ERROR to a unit of EFFORT, or both. In addition, the relevant factors may include characteristics of the FSP caseload (e.g., the proportion of households with earnings) and the socioeconomic context (e.g., the unemployment rate). In the following section, we present our hypotheses about the X's and Z's in (4-2)—the factors that shaped the FSP error index (ERROR) during the study period. The hypotheses focus on the variables representing choices by States, including EFFORT and policy options. The additional factors and their expected relationships to ERROR are discussed in Chapter Five.

Before proceeding to discuss the hypotheses, we note that a variant of the model of ERROR would have EFFORT interact with all variables. We preferred the model in (4-2) because it does not force this relationship on all variables; instead it allows for factors whose effect on ERROR does not depend on the level of EFFORT. In specifying the model, we focused on the interactions with EFFORT that were theoretically plausible and expected to have identifiable effects. Other variables were specified as having effects independent of EFFORT.

Hypotheses About the Relationships of Certification Effort and Policies to the FSP Error Index

This study tested several hypotheses regarding the determinants of the FSP error index (the ERROR variable). The first two hypotheses were the principal conclusions of the authors' previous study of error rates in 1989-2001.

1. The FSP error index declines as the level of certification-related effort increases, all else equal (i.e., $(\beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_M Z_m) < 0$).

2. There were two changes in the relationship of effort to error after 1996. First, the level of the FSP error index independent of EFFORT $(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_N X_n)$ decreased. Second, the effect of EFFORT on the FSP error index $(\beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_M Z_m)$ became absolutely smaller. These changes were hypothesized to be the results of some combination of the following changes:
   - PRWORA implementation and its effects on local welfare office operations
   - shortened certification periods
   - other FNS and State initiatives to reduce errors
   - changes in cost allocation rules and practices resulting in increases to the FSP's share of common certification costs for FSP households receiving cash assistance (relative to what would have been allocated under previous rules and practices).

The present study sought to confirm both of these conclusions and to clarify the reasons for the changes observed after 1996.

The third hypothesis was a newly formed expectation based on studies of the FSP after the 2002 Farm Bill:

3. FSP simplification initiatives beginning in 2000 enabled States to attain lower error levels of the error index at a given level of certification-related effort.

Hypothesis 1: The Error Index Declines as Certification-Related Effort Increases and Rises as Effort Falls

The previous study found that ERROR declined as EFFORT increased during the 1989-2001 period (i.e., in equation 4-1, $\beta<0$). The present study tested whether this basic relationship continued in 2002-2005, and it re-estimated the quantitative relationship between error and effort.

To explain this observed relationship, we start with the assumption that the likelihood of error for a case depends on the actual level of effort spent on the case, the set of tasks that must be done to certify the case, and the difficulty of these tasks. The difficulty of a task may be related to the number and complexity of the applicable rules, the amount of information required, and the ease or difficulty of processing the information. As the actual level of effort for the case increases, the workers responsible for the case can devote more time to one or more tasks and reduce the likelihood of error for those tasks. The likelihood of error for a case will rise if the level of effort stays constant while a factor increases the number of tasks required for that case or the difficulty of those tasks. Although the analysis was conducted at the State level, this case-level framework provided a basis for identifying the variables that might affect the impact of EFFORT on ERROR, as discussed in the sections on the second and third hypotheses.

This relationship of certification worker time to accuracy only holds if other administrative factors are held constant, however. When workers have less time available relative to the volume and difficulty of the tasks, they may maintain the same level of accuracy by processing applications on a less timely basis. Additional certification worker hours may be used for purposes other than error reduction, such as to improve timeliness or access. For example, agencies may improve access by keeping offices open longer hours or out-stationing workers at locations other than FSP offices; these actions may require more staff hours to serve a given caseload with a given level of accuracy.[33] Since we only observe the level of EFFORT expended on all certification-related activities, it is necessary to control for such unobserved differences among States, which may change systematically over time.

Considering the patterns of FSP administrative cost per case, it appears that the total level of certification-related effort adjusts slowly to changes in the number of cases. It follows that, in the short run, an increase in the caseload (e.g., during a recession), will lead to lower certification-related effort per case (EFFORT), and we would therefore expect more errors (a greater value of ERROR). A simple trend analysis provides mixed evidence in support of this expectation. As shown in table 4-1, the expected pattern can be seen in 1989 through 1992. Although the caseload continued to grow in 1993-1994, the effort per case (EFFORT) also increased, showing evidence of the delayed response of total effort to increasing caseloads. The error index (ERROR) rose during these years from .153 to .198. Conversely, EFFORT continued to rise in 1995 through 2001, as the caseload fell. ERROR fell as would be expected in 1995 and 1996, but then, counter to expectations, spiked up in 1997 and 1998 before returning to the previous pattern of falling as EFFORT rose. Between 2000 and 2005, the caseload rose, and the level of EFFORT fell, but ERROR also fell. The year-to-year changes did not always fit this overall pattern. Decreases in EFFORT were associated with decreases in error in 2001, 2003, and 2004. On the other hand, decreases in EFFORT were associated with increases in ERROR in 2002 and 2005.[34] The trends from 1989 to 2005 suggest two conclusions. First, changes in EFFORT were sometimes, but not always, negatively related to changes in caseload. Second, while the association of trends in EFFORT and ERROR suggests a relationship, other factors beside EFFORT clearly were influencing ERROR, particularly in 1997-1998, 2001, 2003, and 2004.

---

33 An agency may need more staff hours per case to provide expanded hours or to out-station workers, for two reasons. First, there may losses in productivity, if the flow of clients per hour is less than the number that can be served with the available staff. Second, there may be fixed costs to have an office open for a given time period, so the cost per case goes up as the number of clients served per hour goes down.

34 Small changes in the error index may not be statistically significant. Error rates are estimated from sample data. We have not attempted to compute confidence intervals for the error index. As a point of reference, the combined payment error rate (percent of benefits) in 2005 was 5.84 percent, and the 95 percent confidence interval was plus or minus 0.25 percent (in decimal form, .0584 plus or minus .0025). If the coefficient of variation for the error index were the same, the confidence interval of the 2005 estimate would be plus or minus .005, and the change from 2004 to 2005 would not be statistically significant. The significance of the changes in the error index from 2000 to 2002 is also questionable.

Table 4-1

FSP Caseload, Certification-Related Effort per Household, and Error Index, 1989-2005

Year	FSP Caseload (millions of households)	Certification-Related Effort per 1000 FSP Households	Error Index[a]
1989	7.522	10.848	0.153
1990	8.076	10.242	0.170
1991	9.198	9.386	0.174
1992	10.406	7.751	0.198
1993	11.175	8.326	0.197
1994	11.510	8.436	0.192
1995	11.309	8.698	0.176
1996	10.998	9.132	0.165
1997	9.870	10.210	0.174
1998	8.646	11.495	0.191
1999	8.028	12.680	0.173
2000	7.700	13.276	0.159
2001	7.825	13.300	0.156
2002	8.597	12.065	0.158
2003	9.586	10.618	0.132
2004	10.723	9.460	0.116
2005	11.645	8.428	0.118

[a] The error index is the weighted sum of positive and negative error rates (in decimal form) as computed for the study, computed by grid search as described in Chapter Five. The negative error rate includes negative action errors. The denominator for the positive and negative error rates includes active and negative action cases.

Sources: USDA-FNS National Data Bank and Quality Control Annual Reports

A key question for the study, therefore, was whether the declining level of certification effort per FSP household could be reconciled with the decline in the error index. One possible explanation was that the relationship of the error index to certification effort (i.e., β in equation 4-1, or the sum of the β terms in equation 4-2) changed during the period after 2000. (As noted, the previous study found such a change between the 1989-1996 period and the 1997-2001 period.) Another was that other changes in the FSP—particularly the reforms to simplify change reporting—offset the effects of budget constraints by independently shifting the level of the error index (i.e., a change in the sum of the α terms in equation 4-2).

Hypothesis 2: The Level of Error and the Error-Effort Relationship Changed After 1996

The previous study identified three effects associated with the implementation of PRWORA: a reduction in the level of the average error index (i.e., an effect on α), a positive association between the error rate and the percentage of FSP households receiving TANF (another effect on α), and a reduction in the absolute size of the change in the error index associated with a change in certification-related effort (β). The previous study could not distinguish among the effects of the changes in the FSP during the 1997-2001 period that might have caused these effects, because of data limitations. With only five years of data after 1996, it was not possible to separate temporary effects of PRWORA implementation (likely concentrated in 1997-1999) from more long-run changes and from later changes. Furthermore, while national policies to encourage error reduction were known, systematic longitudinal data on State policy choices for error reduction were limited to what could be gleaned from the QC data (namely the length of the certification period).

The present study explored alternate model specifications and used the additional data for 2002-2005 to provide clearer evidence regarding the relationship of error to certification effort, the percentage of FSP households receiving TANF, and the percentage of cases with short certification periods. The study also used newly available longitudinal data on State policy options to test whether those policies accounted for some of the observed results of the previous study. The specific policy options of interest are discussed in the context of the third hypothesis (in the next section of this chapter) and in the further discussion of analysis variables in Chapter Five.

One possible reason for the previously observed reduction in the response of error to effort is that rules used in determining certification costs changed in 1997. The total certification cost for a State is the sum of the cost of certifying FSP-only households and the FSP's allocated share of certification costs for FSP households receiving other State-administered benefits, such as AFDC or TANF and General Assistance. The allocated FSP share is determined by each State's cost allocation plan, which may vary from other States but must meet fairness standards and receive Federal approval. Under cost allocation rules that applied before PRWORA, shared costs for FSP/AFDC cases were usually allocated to the AFDC program as the "primary program". Under Public Law 105-185, enacted in 1997, States were required to prorate these costs between TANF and the FSP. Thus, a State's FSP certification cost (as reported) and its EFFORT measure could increase even if its total certification cost and caseload for all programs did not change.[35] This study tested for evidence of the effects of this discontinuity, as discussed in Chapter Five.

Hypothesis 3: FSP Simplification Policies Reduced Error Rates Beginning in 2000

In particular, the study examined the effects of the following State policy options on the level of error rates (α), in addition to the effects of short certification periods:

- Quarterly reporting of income (instead of monthly reporting or change reporting), as authorized in 1999
- Simplified reporting (as previously described), first authorized for FY2001 and expanded by the 2002 Farm Bill
- Simplified definition of income, excluding income not counted under the State's TANF or Medicaid program, authorized by the 2002 Farm Bill
- Transitional benefits for TANF leavers so that they are not required to recertify for up to 5 months, first authorized by the 2002 Farm Bill.

These options were selected because they varied substantially among the States and over time; they were expected to reduce the likelihood of error; they were expected to simplify certification; and State-level data on the use of these options were available for the period. Other policy options were

[35] TANF replaced the open-ended matching of administrative costs and benefits with a block grant. Thus, there was, in theory, an incentive for States to shift costs away from TANF and toward the FSP, which remained an entitlement. Concerns about this possibility were a factor behind the cost allocation provision in P.L. 105-185 (Carmody and Dean, 1998). No evidence of such a shift was gathered for this study, but if it had occurred, it would have affected the comparability of certification costs between the pre-PRWORA and post-PRWORA periods.

not included because they did not meet one of these three criteria.[36] In addition, we did not attempt to model the effects of variations in policies requiring the reporting of changes in household composition or income ("change reporting"). Following the example of Ratcliffe, McKernan and Finegold (2007), we treated change reporting and monthly reporting as the base case (i.e., the excluded category) in defining variables for reporting policies, so that the effects of quarterly and simplified reporting could be identified.[37,38]

[36] For example, PRWORA authorized "comparable disqualification" of FSP recipients for TANF violations, but data for this policy were not available for 1998-2001. States were allowed to broaden the criteria of categorical eligibility for TANF recipients, but data on this option were also incomplete. Use of biometric identification was documented, but too few States used this option for the effect to be measurable.

[37] The effects of change reporting policies would be difficult to estimate because they vary in dollar thresholds for reporting and other factors. Thus, there are relatively few observations of state-years with a given version of this type of policy. Instead, States with these policies are part of the excluded group.

[38] Available data on monthly reporting are incomplete, particularly before 1996, so States with this policy regime are also part of the excluded group.

Chapter Five: Model Estimation and Results

This chapter describes the modeling of the error index as a function of FSP certification effort per household, State policies, FSP household characteristics, and other factors. The chapter begins with a discussion of the dependent variable and the regressors. Next, the base model specification and results are presented, followed by the alternate specifications and their results. The chapter concludes with a summary of the findings regarding the hypotheses in Chapter Four. Supplementary technical information on the modeling is presented in Appendix B.

Data

This analysis used a panel of the 50 States and the District of Columbia over the 17 years from 1989 to 2005, in order to examine the relationship between administrative effort per food stamp household and food stamp error rates. The variables and their expected effects are described below.

Dependent Variable

The dependent variable used in this analysis was the error index defined in Chapter Four as a weighted sum of annual positive and negative case error rates. Because it is based on case error rates and includes negative action errors, the error index differs from the combined payment error rate that is the focus of FNS policy and quality control (QC) sanctions.

The error index, ERROR, was calculated as:

$$ERROR = ERROR_p + \lambda ERROR_n \qquad (5\text{-}1)$$

where $ERROR_p$ is the positive error rate, $ERROR_n$ is the negative error rate, and λ is a parameter representing the relative difficulty of eliminating negative versus positive errors, estimated via grid search.[39] The mean State positive error rate during the period 1989-2001 was 9.2 percent and the mean negative error rate was 4.2 percent.[40] The error index depends on the value of λ. We introduced the parameter λ into the model because while spending more on the administration of food stamps was expected to reduce errors, positive errors may be more or less difficult to affect than negative errors. We estimated that $\lambda = 1.42$ with a standard error of 0.15, implying that the amount of resources required to reduce the positive error rate by 1 percentage point would reduce the negative error rate by 0.70 (=1/1.42) percentage points (assuming that only one rate changes at a time).[41] For example, assume a simple model such that ERROR=A-.001(RESOURCES). Thus, an additional 10 units of resources would reduce ERROR by .01. If $ERROR_n$ stays constant, $ERROR_p$ is reduced by 1 percentage point; if $ERROR_p$ stays constant, $ERROR_n$ is reduced by 0.70 percentage points.

---

39 In the grid search, we iteratively tested regression models using alternate values of λ to compute the dependent variable and selected the value that maximized the log-likelihood. The grid search process and the log-likelihood function are described in Appendix B.

40 We remind the reader that these do not correspond to published rates, because of our inclusion of negative actions and our use of a common denominator for all types of error.

41 The previous study estimated $\lambda = 1.45$ with a standard error of 0.16.

Conversely, the amount of resources required to reduce negative error by one percentage point are 1.42 times the resources required to reduce positive error by the same amount (again holding one error rate constant while the other changed).[42] The mean of the error index for the study period was .151 (15.1 percent). Tests of the base model using the unweighted sum of positive and negative error rates yielded highly consistent results, as did separate models of positive and negative error rates.

Effort

The main focus of this analysis was the impact of a change in certification-related effort (the variable EFFORT in preceding discussions) on the error index. For each State in each year, EFFORT was computed by dividing the certification-related cost per FSP household by the average annual pay per full-time public welfare worker, as described in Chapter Three. Thus, EFFORT was a proxy for the ratio of full-time food stamp workers to food stamp households. The mean value of EFFORT for the study period was 0.010 (i.e., one unit of effort per 100 food stamp households).[43] The square of EFFORT, *EFFORT_SQ*, was included in the model to test for non-linear effects on error.

TANF and Welfare Reform

Three of the control variables represented potential influences on error rates arising from the fact that food stamp agencies also administer welfare program. The first variable described below is a key caseload characteristic, while the next two are policy variables. These variables were also expected to interact with EFFORT, as discussed below.

PCT_TANF is the percentage of food stamp households receiving Temporary Aid for Needy Families (TANF) or its predecessor, Aid to Families with Dependent Children (AFDC). In the previous study, we hypothesized that these households would be less error-prone than other food stamp households, because the food stamp agency is required to have authoritative information on AFDC/TANF benefits, and because earnings make up a smaller portion of income than for non-AFDC/TANF households. However, further consideration suggested the possibility of an opposite effect. When a high percentage of food stamp households receive AFDC or TANF, workers must more often apply both AFDC/TANF and FSP rules to the same households, thus increasing the likelihood of error due to confusion about applicable rules. The mean percentage of food stamp households receiving AFDC/TANF was 28.7 percent during the study period.

TANFIMP is an indicator for the years when the State operated the TANF program in place of the former AFDC program.[44] This variable was defined as the percentage of months during the year when TANF was fully operational (thus ranging from 0 in years before TANF implementation to 1 in years when the State operated a TANF program for the entire year). According to these data, the date of TANF implementation varied among States, with the earliest date in late 1996 and the latest date in

---

[42] In exploratory analysis, the estimate of λ was robust to choice of model specification, i.e., its value did not change materially when variables were added to or subtracted from the model. The parameter and its standard error were estimated for the fixed effects model described first.

[43] An alternate definition of EFFORT was tested, adding ADP costs. This alternative definition did not yield significantly different estimates for EFFORT or other variables. Model specifications with ADP cost per household as a regressor showed no effect of this variable on the error index.

[44] The data for TANFIMP were obtained from the data set used by Kabbani and Wilde (2003). We thank ERS for providing these data.

early 1999. We considered two ways that implementation of TANF could affect the level of FSP errors independently of EFFORT (i.e., contributing to the series of α variables in equation 4-2 that affect the intercept of the regression line of ERROR as a function of EFFORT). On the one hand, the implementation of TANF had the potential to destabilize operations in local welfare offices, because workers might be more focused on learning new rules and on getting clients to understand the new program. Furthermore, the new rules for eligibility and benefit determination adopted by States for their TANF programs introduced new differences between cash assistance and Food Stamp Program rules, so that certification and case management for FSP/public assistance households became more complicated and thus more error-prone. On the other hand, FSP/TANF agencies reinvented themselves during the period of TANF implementation, and these changes might have improved the overall effectiveness of caseworkers, including the accuracy of FSP certification.

AFDCWAIV is an indicator for the presence of a welfare reform waiver to AFDC rules. These waivers allowed individual States to implement changes that were authorized for all States in 1996 by PRWORA. Thus, the possible effects of this variable on the error index were similar to those of TANFIMP. The AFDCWAIV variable was defined as the percentage of months in the year when the AFDC waiver was in effect; the overall mean was .068. AFDC waivers first appeared in 1993 and last in 1998 (waivers could operate where TANF was not fully implemented).

NOTI_TANF_EFFORT and *TANFIMP_TANF_EFFORT* represented the interaction of EFFORT with PCT_TANF (the percentage of household receiving AFDC/TANF), respectively, for the period before TANF implementation and the period when TANF operated. TANFIMP_TANF_EFFORT was computed as the product of the TANFIMP, PCT_TANF, and EFFORT variables. NOTI_TANF_EFFORT was the product of (1-TANFIMP), PCT_TANF, and EFFORT. The creation of these variables was based on two expectations. First, the level of error with a given level of effort was expected to be different in a State with a high percentage of FSP households receiving AFDC/TANF than in a State with a low percentage of such households. On the one hand, the sharing of costs between the FSP and the AFDC/TANF program might reduce the level of error at a given level of cost (and thus effort) allocated to the FSP for an FSP household receiving AFDC/TANF. On the other hand, generic workers serving the FSP and AFDC/TANF might be less efficient than specialized FSP-only workers, leading to an expectation that a given level of effort would have less impact on error when the percentage receiving AFDC/TANF is high. However, there was a second expectation: given the differences between AFDC and TANF, the effect of the interaction of PCT_TANF and EFFORT might have changed with the implementation of TANF. In particular, we expected that sharing caseworkers between the FSP and TANF would have more of an impact on efficiency than sharing between the FSP and AFDC. Therefore, the separate variables for interactions under AFDC (NOTI_TANF_EFFORT) and under TANF (TANFIMP_TANF_EFFORT) were created.

FSP Policy Variables

As discussed in Chapter Four, one of the goals of the study was to estimate the effects of FSP policies on the error index. These policy variables and their expected effects are described below.

CM13 is the percentage of food stamp cases with certification periods of one to three months. Short certification periods were expected to reduce error rates because more frequent reviews of eligibility. The authors' previous study found that this variable had a highly significant negative effect on the error index. The mean percentage of food stamp households with short certification periods was 0.095 (i.e., 9.5%).

CM13_EFFORT is the product of CM13 and EFFORT. This interaction was included because, when a State conducts more frequent recertifications, it must increase the effort devoted to this process and may reduce the effort devoted to other certification activities. Depending on the relative reduction in error from different activities, this shift could increase or decrease the overall impact of a given level of effort on error. The authors' previous study did not include this interaction. The mean of this variable was .001.

A related variable, *CM1299ER*, was defined as the percentage of households with earnings having certification periods of 12 months or more. Long certification periods were expected to increase the probability of error. Support for this expectation came from the findings of Kabbani and Wilde (2003), who found suggestive but not conclusive evidence that the combined payment error rate was higher when at least 75 percent of FSP households with earnings had certification periods of 12 months or more. Kabbani and Wilde (2003) treated the 75 percent indicator variable as a proxy for monthly reporting,[45] but a high percentage of long certification periods for earners could occur in States using change reporting or (after 2002) simplified reporting.[46] We treat CM1299ER simply as an indicator of long certification periods for households with earnings. The mean of this variable was 0.424 (i.e., 42.4%).

PCTEBT is the percentage of food stamp households that receive benefits via electronic benefits transfer (EBT). Under the coupon issuance system, food stamp case workers dealt with replacement of lost or stolen coupons, but comparable functions under EBT are mainly handled by separate customer service centers. Therefore, greater use of electronic benefits transfer, relative to coupons, was expected to reduce interruptions that might contribute to case worker error. In addition, some research suggested that EBT increased FSP participation and thus might have changed the composition of the caseload (Kornfeld, 2002). The first statewide implementation of EBT was completed in 1993; by 2004 all States had implemented EBT. Over the 17 study years, the mean percentage of food stamp households receiving electronic benefits was 0.402 (i.e., 40.2%).[47] This variable did not have a significant association with the error index in the authors' previous study, but it was included because of the theoretical rationale and to help make the results comparable across studies.

PCT_QUARTREP is an indicator for when a State used the quarterly reporting option. The variable is equal to the percentage of months in the year when this option was in effect. Under this option, households are required to report income and other key data on a quarterly basis, instead of reporting monthly or whenever changes exceeding specified thresholds occur ("change reporting"). This option

---

45 The FSP Rules Database has a variable for presence of monthly reporting, but there is a substantial period for which the data are incomplete, so we did not attempt to use this variable

46 Under change reporting, households are required to report when their circumstances change (within specified parameters), rather than reporting monthly regardless of whether they have changes. Simplified reporting can be implemented with recertification every six months, or a required report at six months and a twelve-month certification period.

47 In most years, PCTEBT was either 0 or 1. Values between 0 and 1 occurred during the transition from coupon to EBT issuance, and when certain States issued a portion of benefits in cash under special waivers.

was first offered in 1999 as a way to simplify the FSP for participants and agencies. It was expected to reduce the error index because unreported changes during the quarter do not give rise to errors, and because it allows agencies to shift effort from processing monthly or change reports to other activities that may be more effective at reducing error. The mean of PCT_QUARTREP for the study period was 0.052 (i.e., 5.2%); thus the use of this option was infrequent.

PCT_SIMPREP is an indicator for when a State used the simplified reporting option. The variable is equal to the percentage of months in the year when this option was in effect. Under this option, unreported changes in income or deductions during the six-month reporting period do not create errors unless the household's gross income exceeds the limit for FSP eligibility. Simplified reporting became an option for households with earnings under FNS waiver authority in 2001, and the 2002 Farm Bill expanded this option to include all FSP households. As described in Chapter Four, previous research estimated that simplified reporting reduced the combined overpayment rate (a different measure of error) by 1.2 to 1.5 percentage points, based on applying State options as of 2004 to case data for 2000. When simplified reporting became available, the utilization of the quarterly reporting option declined. The mean of PCT_SIMPREP for the entire study period was 0.155 (i.e., 15.5%).

PCT_TRANBEN is an indicator for when a State opted to provide transitional benefits for households leaving the TANF program. Under this option, the household is not required to recertify for up to 5 months, and any change in income or other eligibility factors is disregarded. This option was first authorized in 2001 by a waiver for New York; the 2002 Farm Bill authorized transitional benefits for all States. The variable is equal to the percentage of months in the year when this option was in effect. The mean for the entire period was 0.042 (i.e., 4.2%).

SIMPINCOME is an indicator for when a State used an option that simplified income computation by excluding income not counted under the State's TANF or Medicaid program. This option was authorized by the 2002 Farm Bill and thus appears in 2003 through 2005. The variable is equal to the percentage of months in the year when this option was in effect. The mean for the entire period was 0.074 (i.e., 7.4%).

Additional Caseload Characteristics and Socioeconomic Conditions

Although the main focus of this analysis was on the impact of effort and policies on the error index, we included other observable covariates in the analysis for two reasons. First, the probability of error may vary for different case types or under different economic conditions, independently of the level of certification-related effort. Second, the level of effort required to process cases with a given level of accuracy may vary by case type or program conditions. Therefore, we introduced control variables intended to control for time-varying differences in caseload characteristics and program conditions.

EARNINC is the percentage of food stamp households with earned income in their case records.[48] Households with reported earnings are likely to have more volatile income and thus be more prone to underpayment or overpayment error. In our sample, the mean percentage of food stamp households with earned income was 25.6 percent. This variable had a significant positive association with the error index in the authors' previous study.

---

[48] This variable does not count households with only unreported earnings, so it understates the proportion of households with the potential for erroneous information on earnings.

SSINC is the percentage of food stamp households with Social Security Old Age, Disability, and Survivors Insurance (OASDI) or Supplemental Security Income (SSI) benefits. These households were expected to be less error-prone than other food stamp households because the Food Stamp program can easily and definitively verify OASDI and SSI benefits through well-established data exchange systems. These households were also expected to be less likely than others to have unreported earnings, a potential source of error that is not captured by the reported earnings indicator. The mean was 38.1 percent of food stamp households receiving OASDI or SSI benefits.

SINGLEPAR is the percentage of food stamp households with children headed by a single adult. Among households with earnings, single-parent households were expected to be less error-prone than two-parent households, which may be more likely to have unreported earnings because they have two potential earners. There is some risk, however, that single-parent households are more likely than two-parent households to have unreported child support income. The mean percentage of food stamp households with one or more children that had a single adult was 72.2 percent.

FYUN0 is the state-specific annual average unemployment rate. In periods of high unemployment, the work hours and earnings of employed food stamp recipients may be more volatile, therefore creating more potential for errors. (EARNINC controlled for the percentage of food stamp households with earnings.) The mean State unemployment rate from 1989 to 2005 was 5.2 percent.[49]

PCT_CHCASELOAD is the percentage difference between the State's FSP caseload for the fiscal year and the caseload for the prior year (using the annual average of monthly household counts as the caseload measure). As described in Chapter Two, the national average positive and negative error rates were positively associated with the caseload. Our primary hypothesis about this correlation was that it reflected shifting levels of EFFORT. The PCT_CHCASELOAD variable controlled for other ways in which caseload trends might affect ERROR, such as a change in the mix of initial certifications versus recertifications.

PCTBLACK is the percentage of the State population classified as black or African-American, *PCTHISP* is the percentage of the State population classified as Hispanic or Latino, and *PCTMET* is the percentage of the State population in metropolitan areas, derived from Current Population Survey data. These variables appeared in the models of error rates and FSP participation used by Kabbani and Wilde (2003) as controls for the possible effects of demographic differences among States that, if omitted, might bias the estimated effects of policy variables. We included these variables in our models for the same reason, not because we expected the probability of error to be directly related to race, ethnicity, or urbanization. These demographic variables also preclude a possible source of bias in the estimated effect of EFFORT.

The means and standards deviations of the dependent variable and the independent variables, along with their definitions and sources, are found in table 5-1. Variables that were used in the authors' previous study are noted.

---

49 A related variable, the change in the unemployment rate, was also used in alternate specifications. The results were similar to those presented here. An increase in the unemployment rate could increase error rates, because there would be more first-time, short-term food stamp recipients. These recipients might be more prone to error because of having no history of dealings with welfare workers. The change in unemployment rate was not significant when included with the unemployment rate in the model.

Table 5-1

Analysis Variables, Definitions, Means, Standard Deviations, and Data Sources

Variable	Definition	Mean	Std. Dev.	Source
ERROR[a]	Error index computed as the weighted sum of positive and negative case error rates (as defined in Chapter Two)	0.151	0.048	QC & FNS reports
EFFORT[a]	Certification-related cost per FSP household, normalized by the state wage for a full-time public welfare worker	0.010	0.004	SF-269/ Census
EFFORT_SQ	Square of EFFORT (used to test for diminishing or increasing returns to effort)	1.20E-04	8.69E-05	Derived
PCT_TANF[a]	Percent of food stamp households receiving AFDC or TANF	0.287	0.138	QC
TANFIMP	Proportion of months in the year operating TANF program. Ranges from 0 (all states, 1989-1996) to 1 (all States in 1999-2005)	0.515	0.493	K&W
AFDCWAIV	AFDC reform waiver (percent of months in year)	0.068	0.234	K&W
NOTI_TANF_EFFORT	Interaction between (1-TANFIMP), PCT_TANF, and EFFORT	0.002	0.002	Derived
TANFIMP_TANF_EFFORT	Interaction between TANFIMP, PCT_TANF, and EFFORT	0.001	0.002	Derived
CM13[a]	Percent of food stamp households with 1-3 month certification periods	0.095	0.126	QC
CM13_EFFORT	Interaction of CM13 and EFFORT	0.001	0.002	Derived
CM1299ER	Percent of food stamp households with earnings having certification period of 12+ months	0.424	0.365	QC
PCTEBT[a]	Percent of food stamp households that receive electronic benefits	0.402	0.465	FNS data bank
PCT_QUARTREP	State has quarterly reporting for earners (percent of months in year)	0.052	0.219	FSP rules
PCT_SIMPREP	State has simplified reporting for earners (percent of months in year)	0.155	0.355	FSP rules
PCT_TRANBEN	State has transitional FS benefit for TANF leavers (percent of months in year)	0.042	0.198	FSP rules
SIMPINCOME	State uses simplified definition of income	0.074	0.259	FNS Options
EARNINC[a]	Percent of food stamp households with earned income in case record	0.256	0.079	QC
SSINC[a]	Percent of food stamp households with OASDI or SSI benefits	0.381	0.102	QC
SINGLEPAR[a]	Percent of food stamp households with children headed by a single adult	0.722	0.084	QC
FYUN0[a]	Unemployment rate	0.052	0.014	BLS
PCT_CHCASELOAD	Percent change in average number of participating households from previous year	0.025	0.099	FNS data bank
PCTBLACK	Percent of State population black	0.109	0.114	CPS/K&W
PCTHISP	Percent of State population Hispanic	0.068	0.084	CPS/K&W
PCTMET	Percent of State population metropolitan	0.697	0.202	CPS/K&W

Table 5-1

Analysis Variables, Definitions, Means, Standard Deviations, and Data Sources

[a] Variables used in the previous study (Logan, Rhodes, and Sabia, 2006).

Sources:

QC: Tabulations from Quality Control public use files.

SF-269/Census: FSP administrative outlays as reported by States on SF-269 reports, provided from FNS administrative database. Welfare worker wage from Census of State and Local Governments.

K&W: Dataset created by N. Kabbani and P. Wilde and used for their 2003 article; covers 1990-2000. For later years, all states set to 1 for TANFIMP and 0 for AFDCWAIV.

FSP rules: Food Stamp Program Rules Database, compiled by the Urban Institute. Data for 2005 taken from Food Stamp Program Options Report, Fifth Edition.

FNS Options: Food Stamp Program State Options Report prepared by FNS, data for 2001-2005.

BLS: Bureau of Labor Statistics.

FNS data bank: Annual average counts of participating FSP households based on State reports to FNS, data taken from FNS national data bank.

CPS: Current Population Survey.

Variables not listed: State fixed effects, State time trends, and interactions with EFFORT used in alternate models (described in text).

Estimation of Models of the Food Stamp Error Index

We began with a base model to test the association between effort and the food stamp error index using a simple model that included fixed effects for each State and used robust standard errors to control for autocorrelation across time periods. Below, we describe the base model and the results it yielded. Derivations of equations and other details are provided in Appendix B. We then summarize the results of alternative specifications and models, which are described in Appendix C. As discussed below, we preferred the base model to the alternative specifications, so the discussion in this chapter focuses on the base model and its implications.

The Base Fixed Effects Model

Following standard practice for state panel data, we used a fixed effects model to estimate equation (6):

$$ERROR_{it} = \alpha_0 + \alpha_i + t_i\delta_i + f(EFFORT_{j,it})\beta_j + X_{j,it}\gamma_j + e_{it} \qquad (5\text{-}2)$$

where $ERROR_{it}$ is the error index in State *i* at time *t*, α_0 is the intercept, α_i is a time-invariant state-effect, t_i is a state-specific linear time trend, δ_i is the state-specific coefficient on that linear time trend, the β_j are parameter estimates for various specifications of *f(EFFORT)*, X_{it} is a row vector of control variables, and γ_j is a column vector of parameters conformable with *X*. The base model did not include state-invariant time effects (i.e., year dummy variables), but a model with these effects was tested as described later in this chapter.

The linear time-invariant state effect, α_i, controls for unmeasured static factors that vary across States. For example, urban States may have higher error rates than rural States, in which case α_i would control for urbanicity.[50] If those state effects were excluded and the omitted variables were correlated with *EFFORT*, the estimates of β_j would be biased and inconsistent. The time effect, δ_i, controls for state-specific unmeasured factors that have linear trends over time. For example, unmeasured improvements in data processing technology may affect effort rates over time, and linear time trends take this effect into account. With these two types of fixed effects in the model, the remaining control variables account for variation in error rates attributable to within-state changes in measured factors. Differences among States that do not vary over time are already taken into account by α_i. Additionally, factors that vary linearly over time within each state are already taken into account by δ_i. The parameters associated with the control variables capture whatever remaining partial correlation exists between error rates and the measured variables.[51] In Appendix C, we present results for an alternative specification that includes a full set of dummy variables for each year. Results from that specification are qualitatively similar to those for the base specification discussed here in the body of the report.

We recognized that e_{it} were likely correlated within State. Bertrand, et al. (2004) suggests that standard errors are not consistent without taking this correlation into account. Therefore, we used the generalized estimating equations (GEE) suggested by Liang and Zeger (1986), using exchangeable correlations within States. We assumed that e_{it} were block-diagonal across States. If there is unspecified correlation within a State, we would expect the standard errors to increase. As a test, we also ran an ordinary least-squares (OLS) model of our base specification, which yielded standard errors for parameter estimates that were roughly 30 percent smaller than the corresponding GEE standard errors.

Base Model Specification

For the base model, the regressors for *f(EFFORT)* were the following variables (defined in the previous section except as noted):

- EFFORT
- EFFORT_SQ – The square of EFFORT
- TANFIMP_TANF_EFFORT
- NoTI_TANF_EFFORT
- CM13_EFFORT

This specification of *f(EFFORT)* is presented as the base model because it includes all of the hypothesized effects without further elaboration. Models presented subsequently use alternative specifications of *f(EFFORT)*.

The $X_{j,it}$ for the base model are identified in table 5-2, which provides the variable means and parameter estimates for the base model. The $X_{j,it}$ are the same for all specifications in this chapter,

---

50 In a cross-sectional analysis, Puma and Hoaglin (1987) found that the incidence and amount of overpayments were positively related to a State's population density. This variable was not expected to vary greatly over time within a State, so it was not included as a separate variable in the models.

51 In the prior analysis, we considered an alternate specification with random state effects, which would have the advantage of being more efficient if the null hypothesis of no systematic difference between fixed and random effects coefficients were true. We ran a Hausman test, however, and rejected the null hypothesis at $p < 0.001$. Thus, a fixed effects specification appears more appropriate.

Table 5-2

Generalized Estimating Equations (GEE) Fixed Effects Parameter Estimates: Base Model

Variable	Means	Estimated Parameters (with Std. Errors)
INTERCEPT		0.177
		(0.228)
EFFORT	0.010	-7.979***
		(2.737)
EFFORT_SQ	1.20E-04	63.309
		(82.268)
TANFIMP_TANF_EFFORT	0.001	17.811***
		(3.989)
NOTI_TANF_EFFORT	0.002	2.640
		(4.515)
CM13_EFFORT	0.001	7.990***
		(2.832)
TANFIMP	0.515	-0.019*
		(0.010)
PCT_TANF	0.287	-0.037
		(0.065)
CM13	0.095	-0.184***
		(0.042)
CM1299ER	0.424	0.001
		(0.010)
PCT_SIMPREP	0.155	-0.036***
		(0.007)
PCT_QUARTREP	0.052	-0.012*
		(0.007)
PCT_TRANBEN	0.042	-0.027***
		(0.010)
AFDCWAIV	0.068	0.011**
		(0.006)
SIMPINCOME	0.074	-0.009
		(0.008)
EARNINC	0.256	0.225***
		(0.060)
SSINC	0.381	-0.009
		(0.052)
SINGLEPAR	0.722	-0.010
		(0.043)
PCTEBT	0.402	-0.003
		(0.006)
FYUN0	0.052	0.123
		(0.171)
PCT_CHCASELOAD	0.025	0.016
		(0.010)
PCT_MET	0.697	-0.022
		(0.047)
PCTBLACK	0.109	0.115
		(0.296)
PCTHISP	0.068	-0.382
		(0.490)
R-squared		0.7412

except as noted. (See the preceding section on data for variable definitions and rationales for their use.)

Results from the Base Model

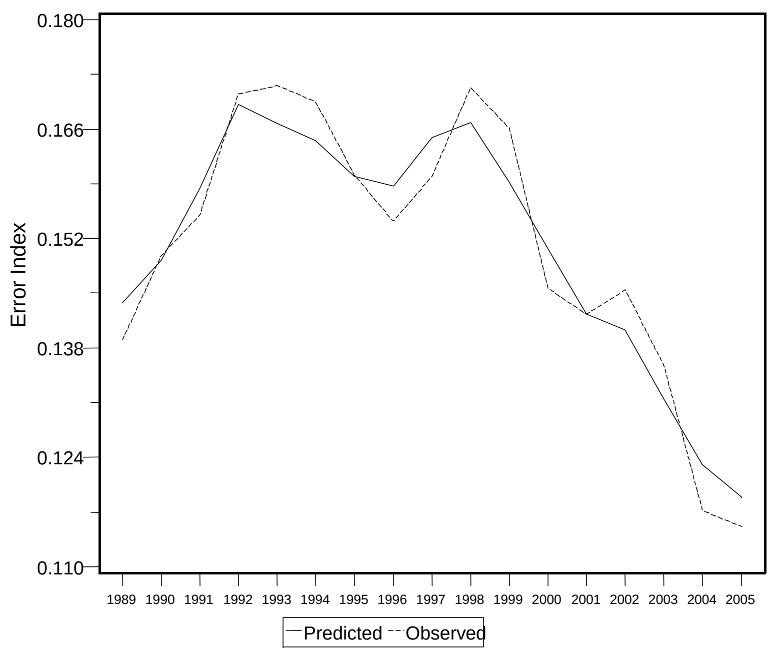
Results of the coefficients associated with *f(EFFORT)* and X_{it} from the base GEE model are reported in the last column of table 5-2. As an illustration of the fit of the model to the data, figure 5-1 graphs the predicted and observed mean error index over time.

Several of the $X_{j,it}$ had significant associations ($p<0.05$) with the error index. Among the policy variables, the error index declined with increases in the rate of short certification periods (CM13), and it was lower in the presence of simplified reporting (PCT_SIMPREP) and transitional benefits for TANF leavers (PCT_TRANBEN). States using an AFDC waiver (AFDCWAIV) had a higher error rate. Considering the caseload characteristics and socioeconomic conditions, only one variable had a significant association with the error index: there were more errors when more FSP households had reported earnings (EARNINC). The associations with CM13 and EARNINC were consistent with the findings of the authors' previous study.

The relationship of the error index to EFFORT comprises several variables, some with significant associations with the error index in the base model estimates. As expected, the EFFORT variable itself was negatively associated with the error index. As a test of diminishing returns of effort on error rates, we included the square of the EFFORT variable. We found no evidence of diminishing returns. The interaction of percentage of States with TANF implementation, the percentage of FSP

Figure 5-1

Predicted and Observed Error Index over Time: GEE Fixed Effects Model with Base Specification



Note: All variables evaluated at their yearly means households receiving TANF (PCT_TANF), and EFFORT (TANFIMP_TANF_EFFORT) was positively associated with the error index. The interaction of "no TANF implementation" (i.e., 1 minus TANFIMP), the percentage of households receiving TANF, and EFFORT (identified as NoTI_TANF_EFFORT) did not appear to be associated with the error index. Finally, the interaction of the percentage of households with 1-3 month certification periods and EFFORT (CM13_EFFORT) was positively associated with the error index.

Figures 5-2, 5-3, and 5-4 illustrate the relationship of EFFORT, PCT_TANF, and CM13 to ERROR. Each figure shows the predicted ERROR as one variable changes while others are held at their means for the time period.

Figure 5-2 plots the predicted ERROR as a function of EFFORT for two periods: 1989-1996 (before PRWORA), and 1997-2005 (after enactment of PRWORA). Both periods have negative slopes, but the pre-PRWORA slope is steeper, indicating higher levels of ERROR at low levels of EFFORT, and lower levels of ERROR at high levels of EFFORT. The change in slope resulted from the positive interactions with EFFORT (TANFIMP_TANF_EFFORT and CM13_EFFORT) that offset the negative effect of EFFORT alone on ERROR. TANFIMP_TANF_EFFORT had a much larger estimated effect than NOTI_TANF_EFFORT (the pre-PRWORA counterpart), and the effect of CM13_EFFORT was larger in the post-PRWORA period because the average value of CM13 was larger. Thus, the interaction of EFFORT with the percentage of FSP households receiving TANF and the percentage with one to three-month certification periods were the reasons for the reduced net effect of EFFORT on ERROR in the 1997-2005 period.

Figure 5-2

Predicted Error Index over Effort, 1989-1996 and 1997-2005

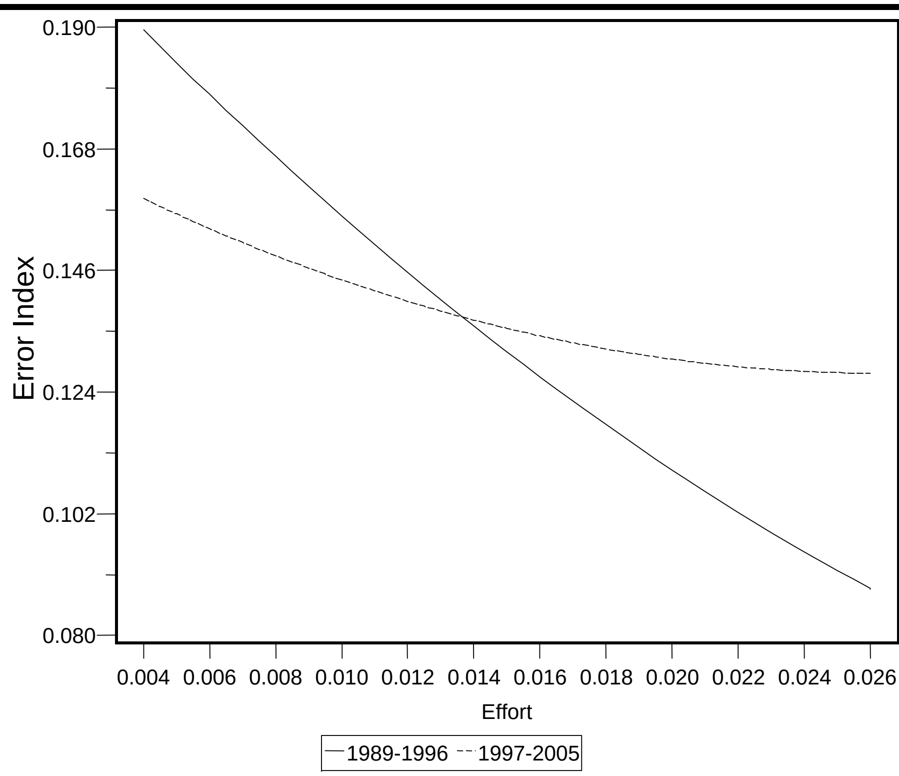
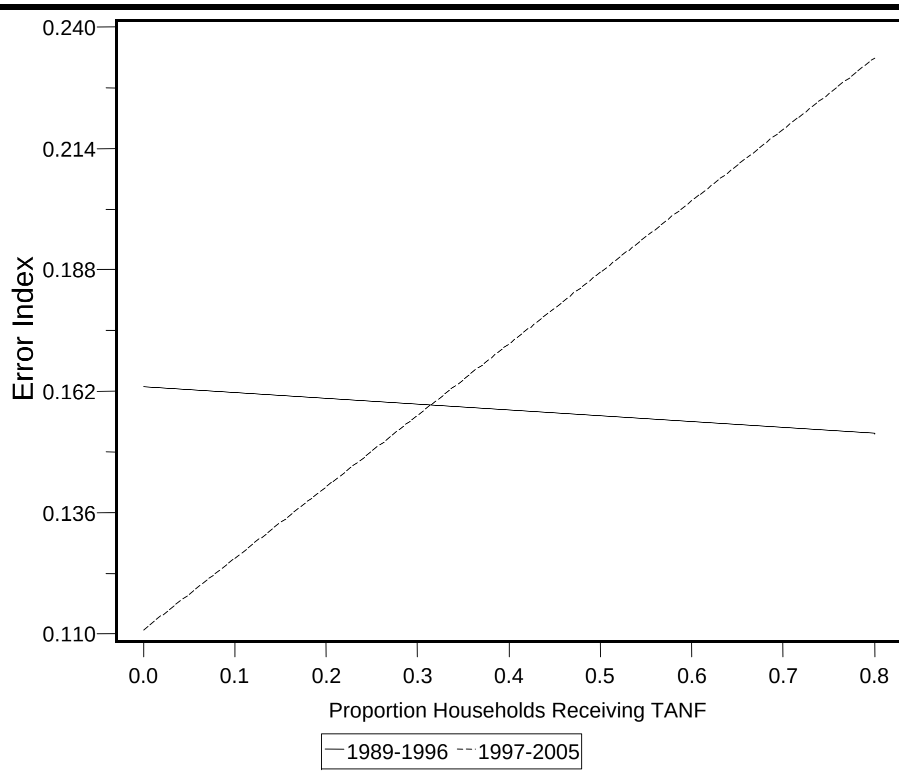


Figure 5-3 shows how ERROR increased with the percentage of FSP households receiving TANF (PCT_TANF) in 1997-2005, contrasted with the nearly flat relationship between ERROR and PCT_TANF in 1989-1996 (under the AFDC program). PCT_TANF had a negative but insignificant effect on ERROR, hence the slightly negative slope of the TANF-ERROR curve for the 1989-1996 period. The reason for the rise in ERROR with PCT_TANF was that the effect of EFFORT on ERROR diminished as PCT_TANF increased with EFFORT in the presence of TANF implementation (i.e., the positive effect of TANFIMP_TANF_EFFORT). Possible reasons for this important effect are discussed later in this report. The relationship of ERROR to PCT_TANF in 1997-2005 was two-sided. On the one hand, States with high percentages of FSP cases receiving TANF had higher error rates; on the other hand, the error rate in a State would fall markedly as the TANF percentage fell. During the 1997-2005 period, an increase of 10 percentage points in the percentage of FSP households with TANF was predicted to raise the error index by an average of 1.6 percentage points.[52]

Figure 5-3

Predicted Error Index over Percent of FSP Households with TANF, 1989-1996 and 1997-2005



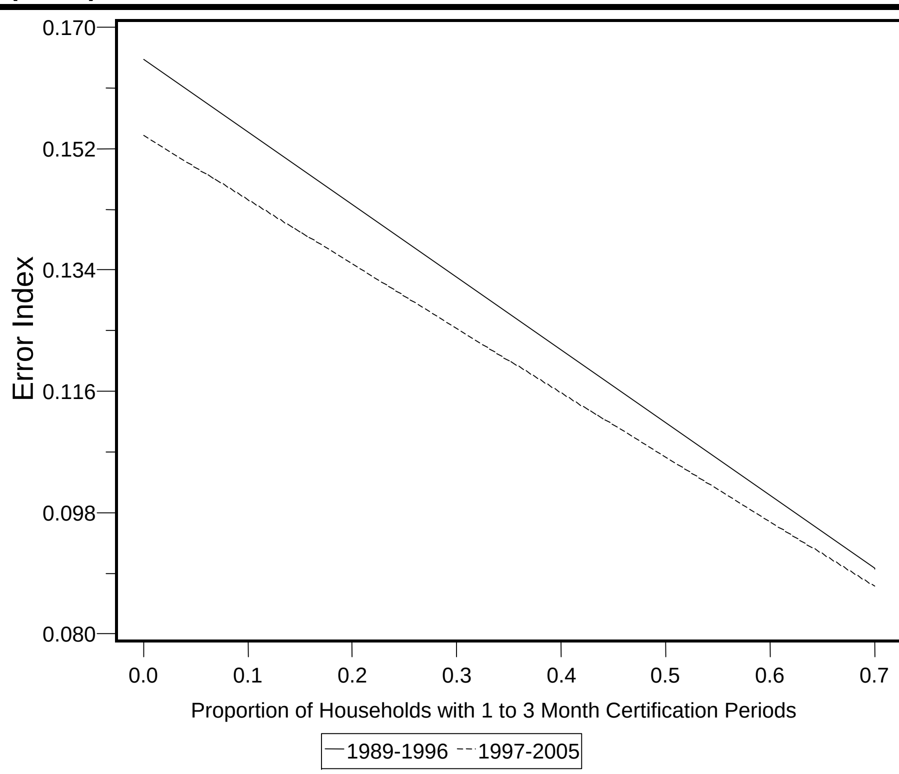
52 A change in PCT_TANF affected ERROR through changes in three variables, each with its own coefficient: PCT_TANF, NOTI_TANF_EFFORT, and TANFIMP_TANF_EFFORT. The estimated effect thus depended on the means of TANFIMP, (1-TANFIMP), and EFFORT, which varied over the period. As a result, the estimated effect of a change of 0.1 in PCT_TANF on ERROR ranged from 0.9 percentage points (i.e., 0.009, for 1997) to 2.0 percentage points (i.e., 0.020, for 2001).

As shown in figure 5-4, ERROR declined as the proportion of cases with short certifications (CM13) increased. The direct negative effect of CM13 on ERROR was greater than the indirect positive effect through the interaction of CM13 with EFFORT. The slope of ERROR as a function of CM13 was almost the same in the two periods; thus, the marginal effect of increasing the use of short certifications was about the same. The lower levels of ERROR at all levels of CM13 in the 1997-2005 period reflect the effects of other variables that shifted the level of error relative to the earlier period. These included the increased level of EFFORT, the reduced percentage of FSP households with AFDC/TANF, and FSP simplification options (most notably simplified reporting and transitional benefits for TANF leavers).

Because of the interactions with EFFORT, it is useful to construct a summary measure of its effect on the error index. We computed the marginal effect of effort on the error index, i.e., the partial derivative of ERROR with respect to EFFORT. As explained in Appendix B, the marginal effect of effort depends on the values of the variables that interact with EFFORT, including PCT_TANF, TANFIMP, CM13, and EFFORT itself (through the squared term, which had a positive effect on ERROR).[53] Therefore, we computed the marginal effect of effort at the yearly means of these

Figure 5-4

Predicted Error Index over Proportion of FSP Households with 1-3 Month Certification Periods (CM13), 1989-1996 and 1997-2005



[53] All variables that interacted with EFFORT were included in computing the marginal effect of effort, even if the variable did not have a significant effect by itself.

variables. The marginal effect represents the number of units that the error index is expected to change with a change of 1 unit in EFFORT. Since the error index is in decimal form, a change of 1 unit in the error index equals 100 percentage points. A change of 1 unit in EFFORT would be equal to 100 times the mean of .01. Below we use the estimated marginal effect to predict the change in error with more plausible changes in EFFORT.

The computed yearly marginal effects, their standard errors, and p-values are reported for the base model in the second column of table 5-3. The table also shows the overall mean marginal effect and the means for three periods: 1989-1996 (pre-PRWORA), 1997-2002 (post-PRWORA, before implementation of the 2002 Farm Bill), and 2003-2005 (after the 2002 Farm Bill). Finally, differences of yearly effects from the overall mean and between the period means are shown with their standard errors and significance levels. The yearly marginal effects and their 95 percent confidence intervals are graphed in figure 5-5. The bars represent the yearly estimates and the lines extending from the bar ends (the "whiskers") represent the confidence intervals.

In the 1989-1996 period, the marginal effect of effort on error varied from year to year by less than a standard error—it was essentially unchanged. The average marginal effect for this period was significantly greater (in absolute value) than the overall average; this was true for each year as well. The marginal effect moved much closer to zero in 1997 and 1998, and then gradually returned to approach pre-TANF levels by 2005. The discontinuity between 1996 and 1997 resulted from the difference between the minimal effect of the interaction of EFFORT and PCT_TANF in the pre-PROWRA period (NOTI_TANF_EFFORT) and the much larger positive effect of this interaction after PRWORA (TI_TANF_EFFORT). In the 1997-2002 period, the marginal effect was significantly smaller (in absolute terms) than in the periods before and after. During this period, the use of short certification periods (CM13) was at its peak for the entire study period, and the percentage of FSP households with TANF (PCT_TANF) was falling but still greater than in 2003-2005. The "recovery" in the marginal effect of effort (i.e., the increase in the absolute value back to pre-1997 levels) began in 1999 and 2000, due to the declining PCT_TANF, which more than offset the effects of increases in EFFORT and CM13 in these years. (Use of short certification periods peaked in 2000.) In 2001-2002, and in 2003-2005, the marginal effect continued to increase in absolute value, due to the continuing decline in PCT_TANF, a reduction in CM13 (as States adopted other approaches to avoid errors in cases with earnings), and a reduction in EFFORT. The marginal effect of effort in 2003-2005 was smaller in absolute value than in 1989-1996, but this difference was not significant.

To aid in understanding the trends in table 5-3 and figure 5-5, table 5-4 shows the contributions of EFFORT and the variables interacting with EFFORT to the marginal effect of effort on error for the three periods. The marginal effect is the sum of the computed effects of the variables that change when EFFORT changes; thus each of these variables has a contribution to the net marginal effect. (The computation of these contributions is explained in Appendix B.) As the table shows, all of the variables that interacted with EFFORT had positive contributions, offsetting the negative contribution of EFFORT by itself and reducing the absolute value of the net marginal effect. The main sources of differences between periods were the contributions of TANFIMP_TANF_EFFORT and

Table 5-3

Marginal Effect of Effort on Error: Base Model Estimates (GEE Fixed Effects)

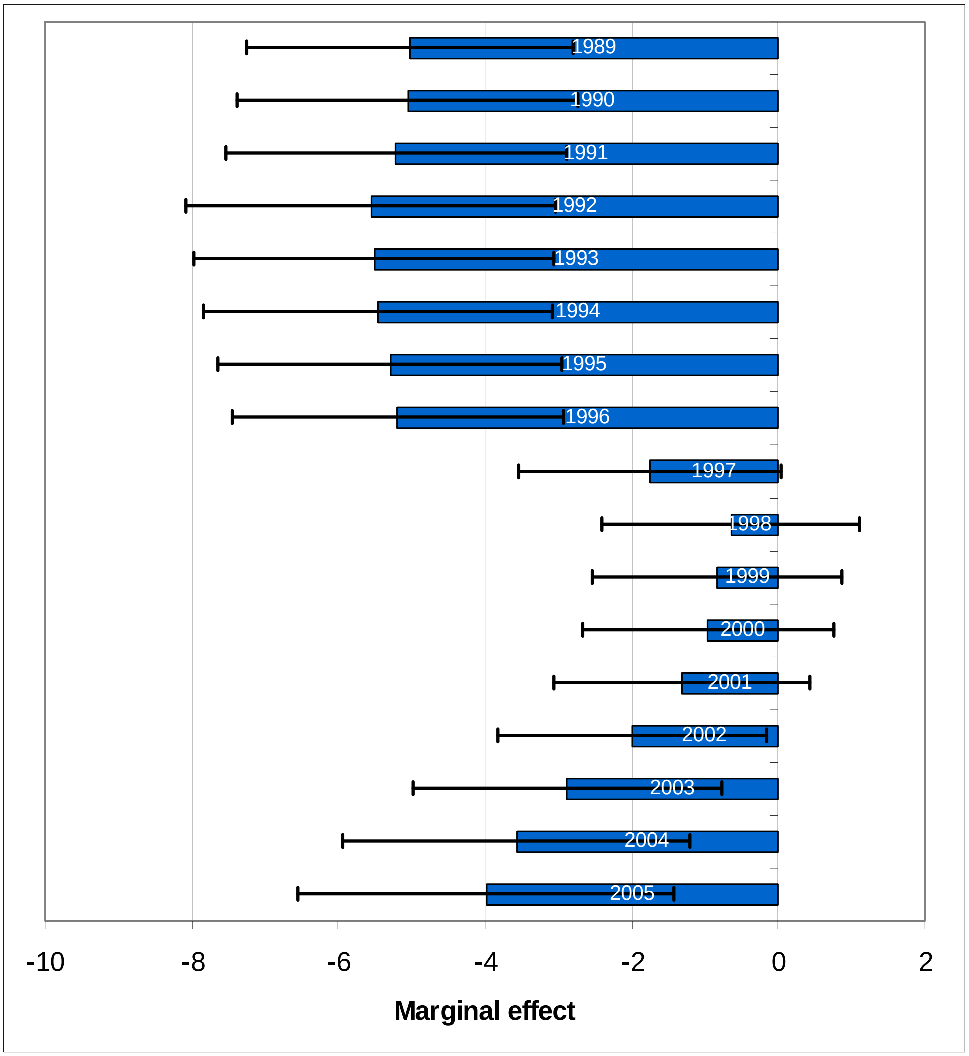
Year	**Marginal Effect**	**Difference from Overall Mean**
Overall Mean	-2.906***	
	(0.902)	
1989	-5.022***	-2.115***
	(1.136)	(0.815)
1990	-5.057***	-2.151***
	(1.181)	(0.805)
1991	-5.218***	-2.312***
	(1.184)	(0.726)
1992	-5.559***	-2.652***
	(1.285)	(0.720)
1993	-5.510***	-2.604***
	(1.251)	(0.696)
1994	-5.463***	-2.557***
	(1.216)	(0.631)
1995	-5.298***	-2.391***
	(1.196)	(0.618)
1996	-5.190***	-2.283***
	(1.148)	(0.538)
1997	-1.747*	1.159***
	(0.915)	(0.186)
1998	-0.646	2.260***
	(0.899)	(0.387)
1999	-0.830	2.077***
	(0.866)	(0.486)
2000	-0.961	1.945***
	(0.877)	(0.588)
2001	-1.317	1.589***
	(0.888)	(0.592)
2002	-2.001**	0.906*
	(0.936)	(0.491)
2003	-2.878***	0.028
	(1.078)	(0.563)
2004	-3.574***	-0.667
	(1.209)	(0.688)
2005	-3.986***	-1.079
	(1.305)	(0.785)

	Marginal Effect	**Difference from 1989-1996**	**Difference from 1997-2002**
1989-1996	-5.327***		-4.139***
	(1.198)		(1.036)
1997-2002	-1.188	4.139***	
	(0.864)	(1.036)	
2003-2005	-3.479***	1.848	-2.291***
	(1.193)	(1.250)	(0.623)

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

Figure 5-5

Marginal Effect of Effort on Error: Base Model Estimates (GEE Fixed Effects)



Note: All variables evaluated at their yearly means

All variables evaluated at their yearly means. Dark lines ("whiskers") extending to the right and left of the ends of the bars mark the upper and lower 95 percent confidence intervals.

Table 5-4

Contribution of Variables to Marginal Effect of Effort on Error

	Period		
	1989-1996	**1997-2002**	**2003-2005**
Contribution of EFFORT to marginal effect	-7.979	-7.979	-7.979
Contribution of EFFORT_SQ to marginal effect	1.206	1.486	1.222
Contribution of TANFIMP_TANF_EFFORT to marginal effect	0.000	4.038	2.773
Contribution of NOTI_TANF_EFFORT to marginal effect	0.950	0.035	0.000
Contribution of CM13_EFFORT to marginal effect	0.496	1.233	0.506
Net Marginal Effect	-5.327	-1.188	-3.479
Additional effort to reduce error index by .01	0.002	0.008	0.003
Annual cost per case to reduce error index by .01[a]	65.96	295.82	100.99
Error reduction for $100 per case per year	-0.015	-0.003	-0.010

[a] Calculated at overall average $35,134 per FTE

NOTI_TANF_EFFORT.[54] Thus, the main reason for the diminished marginal effect of effort on error in 1997-2002 and in 2003-2005 was that, with TANF implementation, there was a change in the relationship of the percentage of FSP households on AFDC/TANF (PCT_TANF) to the effect of effort on error. Without this change, the marginal effect would have been about the same in 2003-2005 as in 1989-1996, since the contributions of the other variables were materially or exactly the same. The contribution of the TANF interaction diminished as the percentage of food stamp households with TANF dropped by more than half, from 31.6 percent in 1997 to 14.2 percent in 2005. The effect of diminishing returns to effort (i.e., contribution of EFFORT_SQ) was slightly larger in 1997-2002 than before or after, reflecting the higher average level of EFFORT during this period. The contribution of short certification periods (CM13_EFFORT) was 1.5 times as large in 1997-2002 as before or after. The percentage of FSP households with short certification periods rose from 4.3 percent in 1992 to 10.9 percent in 1997, jumped to 18.8 percent by 2000, and then fell just as rapidly thereafter, reaching 9.6 percent in 2003 and continuing to fall in 2004-2005.[55]

Table 5-4 also translates the net marginal effect into dollars and cents. First, the table shows the additional effort required to reduce the error index by 1 percentage point (.01) and the real annual cost per food stamp case of this effort (evaluated at the overall mean pay rate per FTE public welfare worker in 2005 dollars). The real cost per case jumped from under $66 in 1989-1996 to almost $296 in 1997-2002, then fell to $101 in 2003-2005. Similarly, an increase in effort costing $100 per case (in 2005 dollars) would have reduced the error index by 1.5 percentage points in 1989-1996, only 0.3 percentage points in 1997-2002, and 1.0 percentage points in 2003-2005. At the overall average marginal effect of -2.906 (shown in table 5.3) and the overall average effort of .01, an increase of 10

[54] The contribution of NOTI_TANF_EFFORT was greater than zero in 1997-2002 because TANF was not fully implemented in all States in 1997 and 1998.)

[55] There was little change in CM13 from 1989 to 1992.

percent in effort (.001 units) would reduce the error index by .3 percentage points, at a real annual cost of just over $35 per household.

Alternative Specifications and Models

We tested the robustness of the base model results in several ways. All of the tests substantially confirmed the results with the base model. We summarize these tests below. The alternate specifications and models, and their results, are presented in Appendix C.

First, we tested the base model using alternative measures of certification error in place of the error index. All of these tests used the base model specification for the control variables. We tested the following error measures:

- overpayment case error rate (percent of cases with overpayments or payment to ineligible household)[56]
- underpayment case error rate (percent of cases with underpayments)
- combined case error rate (unweighted sum of overpayment and underpayment case error rates)
- official overpayment rate (percent of benefits paid in overpayments or to ineligible households)[57]
- official underpayment rate (percent of benefits in underpayments)
- combined payment error rate (unweighted sum of official overpayment and underpayment rates).

All of the variables that were significant (at the 5 percent level or better) in the base model of the error index were also significant and had similar effects in the models of the other combined error rates and one or both of the individual error rates. The trends in the marginal effect of effort on error were also very similar for all the error rates.

Second, we tested two alternative specifications of *f(EFFORT)*, to determine whether TANF implementation (TANFIMP) alone affected the relationship of EFFORT to the error index, or whether the percentage of FSP households with TANF also interacted with EFFORT (through the TANFIMP_TANF_EFFORT variable). One alternative specification removed the TANFIMP_TANF_EFFORT and NoTI_TANF_EFFORT interactions, and instead inserted TANFIMP interacted with EFFORT (TANFIMP_EFFORT), thus framing *f(EFFORT)* as in the author's previous study.[58] The second alternative added the TANFIMP_EFFORT variable to the

---

56 Case error rates were computed using the constant $25 threshold for countable errors, as in the error index.

57 As previously discussed, the threshold for countable errors was $5 prior to 2000 and $25 thereafter. To compensate for this change, we included a dummy variable for years after 2000, expecting that the higher threshold would reduce the error rate. The dummy variable was significant at the .05 level for the underpayment and combined payment error models.

58 In the previous study EFFORT was interacted with an indicator for the years 1997-2001, the period after enactment of PRWORA. TANFIMP changed from 0 to 1 at different times, reflecting the implementation of the legislation.

current study's base model, thus providing a "head to head" test of the variables' effects. These alternatives confirmed that the base model was the preferred specification, underscoring the effect of the percentage of FSP households receiving TANF on the net effect of EFFORT.

Third, we tested alternative specifications with additional policy options that might have reduced error rates by simplifying eligibility determination. These included: excluding one vehicle or all vehicles from the asset test for eligibility, and expanded categorical eligibility for recipients of non-cash TANF services (thus eliminating both vehicle and cash asset tests).[59] Contrary to expectations, excluding one vehicle was associated with a significant increase in the error index, but none of the base model parameters was affected. Given the lack of a rationale for why this policy would increase error, we assumed that it proxied for an unknown variable and left it out of the base model. Expanded categorical eligibility had no effect on the error index. It is possible that the simple indicator variable did not adequately represent the policy differences among States.

Fourth, we introduced variables representing additional time-varying effects, in an effort to refine the understanding the changes in the error index over time. Adding year effect dummy variables did not materially improve the model; only one year effect (for 2003) was significant. We interacted PCT_TANF and EFFORT with two post-PRWORA time periods (1997-2002 and 2003-2005) and with year dummy variables. The standard errors of the coefficient estimates for these time-specific interactions were too large to find any significant differences between periods or years. These alternatives confirmed the results of the base model regarding the changing relationship of EFFORT to ERROR over time.

Finally, we tested alternative modeling approaches to dealing with the potential bias arising from the correlation of panel data within States over time. These modeling approaches and their results are described in Appendix C. The alternative models support the results we found with the fixed effects GEE method. As the GEE is a more robust specification, we concluded that it was the correct specification from which to draw inference

Summary of Results

Below we summarize the findings regarding the three hypotheses presented in Chapter Four. We find substantial support for all three hypotheses, as discussed below.

Relationship of Certification-Related Effort to the Error Index

The consistent results across all of the models lead to the following interpretation of the relationship of certification-related effort to the error index (i.e., the β value in the basic model of ERROR in equation 4-1). Given the inclusion of fixed State effects and the range of alternative specifications, we are confident that the model has sufficient control variables and it is appropriate to treat associations as effects.

---

[59] We expected that relaxing asset requirements would reduce errors by simplifying eligibility determination, but an opposite effect was conceivable. In principle, relaxing asset requirements might increase the number of applications and participating households, and this effect might increase the error index. We did not find that increases in the caseload in general affected the error index, as shown by the lack of effect for the PCT_CHCASELOAD variable seen in table 5-2.

- When the level of effort increased, the error index fell (confirming the first hypothesis).
- On average, an increase of 10 percent in effort reduced the error index by .3 percentage points, at an annual cost of about $35 per household.
- The marginal effect of effort on the error index became smaller (in absolute value) in the early years after enactment of the TANF program (1997-1998), but it grew back to pre-TANF levels from 2001 to 2005.
- In the context of the TANF program, the effect of effort on the error index increased as the percentage of FSP households receiving TANF (PCT_TANF) decreased and vice versa (as indicated by the positive effect of TANFIMP_TANF_EFFORT).
- Thus, if the level of effort and other variables were kept constant, the error index would decrease as the percentage of FSP households with TANF decreased. This effect did not occur under the AFDC program.
- Increasing the percentage of households with short certification periods (CM13) reduced the error index but also reduced the marginal effect of effort on the error index (as indicated by the positive effect of CM13_EFFORT). Apparently, the additional staff time to recertify cases more frequently had an adverse effect on the overall productivity of certification staff, as measured by the level of the error index for a given level of effort.

We find, therefore, that two significant factors explain why the marginal effect of certification-related effort on the error index became smaller after the enactment of TANF: the interaction of the level of effort with the percentage of FSP households receiving TANF (TANFIMP_TANF_EFFORT), and the interaction of the level of effort with the percentage of households with short certification periods (CM13_EFFORT). Both of these effects were stronger from 1997 to 2002 than in earlier and later years. A third factor contributing to the smaller marginal effect in 1997-2002 was the variable in the model (EFFORT_SQ) representing diminishing returns to effort. This variable contributed to the estimated marginal effect, but it was not statistically significant in the model. The average of EFFORT was 20 percent higher in 1997-2002 than in the periods before and after, so contribution of diminishing returns to effort was greater.

We suggest two possible explanations for the effect of the interaction of EFFORT with PCT_TANF on the error index: a transient effect and a more long-term effect. The first explanation is that the implementation of TANF temporarily created a general disruption of local assistance offices. The disruption was larger where more FSP households received TANF, so in such States an additional unit of effort had a smaller effect on the error index. We find some evidence of such a transient effect: the estimated effect of the interaction between PCT_TANF and EFFORT was smaller in 2003-2005, based on the models with three or more periods interacted with these variables. However, we could not conclude that the effect of the interaction was different, because the standard errors are large relative to the size of the estimated coefficients. Additional years of data might provide a sufficiently precise estimate to support a conclusion that the effect was smaller for 2003 and later.

With the data that we have, we cannot rule out the more parsimonious conclusion that the effect of the interaction between PCT_TANF and EFFORT was the same from 1997 to 2005, and that it is evidence of a long-term change. The evidence is consistent with the long-term interpretation that TANF implementation increased the complexity of jointly administering these programs. TANF implementation led to more differences in rules between cash assistance and food stamps, and therefore workers "juggling" both programs were more likely to be confused about the rules. In addition, TANF required workers to focus more on clients' employment status and barriers to work, possibly increasing the time per case spent on "common" activities shared with the FSP (such as taking general case information) and reducing their focus on payment accuracy. As a result, there was a difference in efficiency and accuracy between generalist workers (handling TANF as well as food stamps) and specialized (food stamp-only) workers. One would expect that the proportion of generalist workers rises as the percentage of FSP households with TANF rises. If so, then PCT_TANF would be inversely related to workers' efficiency and accuracy. This effect might have been stronger during the early years of TANF, when the TANF rules were new to all workers, but it could persist. The effect of the interaction between PCT_TANF and EFFORT was mitigated by the decline in PCT_TANF from 34 percent in 1996 to 14 percent in 2005.

The previous study suggested that the reduced effort elasticity of error after 1996 might be the result of changes in how States allocated costs between AFDC/TANF and the FSP. Many States allocated shared ("common") costs for AFDC/FS cases entirely to AFDC. The Agricultural Research, Extension, and Education Reform Act of 1998 (AREERA) required all States to split these shared costs between TANF and the FSP, starting in 1999. Thus, the FSP cost for a FS/TANF case would be greater than the cost for a FS/AFDC case, even if nothing else changed other than the method of cost allocation. States might have changed their cost allocation practices before they were required to, because of the difference in funding between TANF and the FSP.[60]

The only available national data on the extent to which common costs for AFDC/FS cases were allocated to AFDC are the adjustments to certification costs.[61] If these adjustments were highly correlated with the percentage of food stamp households receiving AFDC in the base year for the adjustments (1994), this might suggest that the observed effect of this percentage was the result of changes in cost allocation. However, this correlation is very low—just 0.105.

On balance, we do not find that the evidence supports the hypothesis that changes in States' cost allocation practices contributed to the decline in the effect of EFFORT on the error index after the enactment of the TANF program. The interaction of PCT_TANF with EFFORT could be interpreted as evidence of cost-shifting from TANF to the FSP, which would reduce the efficacy of the worker time allocated to the FSP. Nevertheless, we do not find a significant interaction of PCT_TANF with

---

[60] Federal funding shifted from cost-matching under AFDC to the TANF block grant, and this change altered the financial impact of cost allocation decisions. On the one hand, States would receive more Federal funds if they allocated costs to the FSP, while allocating costs to TANF would not increase their Federal funds. On the other hand, if an incremental cost was allocated to TANF, the State would not need to provide additional non-Federal funds, assuming that the State's maintenance of effort requirement was met. Thus, it is not clear whether States would have an incentive to allocate more costs to the FSP or to TANF. Furthermore, oversight by the Department of Health and Human Services and FNS assures that costs are allocated equitably among programs, in accordance with Federal regulations.

[61] As required by AREERA, FNS deducts a specified amount from the Federal share of certification costs for 44 States, based on a determination of the amount of common costs in 1994 that would have been allocated to the FSP under the "benefiting program method", i.e., if the costs had been shared equally between AFDC, FSP and (if applicable) Medicaid. Thus, the amount of the adjustment, as a percentage of the State's FSP certification expenses in 1994, is an indicator of the degree to which the State allocated common costs to the AFDC program instead of the FSP.

EFFORT in the pre-TANF period. If cost allocation practices under AFDC consistently shifted common costs away from the FSP, one would expect that the incremental effort to reduce error (as charged to the FSP) would have decreased as the percentage of FSP households receiving AFDC increased. This would lead to a negative effect for the NOTI_TANF_EFFORT interaction, provided that "juggling" the FSP and AFDC did not affect errors. To the contrary, the coefficient is not significantly different from zero. Instead, it is possible that any efficiencies from joint administration of AFDC/FS cases were offset by errors resulting from the greater challenge of achieving FSP payment accuracy in this context. It is also possible that variations in cost allocation disrupted the expected relationship between PCT_TANF, EFFORT, and the error index; this interpretation does not fit with the hypothesis that common costs were systematically shifted from AFDC/TANF to the FSP. Furthermore, we do not find evidence that States with higher percentages of food stamp households receiving AFDC were more likely to have substantial common costs that were required to be reallocated to the FSP after 1996. Lacking evidence of a general relationship of cost allocation practices to the EFFORT-ERROR relationship, we prefer the preceding explanation based on short-term or long-term effects of TANF implementation.

The model does not include any indicators for the number of FSP households with immigrants or able-bodied adults without dependents (ABAWDs). Both of these types of households were significantly affected by provisions of PRWORA, and the implementation of these provisions could affected the error index, particularly during the initial implementation of PRWORA and also during the implementation of AREERA, which restored benefits to some immigrants. However, the share of the caseload affected by these changes was small, so their expected impact was small. In 1996, 8.9 percent of FSP households had a permanent resident alien; this figure fell to 7.1 percent in 1997. ABAWD food stamp households were 9.5 percent of the total in 1996 (Cody and Castner, 1999). Furthermore, the rules governing these population groups varied over time and among States, so a series of variables representing these rules would be needed to fully capture their effect on error rates. To avoid over-complicating the model, we opted not to attempt this, but it is unlikely that the estimated effects are materially affected by this omission.

Policies Affecting Error Rates

The models consistently show that several policies were significantly associated with lower levels of error, with the following effects on the error index in the base model:

- Short certification periods had two effects: a reduction in the level of the error index (α in equation 4-1) and a reduction in the absolute value of the marginal effect of effort (β, as discussed above). At the average rate (9.5 percent of food stamp households), the level effect pushed the error index down by 1.7 percentage points, relative to the level with no short certification periods; the net effect was about 1.0 percentage points, after taking into account the interaction with effort. For the period with the greatest increase in use of short certification periods (1992 to 2000), the change in the error index was 14.5 percentage points, and the estimated net effect of short certification periods on the error index was -1.5 percentage points.[62]

---

[62] At the mean of EFFORT for all years (.010), the net effect of a 1 percentage point increase in CM13 on ERROR was (.01*(-.184 + (7.990*.010)= -.001).

- Simplified reporting had an estimated effect of about -0.6 percentage points at the overall mean (15.5 percent of States) and -3.6 percentage points with 100 percent of States using this option.[63]
- Transitional benefits for households leaving TANF had an estimated effect of -0.1 percentage points at the overall mean (4.2 percent of States) and -2.7 percentage points at 100 percent usage.

The finding of a significant effect for simplified reporting is consistent with the results of the previous study (Fink and Carlson, 2005). The size of the estimated effect in our study is larger than the 2.2 percent impact estimated by Fink and Carlson at full usage. However, the methods are sufficiently different that a quantitative comparison of the estimates is not appropriate. Both simplified reporting and transitional benefits create a "hold harmless" condition when most changes in household circumstances do not affect eligibility and benefits, and therefore do not create the potential for error. The size of the effect of transitional benefits is somewhat surprising, given the relatively small percentage of food stamp cases with TANF benefits, but there is no clear benchmark for this effect. We caution readers that operational changes intended to reduce errors (such as specialized change processing units) could be correlated with the policy options and contribute to their estimated effect. The models' fixed effects pick up the effects of uncorrelated operational changes.

We did not find a significant effect for quarterly reporting or simplified definition of income in most models. For quarterly reporting, the small number of observations under this policy could have affected the precision of the estimates; this variable had some estimates significant at the 10 percent level (including the base model) and some at the 5 percent level.

Reconciling the Declines in Error Rates and Effort after 2000

The results provide an explanation of why error rates fell while certification-related effort fell after 2000. The reason was that new policy options reduced the level of error independently of the level of EFFORT.

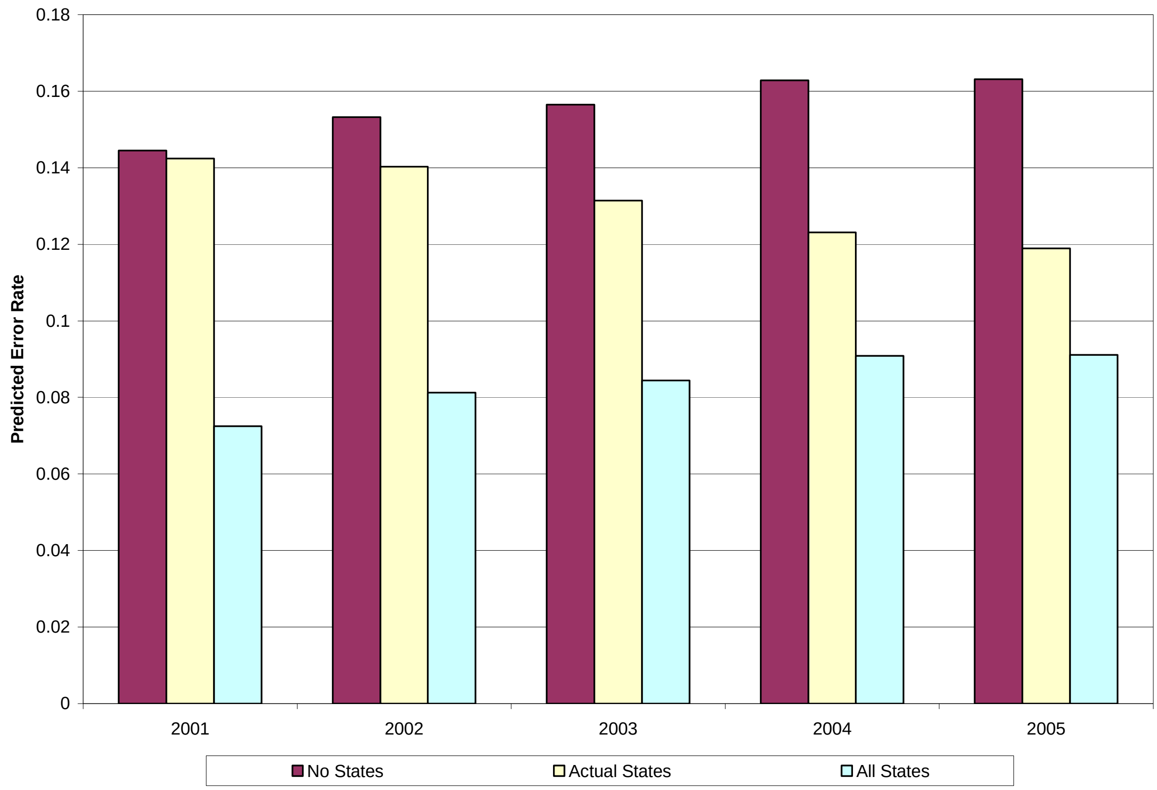
We simulated the combined effects of simplified reporting, transitional benefits, and simplified income definition, as illustrated in figure 5-6. (Simplified definition of income did not have a significant effect as an individual policy, but it was part of the package of "simplification" policies and is therefore included here.) For 2001-2005, the average error index was estimated for three scenarios: "no States" (policy variables equal zero), represented by the left bar for each year; "actual States" (policy variables at actual mean values), the middle bar for each year; and "all States" (policy variables equal one), the right bar for each year. All other variables were valued at their means for the year. The middle bar (actual States) in each set drops by 2.3 percentage points from 2001 to 2005, reflecting the reduction in error as increasing percentages of States used these options. The predicted error index increases by about 1.9 percentage points from 2001 to 2005 for the "no States" scenario (the trend in the left bars for each year) and the "all States" scenario (the trend in the right bars for each year).

---

[63] States vary in the proportion of the FSP caseload assigned to simplified reporting, but statistics on this variation were not available. Therefore, the results reflect the actual proportion of the caseload under simplified reporting, and greater effects might be observed with higher proportions.

Figure 5-6

Estimated Impact of Key Policies on Error Index



Policies include simplified reporting, transitional benefit for TANF leavers, and simplified definition of income.

These scenarios yield the following estimates:

- The actual adoption of these policy options reduced the error index by 2.5 percentage points in 2003, 4.0 percentage points in 2004, and 4.4 percentage points in 2005, as seen by comparing the "actual States" and "no States" scenario for each year. The effect was smaller in 2001-2002, before all of the options were available to all States.
- If all States had adopted these options in 2005, the error index would have been 2.8 percentage points lower, as seen by comparing the "actual States" and "all States" scenarios.
- For each year, the difference between no States and all States with these options is 7.2 percentage points.

Thus, these three policy options jointly explain the reduction in the error index, which would have increased if they had not been available. This counterfactual assumes that States would have expended the same (declining) level of effort as they actually did in these years, and that all other variables (notably the use of short certification periods and quarterly reporting) would have changed, on average, as they actually did. The point estimates of the impact of the three policy options vary among the models, but the qualitative results for these options are robust.

The relatively small increase of 1.9 points in the error index in the no-state vs. all-state scenarios is somewhat surprising, given the 19 percent reduction in certification-related effort from 2001 to 2005 (per table 4-1). Based on the change in effort (-0.027) and the marginal effect of effort on error in 2001 (-1.317), one would expect an increase of 3.6 percentage points in the error index during this period.[64] The main factor that offset this was the decline in the percentage of FSP households with TANF. This trend increased the marginal effect of effort on the error index, due to the previously discussed interaction among these variables. The drop in the percentage of cases with short certification periods also increased the marginal effect of effort on error, but the overall effect of this change was a net upward effect on the error index. Changes in other caseload characteristics or economic conditions may have also offset the decline in EFFORT.

[64] This projection is approximate, holding the marginal effect of effort constant. In fact, if EFFORT declined and nothing else changed, the marginal effect would increase because of the contribution of the EFFORT_SQ term.

Chapter Six: Conclusions and Discussion

This chapter summarizes and explains the trends in certification effort and error rates, using the authors' previous findings confirmed by this study and the new findings. The chapter also discusses the limitations of the study and potential directions for future research.

Trends in FSP Caseloads, Certification-Related Effort, and Error Rates

This study identified several national trends in the FSP from 1989 to 2005:

- In the early 1990's, error rates rose while the FSP caseload rose and the certification-related effort per household fell.
- Error rates fell in the mid-1990's as the trends in the FSP caseload and effort reversed.
- In 1997-1998, as States implemented welfare reform, the FSP caseload continued to decrease and effort continued to increase, but error rates rose.
- From 1998 to 2000, error rates again fell, while the decline in the caseload and the increase in effort continued.
- From 2001 to 2005, error rates continued to fall as the caseload increased and effort declined.

The authors' previous research found a significant and robust negative relationship of effort to case error rates; fluctuations in effort explained much of the rises and falls of error rates from 1989 to 2001. Within this context, the 1997-1998 spike in error rates represented an apparent anomaly. The authors' previous research showed that this anomaly was partly the result of a reduction in the marginal effect of effort on error, significantly associated with and possibly due to the implementation of PRWORA.

The present study sought to better understand the reasons for the trends from 1996 to 2001. In particular, the study sought to identify the variables that changed the marginal effect of effort on error, and to confirm and expand the set of variables independently affecting the level of error.

In addition, the present study sought to explain why error rates fell from 2001 to 2005 while States reduced their effort per case. The principal hypothesis was that new FSP policy options enabled States to reduce error rates. We also considered that the relationship of effort to error might have changed so that less effort was needed to achieve the same level of error.

At the national level, the total level of effort appeared to adjust slowly in response to caseload changes, so that the effort per FSP household moved in the opposite direction from the caseload. This relationship of caseloads to effort was posed as a possible explanation for the parallel trends in FSP caseloads and error rates. However, the change in the FSP caseload also was included in the models to test whether it had an independent effect on error rates.

Explanation of Trends in Error Rates

Like the authors' previous study, this study found a significant and robust negative effect of effort to error, and a reduction in the absolute size of the marginal effect of effort on the error index after 1996. This finding confirms the importance of taking spending or effort into account when modeling error rates and the limitation of studies that do not do so. The present study also confirmed the conclusion of the authors' previous study and others that increased use of short certification periods in the 1990's significantly reduced error rates, and that error rates increased with the percentage of FSP households with earnings.

Considering the negative effect of effort on error rates, the trends in effort played an important role in the fluctuation in the average error index between 1989 and 2005. The role of effort was greatest in 1989-1996, when the marginal effect of effort on error was greatest. Based on this measure, the reduction in effort over this period would be expected to increase the error index by 1.4 percentage points, and the actual total increase from all factors (positive and negative) was 1.5 points. The increase in the percentage of FSP households with earnings also contributed to the increase in error, while the increase in short certifications and the reduction in the percent of FSP households receiving AFDC pushed the error index in the opposite direction. For 1997-1998 and 2001-2005, the trend in the error index was the opposite of what would be expected based on the trend in effort. However, the error index would have been substantially higher in 1997-1998 and lower in 2001-2005 in the absence of the trend in effort.

The spike in the error index in 1997-1998 occurred largely because the implementation of TANF reduced the effect of FSP effort on error. This change was related to the percentage of FSP households receiving TANF, so the larger the State's TANF program was (relative to the FSP), the larger the reduction in the effect of effort on error. As the TANF percentage fell, this effect diminished in 1999-2005. Increases in the percentage of FSP households with earnings also contributed to the 1997-1998 spike in the error index.

Along with the decline in the TANF percentage, policy variables played an important role in driving down the error index from 1998 to 2005. In 1998-2000, the increasing use of short certification periods offset the diminished marginal effect of effort on error due to the TANF interaction effect. The introduction of quarterly reporting may have also contributed to the reduction in error during these years. The decline in the error index in 2001 to 2005, despite declining effort, was mainly due to the introduction of simplified reporting, simplified definition of income, and transitional benefits for TANF leavers.

The data compiled for this study allowed the estimation of the effects of the new policy options of the current decade, including simplified reporting, simplified definition of income, and transitional benefits for FSP households leaving TANF. Unlike the previous study that simulated the effects of simplified reporting with data for an earlier period (Fink and Carlson, 2005), these estimates used actual error rates while simplified reporting was in effect and took into account the role of effort and other factors simultaneously influencing error rates. The simulations for this study suggest that wider adoption of simplified reporting and other simplification policies could further reduce the error index, assuming that barriers to wider adoption (such as reprogramming data systems) can be overcome.

Causes of Changes in the Marginal Effect of Effort on Error

The previous study attributed the change in the marginal effect of effort to PRWORA in general and found a separate, positive effect of the percentage of FSP households receiving TANF on the level of error. Instead, we now attribute the change in the marginal effect primarily to the interaction of TANF implementation, the TANF percentage, and effort. This interaction was the largest contributor to the differences in marginal effect between periods. The additional years of data (2002-2005) and the additional policy variables allowed us to improve our analysis of this important change. The present study also uncovered evidence that short certification periods reduce the net effect of effort, underscoring that the benefits of short certification periods are better understood in the context of their costs.

The new results helped us focus on the effects of TANF implementation in understanding why we observe this relationship. While the previous study explored the possible benefits of cost-sharing between cash assistance and the FSP, the present study suggests that it is challenging to maintain FSP accuracy while operating TANF and the FSP, with their different rules and priorities, in the same offices. There is no evidence that this challenge existed when States operated the AFDC program under standard rules, but presence of AFDC reform waivers was associated with higher error rates. We are not able to conclude, however, whether this challenge was temporary or is likely to persist. The present study also casts doubt on the possibility, suggested by the previous study, that changes in cost allocation contributed to the observed reduction in the effect of effort on error.

Limitations of the Study

One important limitation of the study is that the measure of certification-related effort is a proxy for the ideal measure of actual staff time devoted to certification. The effort measure varies not only because of differences in actual certification worker time per case but also because of differences in staffing mix and overhead. The definition of certification-related costs includes some costs that may not be related to certification, because of variations in how States report costs. Further, certification costs may be spent on improvements in other dimensions besides accuracy, such as timeliness of application processing and facilitating access to the FSP. In addition, the certification-related costs include non-personnel costs for which wage rates may not be the appropriate index to normalize costs across States. The average public welfare worker wage rate used in the effort measure is a proxy for the average pay of workers who administer the FSP. There is still uncertainty about how variations in cost allocation methods, both over time and among States, affect reporting of FSP administrative costs. Thus, more detailed studies of FSP administrative costs may be needed to validate that the effort measure—as computed from the available data—is a good proxy of actual certification staff time.

Another limitation of the study is the lack of success in using automated data processing (ADP) costs as a measure of a State's degree of automation when modeling error rates. These costs had no identifiable effect on their own, and results did not change when they were combined with certification-related costs in the effort measure. It is possible, of course, that automation does not have a detectable effect on error once effort is taken into account. The alternative explanation is that ADP costs are a poor proxy for the capabilities of a State's ADP system to reduce errors. In principle, automation could reduce errors directly, by guiding workers to collect and evaluate information correctly, and by automatically identifying discrepancies. In addition, automation could reduce errors indirectly, by saving time on routine tasks for workers so that they can focus on improving their accuracy. Because of the lack of a direct measure of automation, the effects of this factor on error are likely captured by the fixed State effects or State-specific time trends. Better understanding of the variation in automation among States might permit development of a proxy measure that could be used more successfully in modeling error rates. An alternative would be a more in-depth approach to relate specific changes in automation to the likelihood of error in activities affected by the automation.

We note that an important but unaddressed feature of the 2002 Farm Bill was the reform of quality control sanctions, so that State liabilities for excessive error rates were based on two years' performance, not just one. With only three years of data after the enactment of this provision, we did not attempt to model its impact.[65]

Last but not least, the study's results are inconclusive on a key question: whether the reduced effect of effort on error after 1996 was transient or more long-term. The marginal effect estimates suggest that the effect was transient: the estimated marginal effect grew every year in absolute value from only -0.646 at its lowest (in 1998) to -3.986 in 2005, cutting the effort required to achieve a specified error reduction by 84 percent. The average marginal effect of effort for 2003-2005 was significantly larger (in absolute value) than the effect for 1997-2002. The estimated effect of -3.479 for 2003-2005 was numerically smaller than the estimate of -5.327 for 1989-1996. Alternate models suggest that the interaction effect between the TANF percentage and effort became smaller after 2002. However, the standard errors of estimates of effect coefficients and marginal effects for portions of the study period were large, relative to the estimates, so we could not detect changes over time. Thus, we have conflicting evidence about whether the "TANF effect" is likely to persist. Additional years of data might help, but it is also possible that, as error rates continue to drop to historically low levels, the power to detect the effect of effort—and other factors—may diminish.

Suggestions for Future Research

Given the impact of certification-related effort on error rates, a better understanding of the variation in effort would be a useful direction for FSP research. One line of investigation would be to understand the relationship of spending and effort to the caseload cycle. Another line of research would be to examine the long-term differences in spending and effort across States. Finally, more detailed studies of worker time and the composition of certification-related costs would help to validate and improve the effort measure. Studies of differences in worker time to administer similar cases might help to understand the interactions with effort and the role of automation.

Future studies might investigate the extent to which the organization of FSP agencies may affect error rates. There are several dimensions of variation in organization that could have effects on error rates, including: whether the FSP is administered by the State agency or by Counties under State

[65] The change in sanctions policy applied equally to all States. One might attempt to model the impact by introducing an indicator for the years when the new policy was in effect. This approach was not feasible with only three years of data after the enactment of the 2002 Farm Bill. It was not clear whether the change would affect States' behavior immediately or with a lag; furthermore, States did not learn their official error rates for 2003 until several months after the start of FY2004. Given the uncertainty about the timing of the sanctions policy's effect, an attempt to model the impact would likely require several additional years of data beyond what was available.

supervision; the size of FSP offices, in terms of the number of staff or the caseload served; the degree to which workers specialize in the FSP (versus other programs), in the type of certification task (intake, case maintenance and changes, recertification) or type of case; and the use of traditional face-to-face methods versus newer approaches (web-based or mail-in applications, telephone interviews, centralized call centers, etc.).

Finally, a case-level approach to analyzing error rates might provide different insights from those of the State-level analysis in the present study. Previous studies have used case-level data to model the factors affecting error rates for different types of cases (such as with and without earnings) and for different parts of the case cycle (intake, interim changes, recertification, and closure). A new direction would be to introduce a measure of effort to these models; this may require a measure of effort that varies by case type or activity. In addition, research on the relationship of worker characteristics to errors could be useful. This could be done with State-level measures or with linked data on case actions and worker characteristics (such as pay, seniority, and programs served).

Citations

American Public Human Services Association (APHSA), 2006. *Successful Food Stamp Innovations.* www.aphsa.org, accessed October 27, 2007.

Arellano, M., and S. Bond, 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58(2), 277-297.

Baltagi, B.H., 1995. *Econometric Analysis of Panel Data.* John Wiley & Sons, New York.

Bertrand, M., E. Duflo, and S. Mullaniathan (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics, 119*(1), 249-275.

Carmody and Dean, 1998. *New Federal Food Stamp Restoration for Legal Immigrants: Implications and Implementation Issues.* Center on Budget and Policy Priorities, Washington, D.C.

Center on Budget and Policy Priorities (CBPP), 2005. *Food Stamp Overpayment Error Rate Hits Record Low.* Washington, D.C.

Cody, S. and L. Castner, 1999. Characteristics of Food Stamp Households Fiscal Year 1997. Food and Nutrition Service, USDA, Alexandria, VA.

Economic Research Service (ERS), U.S. Department of Agriculture, 2007. *2002 Farm Bill Title IV: Nutrition Programs.* www.ers.usda.gov/features/farmbill/titles/ titleIVnutritionprograms.htm. Accessed November 5, 2007.

Finegold, K., W. Margrabe, and C. Ratcliffe, 2007. *Food Stamp Program Rules Database Documentation*. Prepared by The Urban Institute for the U.S. Department of Agriculture, Economic Research Service, Food Assistance and Nutrition Research Program, Washington, D.C..

Fink, K., and S. Carlson, 2005. *The Effect of Simplified Reporting on Payment Accuracy.* Food and Nutrition Service, U.S. Department of Agriculture, Alexandria, VA.

Food and Nutrition Service (FNS), U.S. Department of Agriculture, 1994. *Food Stamp Program Quality Control Annual Report, Fiscal Year 1993.*

FNS, 2003. *Food Stamp Program State Options Report, Second Edition.* Alexandria, VA.

FNS, 2007a. Food Stamp Program Participation and Costs (Data as of October 26, 2007). http://www.fns.usda.gov/pd/fssummar.htm, accessed on November 2, 2007.

FNS, 2007b. Food Stamp Program: Average Monthly Participation (Households) (Data as of October 26, 2007). http://www.fns.usda.gov/pd/16fsfyhh.htm, accessed on November 2, 2007.

FNS, 2007c. *Food Stamp Program Quality Control Annual Report, Fiscal Year 2005.* Alexandria, VA.

FNS, 2007d. Electronic Benefits Transfer (EBT) Status Report, June 2007. http://www.fns.usda.gov/fsp/ebt/ebt_status_report.htm, accessed on November 15, 2007.

Holcomb, P.A., and K. Martinson, 2002. *Implementing Welfare Reform across the Nation.* Series A, No. A-53. Urban Institute, Washington D.C.

Kabbani, N.S., and P.E. Wilde, 2003. Short Certification Periods in the U.S. Food Stamp Program. *Journal of Human Resources 38*, Supplement 2003.

Liang, K.-Y., and S.L. Zeger (1986). Longitudinal Data Analysis using Generalized Linear Models. *Biometrika*, *73*(1), 13-22.

Logan, C., W. Rhodes, and J. Sabia, 2006. *Food Stamp Program Costs and Error Rates, 1989-2001.* CCR-15. Economic Research Service, U.S. Department of Agriculture, Washington, D.C.

Mills, G., 1991. *The Relationship Between Overpayments and Underpayments in the Food Stamp Program: An Updated Analysis.* Food and Nutrition Service, U.S. Department of Agriculture, Alexandria, VA.

Mills, G., D. Laliberty, and C. Rodger, 2004. *Food Stamp Certification Periods and Payment Accuracy, State Experience During 1997-2001.* Economic Research Service, U.S. Department of Agriculture, Washington, D.C. USDA-ERS E-FAN-04-012.

Puma, M., and D. Hoaglin, 1987. *The Effect of Caseload Characteristics and Socioeconomic Conditions on Food Stamp Program Error Rates.* Food and Nutrition Service, U.S. Department of Agriculture, Alexandria, VA.

Ratcliffe, C., McKernan, S.M, and Finegold, K., 2008. *The Effect of State Food Stamp and TANF Policies on Food Stamp Program Participation*. Economic Research Service, USDA, Washington, D.C.

Rosenbaum, D., 2000. *Improving Access to Food Stamps: New Reporting Options Can Reduce Administrative Burdens and Error Rates.* Center on Budget and Policy Priorities, Washington, D.C.

Rosenbaum, D., 2003. *Food Stamp Overpayment Error Rate Hits Record Low.* Center on Budget and Policy Priorities, Washington, D.C.

U.S. Department of Health and Human Services (HHS), 2007. Temporary Assistance for Needy Families, Separate State Program-Maintenance of Effort, Aid to Families with Dependent Children—Caseload Data. http://www.acf.hhs.gov//programs/ofa/caseload/caseloadindex.htm, accessed on November 5, 2007.

U.S. Government Accountability Office (GAO), 2001. *Food Stamp Program: States Seek to Reduce Payment Errors and Program Complexity.* Washington, D.C.

U.S. Government Accountability Office (GAO), 2002. *Food Stamp Program: States' Use of Options and Waivers to Improve Program Administration and Promote Access.* Washington, D.C.

U.S. Government Accountability Office (GAO), 2004. *Food Stamp Program: Farm Bill Options Ease Administrative Burden, but Opportunities Exist to Streamline Participant Reporting Rules among Programs.* Washington, D.C.

U.S. Government Accountability Office (GAO), 2005. *Food Stamp Program: States Have Made Progress Reducing Payment Errors, and Further Challenges Remain.* Washington, D.C.

U.S. Government Accountability Office (GAO), 2007. *Food Stamp Program: Use of Alternative Methods to Apply for and Maintain Benefits Could be Enhanced by Additional Evaluation and Information on Promising Practices.* Washington, D.C.

Windmeijer, F., 1995. A Note on R2 in the Instrumental Variables Model. *Journal of Quantitative Economics, 11*, 257-261.

Appendix A: Sources and Methods for Analysis of FSP Administrative Costs

Data Sources

For this study, FNS provided the FSP administrative cost data for 1989-2005 in a series of extracts from the agency's National Data Bank. This data warehouse contains FSP administrative costs and other data reported by the States to FNS.[66] The States submit quarterly and annual reports of FSP administrative expenses to FNS on the SF-269 report. FNS enters these data into its Food Stamp Program Integrated Information System (FSPIIS) and periodically extracts the data to the National Data Bank. FNS provided the data extracts for 1989-2001 from the National Data Bank in November 2003. We verified the national totals for each year against reports provided by FNS from the FSPIIS in July 2004. FNS provided extracts for 2002-2005 in October 2006.

On the SF-269 report, States break down their FSP administrative costs in two ways. The report is organized in columns representing specific program functions, such as certification, benefit issuance, and automated data processing. For each column, the State report identifies Federal and non-Federal shares of outlays. Depending on the function, the Federal share (also known as Federal Financial Participation, or FFP) is set by law at 50 to 100 percent of reimbursable expenses (except for one category, Reinvestment, for which the States do not receive Federal reimbursement). Over the period covered by the data, the Federal share for some expenses was reduced from 75 percent to 63 or 50 percent. (The reporting categories and associated statutory FFP rates are presented later in this section.)

States also report Federal and non-Federal shares of unliquidated obligations on the SF-269. Unliquidated obligations represent commitments of funds that have not been formally expended (i.e., liquidated). The liquidation of obligations is part of the process of finalizing expenditures for a fiscal year. In consultation with FNS, we determined that States could no longer liquidate the unliquidated obligations for this period, so these obligations were not counted in the analysis.

Estimation of Total Costs for Analysis Categories

Preliminary analysis indicated that the actual Federal share (i.e., the Federal outlay divided by the sum of the Federal and non-Federal outlays) frequently was different from the statutory FFP rate. Most often, the actual Federal share was less than 50 percent when the statutory FFP rate was 50 percent. FNS officials indicated to us that in these situations, the Federal outlay was more reliable than the non-Federal outlay, because FNS reviews and confirms the allowable Federal cost during the cost reporting and reimbursement process.[67] For certification costs, as required by AREERA, FNS

[66] The State Agencies include the 50 States and the District of Columbia. FSP costs for Guam and the U.S. Virgin Islands are not included in this analysis. Puerto Rico does not operate the FSP.

[67] FNS officials indicated that when a Federal outlay is revised due to identification of an error or unallowable costs, the corresponding non-Federal outlay may not be revised by the State, and FNS does not attempt to make this correction. Also, in some cases States have included unallowable costs in their reported non-Federal outlays, even though they claimed Federal reimbursement only for allowable costs.

deducts a specified amount from the Federal share of certification costs for 44 States, based on a determination of the amount of common costs in 1994 that would have been allocated to the FSP under the "benefiting program method", i.e., if the costs had been shared equally between AFDC, FSP and (if applicable) Medicaid.

Therefore, the costs presented in this report, with some exceptions, were estimated by dividing the statutory FFP rate into the Federal outlays (in dollars) as reported on the SF-269. For example, a Federal outlay of $1000 in a category with a statutory FFP rate of 50 percent would yield an estimated total of $2000, (i.e., 1000/0.5=2000). In a few minor categories, however, this approach was not feasible (due to their being no Federal share) or yielded clearly unacceptable results, so the actual total was used (Federal outlay plus non-Federal outlay).[68] In all cases, the intent was to use the best estimate of the total allowable cost.

Table A-1 shows how the numerous cost reporting categories were grouped into the categories used for the analysis. The largest categories were generally kept separate. The various issuance-related costs were combined, however, to allow for better comparisons over the study period, when most States replaced their coupon issuance systems with electronic benefit transfer (EBT) systems.[69] The various levels of funding for ADP development and fraud control were combined, as were the components of employment and training (E&T).

The SF-269 functional reporting categories correspond to the FSP functions described in Chapter Four. Throughout the study period, each SF-269 category was generally intended to include the direct costs of the function and the indirect costs that are allocated to the FSP as a result of those direct costs.[70] For example, the reported certification cost is expected to include both the direct costs of certification (such as certification worker salaries and travel) and the indirect costs allocated on the basis of direct worker time or costs (such as shares of local office management, equipment and occupancy, and state-level oversight). For issuance costs, however, States are required to report some indirect costs separately from direct costs for EBT or coupon issuance.[71]

---

[68] The actual costs were used for: E&T 100 percent grant, ADP development, Reinvestment, Systematic Alien Verification of Eligibility (SAVE), Research and Demonstration Evaluation Projects, and State/Local Demonstration Projects. SAVE costs are for obtaining alien status information from the Immigration and Naturalization Service.

[69] The conversion from coupon issuance to EBT was usually a gradual process over a year or more. During this period, States incurred a combination of costs in the various issuance categories, with some portions of the FSP caseload receiving coupons while others received benefits via EBT. Thus, it was most useful to compare States' total issuance costs, with the recognition that they reflected a mix of issuance systems. This mix can be quantified for modeling through the use of data on the number of FSP households receiving benefits via coupons, EBT, and cash (which was used for some SSI recipients and participants in demonstration projects).

[70] Indirect costs are expenditures that cannot be efficiently attributed to a specific program or activity. States must have approved cost allocation plans in order to claim Federal reimbursement for indirect costs. These plans may allocate both personnel and non-personnel costs. Indirect personnel costs typically include supervisors and support staff who serve multiple programs and whose time on individual programs cannot be efficiently measured. Typical indirect non-personnel costs include general-purpose supplies, telecommunications, facilities, equipment, and contracted services.

[71] Indirect costs for issuance are reported separately if they are computed by applying an indirect cost rate, i.e., the State multiplies direct issuance costs by a specified percentage to compute the indirect costs. If the

Table A-1

Cost Categories, Federal Financial Participation Rates, and Basis for Estimating Total Outlays for Analysis

Analysis Category and Included Reporting Categories	Statutory Federal Financial Participation (FFP) Rate	Basis for Estimating Total Outlays
Certification	50%	Federal total
Issuance		
Coupon Issuance	50%	Federal total
EBT Issuance	50%, capped on the basis of prior issuance costs	Federal total
Issuance Indirect	50%	Federal total
EBT Startup	50%	Federal total
Fraud Control[1]		
75% Funding Fraud Control	75%	Federal total
50% Funding Fraud Control	50%	Federal total
Reinvestment (100% Non-Federal)	0%	Federal plus Non-Federal
Automated Data Processing (ADP) Operations	50%	Federal total
ADP Development[2]		
75% Funding ADP Development	75%	Federal plus Non-Federal
63% Funding ADP Development	63%	Federal plus Non-Federal
50% Funding ADP Development	50%	Federal plus Non-Federal
Employment and Training (E&T)		
E&T 100% Grant	100%	Federal plus Non-Federal
E&T 50% Grant	50%	Federal total
E&T Dependent Care	50%, with per-participant cap	Federal total
E&T Transportation/Other	50%, with per-participant cap	Federal total
Optional Workfare	50%	Federal total
E&T ABAWD[3] Grant	50%	Federal total
Food Stamp Nutrition Education (FSNE)	50%	Federal total
Unspecified Other (Direct and indirect costs not included elsewhere)	50%	Federal total
Miscellaneous		
Quality Control	50%	Federal total
Fair Hearing	50%	Federal total
SAVE[4]	100%	Federal plus Non-Federal
Outreach	50%	Federal total
Management Evaluation	50%	Federal total

State uses another allocation method, issuance indirect costs may be reported together with direct issuance costs.

Table A-1 (Continued)

Cost Categories, Federal Financial Participation Rates, and Basis for Estimating Total Outlays for Analysis

Analysis Category and Included Reporting Categories	Statutory Federal Financial Participation (FFP) Rate	Basis for Estimating Total Outlays
Research and Demonstration Evaluation Projects[5]	(varies)	Federal plus Non-Federal
State/Local Demonstration Projects[5]	(varies)	Federal plus Non-Federal

Note: Some categories do not appear in all years' cost data.

[1] The FFP rate for Fraud Control switched to 50% as of April 1, 1994, as mandated by the Mickey Leland Hunger Relief Act of 1993.

[2] ADP development was funded at the 75% FFP rate for approved projects in FY1989-1991. The rate of FFP was 63% from October 1991 through March 1994, and 50% thereafter, as mandated by the Mickey Leland Hunger Relief Act of 1993.

[3] ABAWD=Able-Bodied Adult Without Dependents. Some States received special grants for serving this portion of the FSP recipient population.

[4] SAVE=Systematic Alien Verification of Eligibility (costs for obtaining alien status information from Immigration and Naturalization Service). The official rate of FFP for SAVE expenses was 100% until April 1, 1994, and 50% thereafter, as mandated by the Mickey Leland Hunger Relief Act of 1993. The actual rate of FFP differed from the official rate in some years for unknown reasons.

[5] The FFP rate for Research and Demonstration Evaluation Projects and State/Local Demonstration Projects was determined by individual project budgets set by agreement between FNS and the State or local agency.

The "unspecified other" category is used for reporting costs not specifically identified elsewhere on the SF-269. According to FNS, this category may include both direct and indirect costs.[72] For example, the "unspecified other" cost may include state FSP staff whose time is not identified as spent on one or more specific program functions. Indirect costs are included in this category if they are not associated with a specific program funding stream. Early in the history of the FSP, many States assigned their indirect costs for all FSP operations to the unspecified other category. Before the study period, FNS instructed the States to change this practice and allocate indirect costs among the other categories. However, there may be differences among States and within States over time in the extent to which indirect costs are reported as "unspecified other costs" and not assigned to specific categories such as certification. Thus, the costs reported in a specific category in two different States may not be entirely comparable, because one State may have included all related indirect costs while the other did not.[73]

Although most of the FNS reporting categories represent recurring operational costs, there are several categories that are non-recurring and may be considered investments. In particular, EBT start-up and ADP development may be considered investments, because they are one-time costs that are intended to produce program improvements over a period of several years. Demonstration and evaluation costs may be considered an investment in knowledge. We did not attempt to amortize these non-recurring costs, however, for two reasons. First, there was no clear basis for determining the

[72] This information was provided through personal communications with several FNS staff in 2003-2004.

[73] The methods used to attribute direct costs to FSP reporting categories may also vary both among States and over time. For example, methods for attributing ADP costs to the FSP vary depending on system design and other factors.

appropriate amortization period, both because of the nature of these expenditures and because of the timing of when they are reported. Although there is no national database of detailed information on EBT start-up, ADP development, and demonstration and evaluation costs, the authors' past experience suggests that the largest components of these costs are state (or local) personnel and contractor services. Unlike equipment costs, which are usually spread over time through leases or explicit depreciation charges, the state/local personnel and contractor service costs do not have a standard "useful life" that would be suitable as an amortization period. Furthermore, costs of this type may be incurred in one year and reported in a subsequent year. The second reason for not amortizing the non-recurring costs was that, as shown in the results, we believed that the relatively modest share of costs in these categories did not justify the additional effort that would be required to develop and apply an appropriate methodology, particularly if the assignment of the useful life would be arbitrary.

Normalizing Costs for Comparisons

We normalized FSP administrative costs for comparisons of costs among States and over time. First, the total estimated cost for each category in each State and year was divided by the monthly average number of households participating in the FSP; thus the basic unit of measurement is the cost per case-year. Second, to compensate for the impact of inflation on comparisons over time, all costs were adjusted to 2005 dollars, using the price deflator for the Gross Domestic Product (GDP).[74]

[74] The GDP price deflator was used because it reflects the overall rate of inflation in the domestic economy, and because it has been used by FNS for setting EBT cost reimbursement limits. The Bureau of Economic Analysis (BEA) has a specific price deflator for state and local income security expenses. Use of the BEA deflator would have resulted in slightly higher normalized costs for years prior to 2005, with the greatest difference in the earliest years.

Appendix B: Additional Explanation of Multivariate Analysis

This appendix provides technical details to explain the fixed effects model that was used as the primary model in the multivariate analysis. The appendix provides the derivation of the likelihood function used in the grid search for the weighting parameter (λ) for the error index, and it also explains the computation of the marginal effect of effort on the error index. Technical information on alternative models is presented in Appendix D.

Fixed Effects Model

The fixed effects model in equation (5-2) in the text could be expressed more fully by substituting equation (5-1) into equation (4-2), yielding:

$$\begin{aligned}&ERROR_{p,it} + \lambda ERROR_{n,it} = \alpha_i + t_i'\delta_i + EFFORT'_{it}\beta_1 + EFFORT\_SQ'_{it}\beta_2 + \\ &TANFIMP\_TANF\_EFFORT'_{it}\beta_3 + NoTI\_TANF\_EFFORT'_{it}\beta_4 + \\ &CM13\_EFFORT'_{it}\beta_5 + X_{it}'\gamma + e_{it}\end{aligned} \qquad \text{(B-1)}$$

where the parameter λ was estimated via grid search by maximizing the log-likelihood function. The estimated value of λ was estimated at 1.45 in the fixed effects model. This estimate was robust to changes in the specification of the model, i.e., the inclusion or exclusion of variables and changes in functional form.

We optimized λ by choosing a value for λ, computing a new error measure (equal to positive error +(λ*negative error)), running a regression with the new error measure, and retaining both the value for λ and the log-likelihood. The regression with the largest log-likelihood indicated which value for λ was best. The lower bound for λ was -1, and the upper bound was 1. The likelihood function and its derivation are explained below.

Derivation of the Likelihood Function for the Estimation of λ

The likelihood function for the estimation of λ had the form:

$$\Pi \frac{A+\alpha B}{A+B} e^{-\frac{1}{2}(\frac{A+\alpha B-\mu}{\sigma})^2} / \sqrt{2\pi\sigma^2} \qquad \text{(B-2)}$$

where A is positive errors, B is negative errors, and subscripts (representing states and time) on A and B are implied. Also μ is the estimated mean. This portion of the appendix proves that assertion.

We write the normal density for the random variable Y as:

$$\varphi(Y) = e^{-\frac{1}{2}\left(\frac{Y-\mu}{\sigma}\right)^2} / \sqrt{2\pi\sigma^2} \qquad \text{(B-3)}$$

To convert this into (B-2), change units from Y to a function of T by substituting:

$$Y = A + \alpha B = (A + B)(1 - (1-\alpha)R) = T(1-(1-\alpha)R) \qquad \text{(B-4)}$$

where:

$R = \frac{B}{A+B} = \frac{B}{T}$ is the proportion of negative errors to total errors and $T = A + B$ is the total number of positive and negative errors.

Given the variable transformation from Y to $T(1-(1-\alpha)R)$, we have to introduce a Jacobian into the density function. Assuming that the ratio B/(A+B) is random or at least independent of T, the derivative is:

$$\frac{\partial Y}{\partial T} = (1-(1-\alpha)R) \qquad \text{(B-5)}$$

Substituting $B/(A+B)$ for R, and rearranging, equation (B-5) becomes:

$$\frac{\partial Y}{\partial T} = \frac{A+\alpha B}{A+B} \qquad \text{(B-6)}$$

Thus, changing Y into $A+\alpha B$ in (B-3), and incorporating the Jacobian into the expression, leads to (B-2).

After concentrating the likelihood for (B-2) and dropping constants, maximum likelihood requires that we maximize:

$$\sum \ln(A+\alpha B) - (A+\alpha B-\mu)^2 \qquad \text{(B-7)}$$

We used maximum likelihood to estimate α in the simple model, and we confirmed this solution using a grid search based on OLS, where (B-7) was the criterion function to select an α that maximized the likelihood. Both solutions are identical, but the first leads to an estimate for the standard error of α. Because this estimate of α was consistent for all the other models, we use the estimate for α in subsequent models.

Marginal Effect of Effort on Error Index

We defined *f(EFFORT)* found in (5-2) as:

$$f(EFFORT) = \beta_1 \times EFFORT + \beta_2 \times EFFORT^2 + \beta_3 \times TANFIMP \times PCT\_TANF \times EFFORT + \beta_4 \times NoTI \times PCT\_TANF \times EFFORT + \beta_5 \times CM13 \times EFFORT$$
(B-8)

Letting Z represent the marginal effect of effort on the error index, we defined Z as the partial derivative of (5-2) with respect to effort:

$$Z = \frac{\partial ErrorIndex}{\partial EFFORT} = \beta_1 + 2\beta_2 \times EFFORT + \beta_3 \times TANFIMP \times PCT\_TANF + \beta_4 \times NoTI \times PCT\_TANF + \beta_5 \times CM13 \quad \text{(B-9)}$$

We evaluated the variables that appear in *Z* at their yearly means. For the variance of *Z*, we used the following. Define *D* as:

$$D = \begin{bmatrix} \partial Z / \partial \beta_1 \\ \partial Z / \partial \beta_2 \\ \partial Z / \partial \beta_3 \\ \partial Z / \partial \beta_4 \\ \partial Z / \partial \beta_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \times EFFORT \\ TANFIMP \times PCT\_TANF \\ NoTI \times PCT\_TANF \\ CM13 \end{bmatrix} \quad \text{(B-10)}$$

Then, by the delta method, we computed the variance of Z as:

$$Var(Z) = D^T V D \quad \text{(B-11)}$$

where V is the estimated variance-covariance matrix of the β coefficients.

Estimation of Variable Contributions to the Marginal Effect of Effort on Error

Table 5-4 presents the contributions of several variables to the marginal effect of effort. The computation of these contributions used equation B-9. The values used in these computations and the steps in the computation are provided in table B-1.

Table B-1

Decomposition of Marginal Effect of Effort on Error

	1989-1996	1997-2002	2003-2005
β_1 (Coefficient of *EFFORT*)	-7.979	-7.979	-7.979
β_2 (Coefficient of *EFFORT_SQ*)	63.309	63.309	63.309
Mean*(EFFORT)*	0.010	0.012	0.010
$2\beta_2 \times EFFORT$	1.206	1.486	1.222
β_3 (Coefficient of *TANFIMP_TANF_EFFORT*)	17.811	17.811	17.811
Mean*(TANFIMP)*$\times$Mean*(TANF)*	0.000	0.227	0.156
$\beta_3 \times TANFIMP \times PCT\_TANF$	0.000	4.038	2.773
β_4 (Coefficient of *NOTI_TANF_EFFORT*)	2.640	2.640	2.640
Mean*(NOTI)*$\times$Mean*(TANF)*	0.360	0.013	0.000
$\beta_4 \times NoTI \times PCT\_TANF$	0.950	0.035	0.000
β_5 (Coefficient of *CM13_EFFORT*)	7.990	7.990	7.990
Mean*(CM13)*	0.062	0.154	0.063
$\beta_5 \times CM13$	0.496	1.233	0.506
Net Marginal Effect	-5.327	-1.188	-3.479

Appendix C: Alternate Models of Error Rates

This appendix presents the alternate models of error rates and their results. First, we present the results from using the base model (as presented in Chapter Five) with alternate measures of error as dependent variables in place of the error index. Next, we discuss results from alternate specifications of interactions with EFFORT and models with additional time-varying relationships to the error index. Finally, we describe models that use alternatives to the fixed effects generalized estimating equations (GEE) approach to correcting for serial correlation within States over time. Additional technical information on the assumptions, rationale, and computations for this last set of models is provided in Appendix D.

Alternative Measures of Error

The base model specification (as in table 5-2) for the control variables uses the error index. We also reestimated the model using alternative measures of certification error in place of the error index. Table C-1 shows the results using each of the following error measures as the dependent variable:

- Column A: base model of the error index (weighted sum of positive and negative case error rates)
- Column B: overpayment case error rate (percent of cases with overpayments or payment to ineligible household)[75]
- Column C: underpayment case error rate (percent of cases with underpayments)
- Column D: combined case error rate (unweighted sum of overpayment and underpayment case error rates)
- Column E: official overpayment rate (percent of benefits paid in overpayments or to ineligible households)[76]
- Column F: official underpayment rate (percent of benefits in underpayments)
- Column G: combined payment error rate (unweighted sum of official overpayment and underpayment rates).

[75] Case error rates were computed using the constant $25 threshold for countable errors, as in the error index.

[76] As previously discussed, the threshold for countable errors was $5 prior to 2000 and $25 thereafter. To compensate for this change, we included a dummy variable for years after 2000, expecting that the higher threshold would reduce the error rate. The dummy variable was significant at the .05 level for the underpayment and combined payment error models.

Table C-1

GEE Fixed Effects Parameter Estimates: Base Model of Error Index and Alternate Models of Case and Payment Error Rates

	Computed Case Error Rates				Official Payment Error Rates		
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
Variable	Base Model Error Index	Overpayment Case Error Rate	Underpayment Case Error Rate	Combined Case Error Rate	Official Overpayment Rate	Official Underpayment Rate	Combined Payment Error Rate
Dependent variable mean	*0.151*	*0.096*	*0.042*	*0.138*	*0.064*	*0.021*	*0.085*
INTERCEPT	0.177	0.204	-0.011	0.193	0.216	-0.037	0.179
	(0.228)	(0.149)	(0.090)	(0.211)	(0.145)	(0.041)	(0.169)
Y2000 Plus					-0.006*	-0.003**	-0.009**
					(0.003)	(0.001)	(0.004)
EFFORT	-7.979***	-5.345**	-2.280***	-7.625***	-3.248*	-1.306**	-4.554**
	(2.737)	(2.168)	(0.838)	(2.564)	(1.941)	(0.553)	(2.250)
EFFORT_SQ	63.309	38.344	21.081	59.425	-19.452	5.819	-13.633
	(82.268)	(69.739)	(29.881)	(77.704)	(60.795)	(15.890)	(67.220)
TANFIMP_TANF_EFFORT	17.811***	11.775***	5.332***	17.107***	9.859***	2.935***	12.794***
	(3.989)	(3.248)	(1.182)	(3.842)	(2.822)	(1.066)	(3.724)
NOTI_TANF_EFFORT	2.640	2.251	0.708	2.959	3.830	1.500	5.330
	(4.515)	(3.502)	(1.463)	(4.367)	(2.682)	(1.123)	(3.548)
CM13_EFFORT	7.990***	3.760*	3.117**	6.877***	4.081**	2.749***	6.830***
	(2.832)	(1.943)	(1.248)	(2.611)	(1.687)	(0.745)	(2.057)
TANFIMP	-0.019*	-0.014*	-0.004	-0.018*	-0.007	0.002	-0.005
	(0.010)	(0.008)	(0.003)	(0.010)	(0.006)	(0.002)	(0.007)
TANF	-0.037	-0.026	-0.010	-0.036	-0.051	-0.016	-0.067
	(0.065)	(0.051)	(0.016)	(0.062)	(0.043)	(0.015)	(0.057)
EARNINC	0.225***	0.105**	0.094***	0.199***	0.040	0.048***	0.087**
	(0.060)	(0.042)	(0.022)	(0.056)	(0.034)	(0.013)	(0.039)
SSINC	-0.009	-0.012	0.002	-0.010	0.021	0.024***	0.046
	(0.052)	(0.038)	(0.016)	(0.049)	(0.029)	(0.009)	(0.035)
SINGLEPAR	-0.010	-0.045	0.022	-0.022	-0.028	0.006	-0.022
	(0.043)	(0.032)	(0.014)	(0.040)	(0.029)	(0.009)	(0.034)
PCTEBT	-0.003	0.001	-0.003	-0.001	0.003	-0.002	0.002
	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.001)	(0.005)
FYUN0	0.123	0.068	0.045	0.112	-0.049	-0.005	-0.054
	(0.171)	(0.115)	(0.061)	(0.157)	(0.092)	(0.032)	(0.101)
CM13	-0.184***	-0.097***	-0.065***	-0.162***	-0.086***	-0.043***	-0.129***
	(0.042)	(0.031)	(0.014)	(0.038)	(0.025)	(0.009)	(0.029)
CM1299ER	0.001	0.008	-0.005	0.004	0.005	-0.003	0.002
	(0.010)	(0.007)	(0.003)	(0.010)	(0.005)	(0.002)	(0.007)

Table C-1

GEE Fixed Effects Parameter Estimates: Base Model of Error Index and Alternate Models of Case and Payment Error Rates

	Computed Case Error Rates				Official Payment Error Rates		
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
Variable	Base Model Error Index	Overpayment Case Error Rate	Underpayment Case Error Rate	Combined Case Error Rate	Official Overpayment Rate	Official Underpayment Rate	Combined Payment Error Rate
PCT_CHCASELOAD	0.016	0.007	0.003	0.010	0.006	-0.002	0.004
	(0.010)	(0.008)	(0.004)	(0.010)	(0.009)	(0.002)	(0.009)
PCT_MET	-0.022	0.027	-0.035*	-0.008	0.052**	-0.017	0.035
	(0.047)	(0.033)	(0.020)	(0.042)	(0.027)	(0.013)	(0.031)
PCTBLACK	0.115	-0.080	0.137	0.057	-0.188	0.122**	-0.067
	(0.296)	(0.188)	(0.119)	(0.274)	(0.182)	(0.055)	(0.214)
PCTHISP	-0.382	-0.415	-0.014	-0.430	-0.500*	0.111	-0.389
	(0.490)	(0.302)	(0.186)	(0.445)	(0.270)	(0.080)	(0.320)
PCT_SIMPREP	-0.036***	-0.026***	-0.009***	-0.035***	-0.015***	-0.005***	-0.020***
	(0.007)	(0.004)	(0.002)	(0.006)	(0.004)	(0.001)	(0.005)
PCT_QUARTREP	-0.012*	-0.008	-0.004	-0.012*	-0.004	-0.003**	-0.006
	(0.007)	(0.005)	(0.003)	(0.007)	(0.004)	(0.001)	(0.005)
PCT_TRANBEN	-0.027***	-0.019***	-0.006**	-0.025***	-0.013**	-0.002	-0.015**
	(0.010)	(0.006)	(0.003)	(0.009)	(0.005)	(0.002)	(0.007)
AFDCWAIV	0.011**	0.006	0.004**	0.010*	0.004	0.002	0.005
	(0.006)	(0.004)	(0.002)	(0.005)	(0.003)	(0.001)	(0.004)
SIMPINCOME	-0.009	-0.004	-0.004	-0.008	-0.005	-0.002	-0.006
	(0.008)	(0.006)	(0.003)	(0.008)	(0.004)	(0.002)	(0.006)
R^2	0.7412	0.8406	0.8231	0.8580	0.8061	0.8170	0.8288

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

Unlike the error index used in the base model, all of the alternative error rates exclude negative action errors, both from the numerator and from the denominator. For columns B, C, and D, the denominator is the number of active cases. For columns E, F, and G, the denominator is the total benefits issued to active cases. Each column shows the mean of the dependent variable in addition to the parameter estimates.

All of the control variables that were significant (at the 5 percent level or better) in the base model of the error index were also significant and had similar effects in the models of the other combined error rates. The significant control variables in the base model also had similar effects and significance levels in one or both of the models of individual error rates. The trends in the marginal effect of effort on error (not shown) were also very similar for all the error rates.

There were a few significant variables in the base error index model (column A) that had lower significance levels or were not significant in the alternative model of the overpayment case error rate (column B) or the official overpayment rate (column E). Most notably, in the official overpayment rate model, the EFFORT variable had a lower significance level (0.10), and the percentage of FSP households with earned income (EARNINC) did not have a significant estimated effect. The interaction of EFFORT with CM13 had a significance level of 0.10 in the overpayment case error model. All of these variables were significant at the 0.01 level in the base error index model. In the official underpayment rate model, the percentage of States with transitional benefits for TANF leavers (PCT_TRANBEN) was not significant, while the percentage of FSP households with Social Security or SSI income (SSINC) was significant at the 0.01 level. Both of these results differed from the base error index model. These differences across dependent variables suggest that there may be some factors that affect overpayments but not underpayments, or vice versa. We note that the differences across models were relatively small. Differences in statistical significance do not always indicate substantive differences, as the significance cutoffs are arbitrary.

Alternative Specifications of Effort Variables

Table C-2 compares the base model (in the second column) with two alternative specifications of *f(EFFORT)*, i.e., the relationship of effort to the error index. In the base model, the variables that captured the effects of effort on the error index were EFFORT, the square of EFFORT (EFFORT_SQ), the interaction of EFFORT with the percentage of households with short certifications (CM13_EFFORT), and the interactions of EFFORT with the percentage of households receiving AFDC/TANF with and without TANF implementation (TANFIMP_TANF_EFFORT and NOTI_TANF_EFFORT). A key question was whether the percentage of FSP households receiving AFDC/TANF interacted with EFFORT.

Therefore, alternative specification 1 eliminated the hypothesis that the percentage of FSP households with AFDC/TANF (PCT_TANF) interacted with EFFORT before and after the enactment of PRWORA and the TANF program. The simpler model for alternative specification 1 instead posed the hypothesis that the net effect of EFFORT changed with TANF implementation, and that the percentage of FSP households with AFDC/TANF only affected the level of the error index. In the third column of table C-2, estimates for alternative specification 1 are shown. In place of the TANFIMP_TANF_EFFORT and NoTI_TANF_EFFORT interactions, this specification included a variable computed by interacting the indicator for TANF implementation (TANFIMP) with EFFORT

Table C-2

GEE Fixed Effects Parameter Estimates: Alternative Specifications of Error as a Function of Effort

Variable	Base	f(EFFORT) Specification 1	f(EFFORT) Specification 2
INTERCEPT	0.177	0.142	0.178
	(0.228)	(0.242)	(0.232)
EFFORT	-7.979***	-8.829***	-5.732
	(2.737)	(2.483)	(3.557)
EFFORT_SQ	63.309	192.862**	91.992
	(82.268)	(86.382)	(83.893)
TANFIMP_EFFORT	-	1.865	-2.782
		(1.425)	(2.058)
TANFIMP_TANF_EFFORT	17.811***	-	15.640***
	(3.989)		(4.511)
NOTI_TANF_EFFORT	2.640	-	-3.585
	(4.515)		(7.425)
CM13_EFFORT	7.990***	5.968*	8.576***
	(2.832)	(3.401)	(2.736)
TANFIMP	-0.019*	0.007	-0.006
	(0.010)	(0.012)	(0.012)
PCT_TANF	-0.037	0.108***	-0.004
	(0.065)	(0.039)	(0.068)
CM13	-0.184***	-0.161***	-0.189***
	(0.042)	(0.050)	(0.042)
CM1299ER	0.001	0.006	0.001
	(0.010)	(0.010)	(0.010)
PCT_SIMPREP	-0.036***	-0.038***	-0.036***
	(0.007)	(0.008)	(0.007)
PCT_QUARTREP	-0.012*	-0.018**	-0.012
	(0.007)	(0.008)	(0.007)
PCT_TRANBEN	-0.027***	-0.033***	-0.026***
	(0.010)	(0.009)	(0.009)
AFDCWAIV	0.011**	0.009	0.009*
	(0.006)	(0.006)	(0.006)
SIMPINCOME	-0.009	-0.009	-0.010
	(0.008)	(0.009)	(0.008)
EARNINC	0.225***	0.264***	0.230***
	(0.060)	(0.067)	(0.060)
SSINC	-0.009	-0.013	-0.010
	(0.052)	(0.053)	(0.052)
SINGLEPAR	-0.010	-0.022	-0.011
	(0.043)	(0.045)	(0.042)
PCTEBT	-0.003	-0.005	-0.002
	(0.006)	(0.007)	(0.006)
FYUN0	0.123	0.002	0.148
	(0.171)	(0.168)	(0.174)
PCT_CHCASELOAD	0.016	0.006	0.018*
	(0.010)	(0.011)	(0.010)
PCT_MET	-0.022	-0.060	-0.024
	(0.047)	(0.048)	(0.047)
PCTBLACK	0.115	0.139	0.102
	(0.296)	(0.311)	(0.304)
PCTHISP	-0.382	-0.287	-0.439
	(0.490)	(0.511)	(0.502)
R^2	0.7412	0.7230	0.7423

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

(TANFIMP_EFFORT). The TANFIMP_EFFORT variable had a significant positive relationship to the error index in the authors' previous study. In the estimates of alternative specification 1 for the current study, the results were mostly the same as for the base model. However, the coefficient for the coefficient for TANFIMP_EFFORT was not statistically different from zero, but the coefficient for EFFORT_SQ was positive and significant. (EFFORT rose during the period when PCT_TANF fell by the greatest amount, so EFFORT_SQ may be proxying for TANFIMP_TANF_EFFORT in this specification.) In addition, the coefficient for the PCT_TANF variable was positive and significant. The results from the base model suggest that the coefficients on EFFORT_SQ and PCT_TANF became not statistically different from zero when the TANFIMP_TANF_EFFORT variable was added.

In the fourth column of table C-2, estimates are shown for alternative specification 2, with the TANFIMP_EFFORT variable added to the base specification of *f(EFFORT)*. This model provided a further test of whether TANFIMP_EFFORT should be interacted with PCT_TANF. The coefficient for TANFIMP_EFFORT was not statistically significant; neither was the coefficient of EFFORT. (The coefficient for EFFORT may have been affected by multicollinearity with some combination of TANFIMP_EFFORT, TANFIMP_TANF_EFFORT, and NOTI_TANF_EFFORT). On the other hand, the coefficient for TANFIMP_TANF_EFFORT remained highly significant, and the other significant variables in the base model were significant in alternative specification 1, except for the AFDC reform waiver indicator (AFDCWAIV). Therefore, we concluded that the base model specification with the TANFIMP_TANF_EFFORT variable was preferable to the alternative specifications, and to the previous study's model with the TANFIMP_EFFORT variable, which left out the role of the percentage of FSP households with TANF (PCT_TANF).

Alternative Models with Additional Time-Varying Relationships

Table C-3 compares the base model with three alternative models that introduce additional time-varying relationships. The previous study raised the question of whether the effects of TANF implementation on the error index were transient disturbances or more long-term.

To control for potential unobserved year-to-year variation, year dummy variables were added to the base model. (The base model estimates are shown in the second column of table C-3. Not shown in this or other exhibits are the state-specific time trend variables that were used in all models, as discussed in Chapter Five.) The results for this specification are shown in the third column of table C-3. None of the dummy variables had a significant coefficient at the 5 percent level, and only the coefficients for the 2002 and 2003 year dummy variables were significant at the 10 percent level. Thus, there was no strong evidence of an effect associated with a specific year. We tested a linear hypothesis that the parameter estimates for the dummy variables were jointly zero, using a score test.[77] The test statistic was χ^2 (15) = 25.8, p=0.040. Although the p-value was less than 0.05, the improvement in model fit was small (the R^2 increased by 0.01), and the addition of the dummy variables did not qualitatively affect any results of interest. Estimated coefficients for the EFFORT variables were larger in absolute value, but the confidence intervals overlapped with those of the base model estimates. Since the main findings for the coefficients of the variables affecting the level of error (the X_{it} variables) remained consistent, we elected to retain the base specification.

---

[77] This test was generated with the SAS 9.1 CONTRAST statement in PROC GENMOD.

Table C-3

GEE Fixed Effects Parameter Estimates: Alternative Specifications of Time-Varying Relationships

Variable	Base	Base with Year Dummies	Three Periods in TANFIMP	Year Dummies in TANFIMP
INTERCEPT	0.177	0.195	0.223	0.251
	(0.228)	(0.285)	(0.232)	(0.240)
EFFORT	-7.979***	-11.319***	-7.511***	-8.547***
	(2.737)	(2.995)	(2.718)	(2.684)
EFFORT_SQ	63.309	131.222*	52.165	80.939
	(82.268)	(78.421)	(80.374)	(79.115)
TANFIMP_TANF_EFFORT	17.811***	19.244***	-	-
	(3.989)	(4.574)		
NOTI_TANF_EFFORT	2.640	10.271*	2.793	3.892
	(4.515)	(5.730)	(4.612)	(4.688)
TI9702_TANF_EFFORT	-	-	16.690***	-
			(3.909)	
TI0305_TANF_EFFORT	-	-	11.827**	-
			(4.714)	
TI97_TANF_EFFORT	-	-	-	13.121***
				(4.590)
TI98_TANF_EFFORT	-	-	-	18.270***
				(4.438)
TI99_TANF_EFFORT	-	-	-	18.840***
				(4.317)
TI00_TANF_EFFORT	-	-	-	15.248***
				(4.504)
TI01_TANF_EFFORT	-	-	-	17.014***
				(4.683)
TI02_TANF_EFFORT	-	-	-	17.637***
				(4.059)
TI03_TANF_EFFORT	-	-	-	14.574***
				(4.707)
TI04_TANF_EFFORT	-	-	-	9.192*
				(5.567)
TI05_TANF_EFFORT	-	-	-	10.543
				(7.193)
CM13_EFFORT	7.990***	9.801***	7.862***	7.775***
	(2.832)	(3.108)	(2.821)	(3.007)
TANFIMP	-0.019*	-	-0.016	-0.013
	(0.010)		(0.010)	(0.012)
PCT_TANF	-0.037	-0.083	-0.038	-0.038
	(0.065)	(0.074)	(0.065)	(0.064)
CM13	-0.184***	-0.203***	-0.182***	-0.181***
	(0.042)	(0.046)	(0.041)	(0.045)
CM1299ER	0.001	0.001	-0.000	-0.002
	(0.010)	(0.010)	(0.010)	(0.010)
PCT_SIMPREP	-0.036***	-0.026***	-0.034***	-0.033***
	(0.007)	(0.009)	(0.007)	(0.007)
PCT_QUARTREP	-0.012*	-0.013*	-0.011	-0.009
	(0.007)	(0.007)	(0.007)	(0.007)
PCT_TRANBEN	-0.027***	-0.025***	-0.023**	-0.020*
	(0.010)	(0.009)	(0.010)	(0.011)
AFDCWAIV	0.011**	0.014**	0.010*	0.011**
	(0.006)	(0.006)	(0.006)	(0.006)
SIMPINCOME	-0.009	0.001	-0.008	-0.008
	(0.008)	(0.009)	(0.008)	(0.008)
EARNINC	0.225***	0.194***	0.214***	0.198***
	(0.060)	(0.060)	(0.060)	(0.062)

Table C-3

GEE Fixed Effects Parameter Estimates: Alternative Specifications of Time-Varying Relationships

Variable	Base	Base with Year Dummies	Three Periods in TANFIMP	Year Dummies in TANFIMP
SSINC	-0.009	-0.041	-0.020	-0.026
	(0.052)	(0.049)	(0.052)	(0.052)
SINGLEPAR	-0.010	-0.031	-0.013	-0.020
	(0.043)	(0.046)	(0.043)	(0.045)
PCTEBT	-0.003	-0.003	-0.002	-0.003
	(0.006)	(0.007)	(0.006)	(0.007)
FYUN0	0.123	-0.046	0.138	0.118
	(0.171)	(0.200)	(0.172)	(0.188)
PCT_CHCASELOAD	0.016	0.011	0.018*	0.011
	(0.010)	(0.012)	(0.010)	(0.012)
PCT_MET	-0.022	0.020	-0.021	0.003
	(0.047)	(0.050)	(0.046)	(0.047)
PCTBLACK	0.115	0.106	0.049	-0.001
	(0.296)	(0.352)	(0.303)	(0.315)
PCTHISP	-0.382	-0.225	-0.468	-0.555
	(0.490)	(0.607)	(0.499)	(0.529)
Y1989	-	-0.015	-	-
		(0.068)		
Y1990	-	-0.006	-	-
		(0.064)		
Y1991	-	-0.005	-	-
		(0.059)		
Y1992	-	0.003	-	-
		(0.055)		
Y1993	-	0.008	-	-
		(0.050)		
Y1994	-	0.010	-	-
		(0.047)		
Y1995	-	0.006	-	-
		(0.042)		
Y1996	-	0.004	-	-
		(0.037)		
Y1997	-	0.006	-	-
		(0.032)		
Y1998	-	0.019	-	-
		(0.028)		
Y1999	-	0.025	-	-
		(0.024)		
Y2000	-	0.013	-	-
		(0.021)		
Y2001	-	0.018	-	-
		(0.017)		
Y2002	-	0.022*	-	-
		(0.012)		
Y2003	-	0.016**	-	-
		(0.007)		
R^2	0.7412	0.7531	0.7434	0.7490

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

The models in the fourth and fifth columns of table C-3 refined the time period covered by the TANFIMP_TANF_EFFORT variable. The “Three Periods in TANFIMP” and “Year Dummies in TANFIMP” models tested whether the effect of the TANFIMP_TANF_EFFORT interaction diminished over time. The “Three Periods in TANFIMP” model (Column 4) broke TANFIMP_TANF_EFFORT into two time periods: 1997-2002 (TI9702_TANF_EFFORT) and 2003-2005 (TI0305_TANF_EFFORT). The “Year Dummies in TANFIMP” model (Column 5) interacted TANF_EFFORT with each year from 1997 to 2005 (TI97_TANF_EFFORT to TI05_TANF_EFFORT). In both cases, the estimated coefficients for the 1997-2002 period were larger than for the 2003-2005 period, but we found little statistical evidence that the coefficient for the TANFIMP_TANF_EFFORT interaction variable changed over time. We note, however, that the periods were relatively short, and several other key variables changed, making it difficult to detect a difference between the periods.

Marginal Effect Estimates for Base and Alternative Models

The estimated marginal effect of effort on the error index was remarkably similar across specifications of the GEE fixed effects model, as shown in table C-4. In the alternative models, as in the base model, the marginal effect of effort was statistically identical from 1989 to 1996, decreased markedly in absolute value in 1997-1998, and returned to a level that was not statistically different from the pre-1997 level by 2005. Estimated marginal effects for 1989-1996 did not differ across specifications. The base model with added year dummy variables (column 5) produced larger estimates (in absolute value) than the base model for 1998 through 2005, as did the alternate specification with TANFIMP_EFFORT in place of TANFIMP_TANF_EFFORT (column 3). With a few exceptions, however, the standard errors were too large (relative to the estimates) for differences across models to be detectable. Thus, we did not find evidence from the marginal effects that any of the alternate specifications yielded materially different results.

Alternative Models using the Base Specification

After settling on the base model and checking its robustness to alternative specification, we explored other models for dealing with the potential bias arising from the correlation of panel data within States over time. While the base fixed effects model used generalized estimating equations (GEE), the alternative models support the results we found with the base model. As discussed below, the GEE approach is a more robust specification and therefore we chose it as the best specification from which to draw inference. Table C-5 reports the coefficients from the alternative models, and table C-6 shows the estimated marginal effect of effort on error over time.

Prais-Winsten AR(1) Models

The Prais-Winsten estimate is a Feasible Generalized Least Squares (FGLS) estimator that works to “sweep” first-order autocorrelation from the model. The estimator requires two steps. First, using the fixed effects model, the analyst estimates the regression residuals. Since these are consistent measures of the error term (e_{it}), they are used to estimate the autocorrelation coefficient ρ. The estimate of ρ is used to transform the dependent variable and every independent variable, as described in Appendix D.

Table C-4

GEE Fixed Effects Marginal Effect of Effort on Error: Base Model and Alternative Specifications

Year	Base	f(EFFORT) Specification 1	f(EFFORT) Specification 2	Base with Year Dummies	Three Periods in TANFIMP	Year Dummies in TANFIMP
1989	-5.763***	-6.213***	-5.493***	-5.321***	-5.375***	-5.666***
	(1.136)	(1.281)	(1.164)	(1.181)	(1.178)	(1.149)
1990	-5.744***	-6.355***	-5.594***	-5.233***	-5.344***	-5.642***
	(1.181)	(1.296)	(1.193)	(1.240)	(1.228)	(1.212)
1991	-5.852***	-6.541***	-5.678***	-5.467***	-5.444***	-5.775***
	(1.184)	(1.326)	(1.208)	(1.232)	(1.231)	(1.228)
1992	-6.087***	-6.965***	-5.982***	-5.832***	-5.657***	-6.033***
	(1.285)	(1.435)	(1.316)	(1.325)	(1.329)	(1.350)
1993	-6.060***	-6.884***	-5.915***	-5.814***	-5.636***	-6.007***
	(1.251)	(1.408)	(1.284)	(1.290)	(1.295)	(1.312)
1994	-6.014***	-6.815***	-5.756***	-5.875***	-5.593***	-5.983***
	(1.216)	(1.404)	(1.268)	(1.241)	(1.257)	(1.270)
1995	-5.860***	-6.671***	-5.576***	-5.688***	-5.444***	-5.832***
	(1.196)	(1.390)	(1.249)	(1.218)	(1.237)	(1.245)
1996	-5.761***	-6.532***	-5.324***	-5.715***	-5.353***	-5.763***
	(1.148)	(1.382)	(1.241)	(1.153)	(1.185)	(1.186)
1997	-2.391***	-4.955***	-2.769***	-3.757***	-2.273**	-3.676***
	(0.915)	(0.955)	(0.925)	(0.920)	(0.925)	(1.068)
1998	-1.347	-3.964***	-2.050**	-3.273***	-1.319	-1.627
	(0.899)	(0.979)	(0.893)	(0.926)	(0.914)	(0.989)
1999	-1.596*	-3.596***	-2.198***	-3.459***	-1.551*	-1.743*
	(0.866)	(0.925)	(0.839)	(0.908)	(0.877)	(0.997)
2000	-1.769**	-3.384***	-2.303***	-3.588***	-1.712*	-2.714***
	(0.877)	(0.912)	(0.842)	(0.925)	(0.885)	(0.878)
2001	-2.124**	-3.457***	-2.634***	-3.987***	-2.049**	-2.668***
	(0.888)	(0.906)	(0.850)	(0.953)	(0.896)	(0.860)
2002	-2.734***	-3.942***	-3.281***	-4.770***	-2.629***	-3.160***
	(0.936)	(0.952)	(0.895)	(1.036)	(0.948)	(0.995)
2003	-3.534***	-4.395***	-4.084***	-5.754***	-4.210***	-4.490***
	(1.078)	(1.068)	(1.034)	(1.217)	(1.210)	(1.243)
2004	-4.180***	-4.800***	-4.746***	-6.551***	-4.755***	-5.930***
	(1.209)	(1.192)	(1.164)	(1.372)	(1.320)	(1.299)
2005	-4.556***	-4.984***	-5.116***	-7.007***	-5.042***	-6.005***
	(1.305)	(1.272)	(1.259)	(1.482)	(1.399)	(1.441)

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

Table C-5

Alternative Models using the Base Specification

Variable	Prais-Winsten AR(1) - State-Specific Autocorrelation	Prais-Winsten AR(1) - Common Autocorrelation	Long-Run Partial Adjustment	Long-Run Arellano-Bond
INTERCEPT	0.326*	0.190	-0.150	-0.009***
	(0.187)	(0.195)	(0.297)	(0.003)
EFFORT	-6.788***	-6.213**	-10.545***	-10.757**
	(2.411)	(2.606)	(3.694)	(4.573)
EFFORT_SQ	42.130	18.085	13.441	-11.855
	(88.615)	(90.695)	(118.542)	(128.631)
TANFIMP_TANF_EFFORT	15.757***	15.793***	23.186***	23.167***
	(3.423)	(3.375)	(5.491)	(7.553)
NOTI_TANF_EFFORT	1.859	1.527	8.241	10.566
	(3.487)	(3.617)	(6.759)	(7.722)
CM13_EFFORT	7.816***	7.192**	11.820***	12.079**
	(2.891)	(3.502)	(3.932)	(5.946)
TANFIMP	-0.014	-0.018*	-0.004	-0.001
	(0.010)	(0.010)	(0.015)	(0.017)
PCT_TANF	-0.011	-0.025	-0.066	-0.029
	(0.037)	(0.037)	(0.081)	(0.098)
CM13	-0.166***	-0.168***	-0.232***	-0.254***
	(0.040)	(0.049)	(0.066)	(0.086)
CM1299ER	0.004	0.003	0.010	0.019
	(0.008)	(0.009)	(0.013)	(0.018)
PCT_SIMPREP	-0.035***	-0.034***	-0.047***	-0.052***
	(0.006)	(0.007)	(0.010)	(0.013)
PCT_QUARTREP	-0.013**	-0.013**	-0.015	-0.014
	(0.005)	(0.006)	(0.009)	(0.010)
PCT_TRANBEN	-0.025***	-0.026***	-0.037***	-0.033*
	(0.007)	(0.008)	(0.014)	(0.019)
AFDCWAIV	0.010**	0.010	0.014**	0.010
	(0.005)	(0.006)	(0.007)	(0.010)
SIMPINCOME	-0.011**	-0.010**	-0.005	-0.014
	(0.005)	(0.005)	(0.010)	(0.012)
EARNINC	0.219***	0.217***	0.314***	0.507***
	(0.041)	(0.046)	(0.087)	(0.108)
SSINC	-0.022	-0.019	-0.028	0.053
	(0.033)	(0.036)	(0.077)	(0.110)
SINGLEPAR	-0.006	0.016	0.044	0.130**
	(0.029)	(0.032)	(0.059)	(0.062)
PCTEBT	0.001	-0.001	0.004	0.011
	(0.004)	(0.005)	(0.007)	(0.010)
FYUN0	0.208	0.159	0.116	0.621*
	(0.153)	(0.160)	(0.237)	(0.329)
PCT_CHCASELOAD	0.012	0.012	0.035**	0.021
	(0.013)	(0.013)	(0.015)	(0.022)
PCT_MET	-0.039	-0.054	-0.083	-0.123
	(0.053)	(0.057)	(0.083)	(0.108)
PCTBLACK	-0.108	0.081	0.573	1.152**
	(0.230)	(0.250)	(0.365)	(0.451)
PCTHISP	-0.570*	-0.306	0.400	1.368**
	(0.306)	(0.318)	(0.595)	(0.696)
ψ (Lagged ERROR)[a]	-	-	0.359***	0.404***
			(0.046)	(0.054)

Table C-5

Alternative Models using the Base Specification

Variable	Prais-Winsten AR(1) - State-Specific Autocorrelation	Prais-Winsten AR(1) - Common Autocorrelation	Long-Run Partial Adjustment	Long-Run Arellano-Bond
R-squared	0.8632	0.6608	0.7853	-
Goodness of fit - Corr(Error,Xb)[b]	-	-	-	0.6213
2nd Order Serial Correlation	-	-	-	z = 1.27 p = 0.204

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10
[a] ERROR lagged one year in Partial Adjustment model. Arellano-Bond model uses ERROR lagged 2 and 3 years as instruments for one-year lagged ERROR.
[b] Squared correlation of short-term predictions with actual error rates.

Table C-6

Alternative Models: Estimated Marginal Effects of Effort on Error

Year	Prais-Winsten AR(1) - State-Specific Autocorrelation	Prais-Winsten AR(1) - Common Autocorrelation	Long-Run Partial Adjustment	Long-Run Arellano-Bond
	-5.125***	-4.994***	-6.572***	-6.187***
1989	(0.916)	(0.921)	(1.748)	(1.877)
	-5.101***	-4.957***	-6.405***	-5.961***
1990	(0.900)	(0.935)	(1.756)	(1.901)
	-5.184***	-5.015***	-6.534***	-6.088***
1991	(0.872)	(0.933)	(1.684)	(1.898)
	-5.381***	-5.162***	-6.738***	-6.250***
1992	(0.923)	(1.025)	(1.687)	(1.997)
	-5.359***	-5.148***	-6.748***	-6.279***
1993	(0.900)	(0.994)	(1.663)	(1.963)
	-5.301***	-5.092***	-6.752***	-6.320***
1994	(0.877)	(0.978)	(1.603)	(1.945)
	-5.153***	-4.959***	-6.546***	-6.117***
1995	(0.862)	(0.970)	(1.593)	(1.935)
	-5.044***	-4.857***	-6.508***	-6.130***
1996	(0.832)	(0.951)	(1.522)	(1.913)
	-1.952**	-1.716**	-3.325***	-3.532**
1997	(0.766)	(0.848)	(1.243)	(1.721)
	-0.965	-0.739	-2.457**	-2.912*
1998	(0.789)	(0.874)	(1.239)	(1.755)
	-1.179	-0.995	-2.824**	-3.301*
1999	(0.777)	(0.845)	(1.187)	(1.723)
	-1.330*	-1.171	-3.074***	-3.563**
2000	(0.801)	(0.851)	(1.192)	(1.730)
	-1.653**	-1.487*	-3.552***	-4.044**
2001	(0.809)	(0.854)	(1.204)	(1.769)
	-2.210***	-1.989**	-4.329***	-4.803**
2002	(0.809)	(0.880)	(1.265)	(1.891)
	-2.928***	-2.656***	-5.338***	-5.790***
2003	(0.892)	(0.983)	(1.442)	(2.129)
	-3.520***	-3.203***	-6.184***	-6.627***
2004	(0.987)	(1.077)	(1.606)	(2.339)
	-3.854***	-3.516***	-6.652***	-7.083***
2005	(1.059)	(1.155)	(1.728)	(2.485)

Notes: Standard Errors are in parentheses. *** - p<=0.01; ** - 0.01<p<=0.05; * - 0.05<=p<=0.10

The Prais-Winsten model was estimated in two ways—first, with a common estimate of the ρ parameter across States, and second with a state-specific estimate of ρ. Relative to a Prais-Winsten model using a common ρ, a model using state-specific estimates of ρ may produce less biased estimates if the autocorrelation parameters are not equal across States. It may be less efficient, however, because it requires additional parameter estimates. Because of this trade-off, we present and compare results using both Prais-Winsten models. Our panels are quite short (17 periods) to be estimating separate parameters for each state. On a priori grounds, the pooled estimator seems preferable.

The second and third columns in tables 5-7 and 5-8 report results for the fixed effects specification using Prais-Winsten AR(1) models rather than a GEE to control for within-State correlations. The Prais-Winsten model can be more efficient if the AR(1) assumptions are met, but standard errors may not be consistently estimated if they are not.[78] The second column reports parameter estimates when computing separate correlation parameters for each state, while the third column reports parameter estimates when computing a common correlation parameter across all states.

The results were similar between the state-specific and common correlation parameters. The coefficients of the X_{it} variables were consistent in size and significance to the fixed effects GEE coefficients. The coefficients on the EFFORT, TANFIMP_TANF_EFFORT, and CM13_EFFORT variables were smaller in absolute value, and the estimated marginal effects of effort were slightly smaller for the Prais-Winsten models versus the GEE. However, the confidence intervals of estimates overlapped, and the changes in the marginal effects over time were consistent across all of the models.

Although there was a reduction of around 20 percent in the standard errors, given the similarities in conclusions between the GEE and the Prais-Winsten models, this reduction is only appropriate if the autoregression is only AR(1). If it was AR(2) or higher, the Prais Winston standard errors are incorrect. Since we have no a priori reason to exclude AR(2) or higher, we prefer the more robust GEE model. Bertrand et al. (2004) argue that, in the presence of serial correlation higher than order one, the standard errors will be too small when fitting an AR(1) model to the data, as in the Prais-Winsten estimators. An estimator that models an arbitrary correlation of within-State error terms, such as the GEE, performs well when the number of groups (in our case, States) is "sufficiently large." In a similar model using State-level data, Bertrand et al. suggest that the group of all 50 States is "sufficiently large."

Partial Adjustment Model

The previous models treat the availability of resources as exogenous to the error index. A dynamic model, however, might better explain the data generating process. One form of dynamic model that we employed is the partial adjustment model.

The partial adjustment model assumes that States adjust their resources so as to achieve a desired level of errors, but only make these adjustments gradually by a fraction (ψ), closing part of the gap between the actual and target error index each year. This partial adjustment could reflect budget

[78] The AR(1) specification assumes that there is a correlation between the present year and the previous one, but not between the present year and any earlier year (2nd order or higher).

constraints, or the time needed to implement process improvements. Details on the calculation of the parameters and their standard errors are found in Appendix D.

The simple partial adjustment model estimator may have several problems. First, if the residuals are autocorrelated, the estimated standard errors may be inconsistent, leading to misleading hypothesis testing. Second, if the time series does not satisfy stationarity, coefficient estimates will be biased and inconsistent. When this model was estimated in the previous study, however, the process was not found to be autocorrelated, and it was clearly stationary.

More importantly, the partial adjustment model is biased because the lagged value of the dependent variable is necessarily correlated with the error term. The model is consistent as the number of time periods approaches infinity, but with a period of 17 years, we expect a negative bias for a positive coefficient on the order of 1/17 (Baltagi, 1995, p. 126).

The fourth column of table C-5 reports long-run estimates and standard errors for the partial adjustment model, comparable to the estimates reported for the GEE fixed effects model. Estimates were computed with clustering on States, as in the GEE model, to correct the standard errors for autocorrelation. The additional parameter reported, ψ, is the parameter associated with the lagged dependent variable appearing in the model. One minus ψ estimates the rate of convergence of the short-run to the long-run; in this case, in any year the error rate achieves roughly 60 percent of its estimated long-run error rate.

The results for the partial adjustment model were remarkably similar to those from the GEE fixed effects model. The significant variables in the GEE model were all significant in the partial adjustment model. Almost all of the coefficients on the statistically significant X_{it} variables were larger in absolute value in the partial adjustment model than the fixed effects GEE. The exception was the coefficient on *PCT_QUARTREP*, which is roughly the same across models. In addition, the coefficients on the statistically significant effort variables *EFFORT*, *TANFIMP_TANF_EFFORT*, and *CM13_EFFORT* were all larger in absolute value in the partial adjustment model than the fixed effects GEE. The differences between models in the estimated coefficients were less than the width of the confidence intervals, so we could not detect a difference between the estimates. As shown in the fourth column in table C-6, the trends in the marginal effect of effort on error were similar between the partial adjustment and fixed effects GEE models; the levels were larger in absolute value for the partial adjustment model but the differences do not appear to be material, give the size of the standard errors of the estimates.

Arellano-Bond Dynamic Model

The Arellano-Bond (Arellano and Bond, 1991) approach provides a dynamic model as an alternative to the partial adjustment model. The Arellano-Bond model uses instrumental variables to surmount the problem of bias and inconsistency introduced when using the lagged dependent variable as a regressor. This method uses dependent variables lagged two and three periods as instruments for the one-period lagged dependent variable. The instruments are obtained in a dynamic panel model by using orthogonality conditions between lagged values of ERROR and the disturbances, v_{it} (Baltagi, 1995). The model is estimated via generalized method of moments (GMM) using the STATA command xtabond.[79] Details on the calculation of the parameters and their standard errors are found in Appendix D.

Using the Arellano-Bond model has two benefits. First, the model allows for first-order serial correlation of the error terms, although it requires the absence of second-order and higher serial correlation. Second, it uses an instrumental variable method to account for the fact that one-period lagged values of the dependent variable will be correlated with the residuals. A drawback of the Arellano-Bond estimator, however, is that it requires three years of data to be dropped to utilize the instruments. The loss of information results in less efficient parameter estimates.

The Arellano-Bond model was run with variations in the treatment of EFFORT and the use of additional instruments. Estimates were produced with and without using the variables other than the fixed effects and EFFORT variables as instruments, and with and without EFFORT as a predetermined variable. In general, the coefficients were not statistically different between those options. We chose the model with the additional instruments and EFFORT as predetermined because it improved the tests of 2nd order serial correlation.

The fifth column in tables C-6 and C-7 reports long-run estimates and standard errors for the Arellano-Bond model. At roughly 0.4, the coefficient for the lagged error index (ψ) was similar to that estimated in the partial-adjustment model. We did not find significant second-order serial correlation in a test of the first-differential residuals ($p=0.204$), so the GMM requirement of a zero second-order correlation appears to be met.[80] For a goodness-of-fit measure we report the squared correlation between the predicted and observed error rates, as suggested by Windmeijer (1995).

The results for the Arellano-Bond model were very similar to the partial adjustment model and confirmed the findings from the base model as to which variables were significant and the direction of their association with the error index. Compared to the fixed effects GEE model, in the Arellano-Bond model all of the statistically significant coefficients were larger in absolute value, as were the estimated marginal effects of effort on error. For most of the coefficients and all of the marginal effects, these differences did not appear to be significant (based on the size of the standard errors), but the coefficient for EARNINC was significantly larger in the Arellano-Bond model. Three additional variables had positive and statistically significant coefficients at the 5 percent level: the percent of households with children that have one adult (SINGLEPAR), the percent of households that are black (*PCT_BLACK*) and the percent of households that are Hispanic (*PCT_HISP*). However, since the coefficients on these variables were not statistically significant in any of the other models, the significant coefficients in the Arellano-Bond model should probably be viewed with skepticism.

[79] Details of the model are provided in Baltagi (1995) and in Appendix D.

[80] The test statistic was computed by STATA.

Appendix D: Explanation of Alternative Models

This appendix provides technical details regarding the assumptions, rationale, and computation of the alternative models presented in the last section of Appendix C. These models differed from the base fixed effects GEE model presented in the body of the report in their approach to dealing with the potential bias arising from the correlation of panel data within States over time.

Prais-Winsten FGLS Model

Allowing for first-order autocorrelation and heteroskedasticity in the fixed effects model described in (B-1), the residuals are given by:

$$e_{it} = \rho e_{it-1} + \varepsilon_{it} \qquad \text{(D-1)}$$

$$\varepsilon_{it} \sim N(0, \sigma_i^2) \qquad \text{(D-2)}$$

where ρ is the autocorrelation coefficient.

For periods $t > 1$, the Prais-Winsten model is expressed as:

$$\begin{aligned} ERROR_{it}^* &= \alpha_i^* + t_i^{*\prime}\delta_i + EFFORT^*_{it}\beta_1 + EFFORT\_SQ'^*_{it}\beta_2 + \\ &TANFIMP\_TANF\_EFFORT^*_{it}\beta_3 + NoTI\_TANF\_EFFORT^*_{it}\beta_4 + \\ &CM13\_EFFORT^*_{it}\beta_5 + X_{it}^{*\prime}\gamma + e_{it}^* \end{aligned} \qquad \text{(D-3)}$$

where the asterisks on each of the independent variables and the dependent variable denote the transformation given by:

$$V^*_t = V_t - \rho V_{t-1} \qquad \text{(D-4)}$$

and the error structure is given by:

$$e^*_{it} = \varepsilon_{it} \qquad \text{(D-5)}$$

Note that equation (D-4) requires subtracting a weighted lagged-value of an observation from that same variable's current period value. This cannot be done for the first observation, which might be discarded from the estimation, as in the Cochrane-Orcutt method. Prais-Winsten provide an alternative transformation for the first time period information. We employ the Prais-Winsten method of weighting the first-year's observations by $(1-\rho^2)^{1/2}$. Thus, we rewrite equation (D-4) for t = 1 as:

$$V^*_1 = (1-\rho^2)^{1/2} V_1 \qquad \text{(D-6)}$$

and the error term in period 1 is given by:

$$e^*_{it} = (1-\rho^2)^{1/2} \varepsilon_{it} \qquad \text{(D-7)}$$

After transforming the data, the regression for estimation can be written for all t as:

$$ERROR_{it}{}^* = \alpha_i{}^* + t_i{}^*{}'\delta_i + EFFORT_{it}{}^*{}'\beta_1 + PEFFORT_{it}{}^*{}'\beta_2 + X_{it}{}^*{}'\gamma + e_{it}{}^* \quad \text{(D-8)}$$

The model described by equation (D-8) is a special case of a more general model where the autocorrelation is expressed as:

$$e_{it} = \rho_i e_{it-1} + \varepsilon_{it} \quad \text{(D-9)}$$

where ρ is subscripted by state i. That is, state-specific autocorrelation parameters are estimated, allowing there to be differences in autocorrelation across states. Equation (D-8) is then estimated as above, using a state-specific autocorrelation parameter.

Relative to the Prais-Winsten model using a common ρ the model using a state-specific estimate of ρ may produce less biased estimates if the autocorrelation parameters are not equal across states. It may be less efficient, however, because it requires additional parameter estimates. Thus, the estimates using a common ρ will be consistent and efficient if the autocorrelation coefficient does not vary across states, while the estimates using state-specific values of ρ will be consistent when the autocorrelation coefficient does vary across states, but will be inefficient if it does not.

Partial Adjustment Model

The partial adjustment model assumes that states adjust their resources so as to achieve a desired level of errors, but only make these adjustments gradually. That is, we assume:

$$ERROR_{it} - ERROR_{it-1} = (1-\psi)\left(ERROR^*_{it} - ERROR_{it-1}\right) \quad \text{(D-10)}$$

where ψ is the fraction of the gap that is closed within a year and $ERROR_{it}$* is the desired error rate of state i at time t. Then, rewriting equation (5-2) in the text as the target level of ERROR, we have:

$$\begin{aligned} ERROR^*_{it} &= \alpha_i + t_i'\delta_i + EFFORT'_{it}\beta_1 + EFFORT\_SQ'_{it}\beta_2 + \\ &TANFIMP\_TANF\_EFFORT'_{it}\beta_3 + NoTI\_TANF\_EFFORT'_{it}\beta_4 + \\ &CM13\_EFFORT'_{it}\beta_5 + X_{it}'\gamma + e_{it} \end{aligned} \quad \text{(D-11)}$$

Because we cannot observe the targeted level of error, however, we substitute equation (D-10) into (D-11) and solve for the observed error rate:

$$\begin{aligned} ERROR_{it} &= \tilde{\alpha}_i + t_i'\tilde{\delta}_i + \psi e_{it-1} + EFFORT'_{it}\tilde{\beta}_1 + EFFORT\_SQ'_{it}\tilde{\beta}_2 + \\ &TANFIMP\_TANF\_EFFORT'_{it}\tilde{\beta}_3 + NoTI\_TANF\_EFFORT'_{it}\tilde{\beta}_4 + \\ &CM13\_EFFORT'_{it}\tilde{\beta}_5 + X_{it}'\tilde{\gamma} + \nu_{it} \end{aligned} \quad \text{(D-12)}$$

where the above coefficients with the tildes (such as $\tilde{\beta}_1$) relate to the original coefficients in equation (5-2) in the text by a factor of (1/1-ψ), with $|\psi|<1$. The long-run effect of EFFORT is then given by the following relationship:

$$\beta_1 = \frac{\tilde{\beta}_1}{1-\psi} \qquad \text{(D-13)}$$

The variances for the long-run estimates are calculated via the delta method. Using a linear expansion, $Var(\beta_1)$ is given as dVd' where:

$$d \approx \left[\frac{1}{1-\psi} \quad \frac{\tilde{\beta}_1}{(1-\psi)^2} \right] \qquad \text{(D-14)}$$

is a row vector, approximated with estimates of ψ and $\tilde{\beta}_1$, and V is a 2×2 matrix whose elements are the estimated sampling variances and covariances for $\tilde{\beta}_1$ and $\hat{\psi}$. The calculation of parameter estimates and sampling variances of long-run effects of other covariates are analogous to that described in equation (D-13) and equation (D-14) above.

Arellano-Bond Model

The Arellano-Bond model is based on a method of instrumental variables to surmount the problem of bias and inconsistency introduced when using the lagged dependent variable as a regressor. The model is based on equation (D-12). The disturbances, ν_{it}, are assumed to have finite moments with $E(\nu_{it}) = E(,\nu_{is},\nu_{it}) = 0$ for s≠t. This assumption assumes that there is no serial correlation, but does not require independence over time.

Under these assumptions, values of the dependent variable, ERROR, lagged two periods can be used as valid instruments. For simplicity, we re-write equation (D-12) as:

$$ERROR_{it} = \tilde{\alpha}_i + \psi ERROR_{it-1} + W'_{it}\pi + \nu_{it} \qquad \text{(D-15)}$$

The equation in (D-15) is then first-differenced, thus removing $\tilde{\alpha}_i$ and producing an equation that is estimable via instrumental variables, using two-period lagged values of $ERROR_{it}$. Arellano and Bond (1991) note that for panels with at least three time periods, the model implies m = (T-2)(T-1)/2 linear moment restrictions:

$$E\left[\left(\overline{ERROR}_{it} - \psi \overline{ERROR}_{it-1} - \overline{W}'_{it}\pi\right)\nu_{it-j}\right] = 0 \qquad j = 2,\ldots,(t-1); \quad t = 3,\ldots.,T \qquad \text{(D-16)}$$

where $\overline{ERROR}_{it} = ERROR_{it} - ERROR_{it-1}$. The estimates of the coefficients in (D-15) are obtained via generalized methods of moments (GMM). For further simplicity, including the lagged values of $ERROR_{it}$ as instruments, we rewrite equation (D-15) as:

$$ERROR_{it} = K'_{it}\kappa + \nu_{it} \qquad \text{(D-17)}$$

Then, following Arellano and Bond (1991), the GMM estimator $\hat{\kappa}$ is given by the following kx1 coefficient vector:

$$\hat{\kappa} = \left(\vec{K}'ZA_N Z'\vec{K}\right)^{-1}\vec{K}'ZA_N Z'\vec{e} \tag{D-18}$$

where $\vec{K}$ is a stacked (T-2)Nxk matrix of observations on $\overline{ERROR}$, Z_i = diag ($ERROR_{i1},\ldots,ERROR_{is}, K_{i1},\ldots,K_{is}$) for s=1,…,T-2, and A_N is given by V^{-1}_N, where:

$$\hat{V}_N = N^{-1}\sum_i Z_i'\hat{v}_i\hat{v}_i'Z_i \tag{D-19}$$

The long-run estimate of the effect of effort on error is computed analogously to equation (D-14) above, where ψ is now the parameter estimate associated with the instrument. The standard error associated with the long-run estimate is calculated via the delta method, analogously to equation (D-14).