2012

Pulling Together: Linking Unemployment Insurance and Supplemental Nutrition Assistance Program Administrative Data to Study Effects of the Great Recession

Theresa Anderson
U.S. Department of Agriculture, TAnderson@ers.usda.gov

John A. Kirlin
U.S. Department of Agriculture, JKirlin@ers.usda.gov

Michael Wiseman
The George Washington University, wisemanm@gwu.edu

Pulling Together: Linking Unemployment Insurance and Supplemental Nutrition Assistance Program Administrative Data to Study Effects of the Great Recession*

Theresa Anderson
Economic Research Service, U.S. Department of Agriculture
TAnderson@ers.usda.gov

John A. Kirlin
Economic Research Service, U.S. Department of Agriculture
JKirlin@ers.usda.gov

Michael Wiseman
The George Washington University
WisemanM@gwu.edu

Abstract:

The SNAP-UI Data Linkage Project is an effort coordinated by the U.S. Department of Agriculture's Economic Research Service (ERS) to link state-level administrative data from the Supplemental Nutrition Assistance Program (SNAP) and Unemployment Insurance (UI) program to examine the concurrent and sequential patterns in use of these program before and during the Great Recession. The project focuses on calendar years 2006 through 2009 and utilizes data from seven states: California, Florida, Georgia, Illinois, Maryland, Michigan, and Texas. The project has illuminated various issues with administrative data linkage, which this paper characterizes as the "Three C's" of administrative data: custody, confidentiality, and consistency.

From the outset, ERS had three primary hypotheses: 1) The low rate of concurrent SNAP-UI receipt in existing data understates the total connection between SNAP and UI benefits because people tend to take up nutrition benefits only after UI claims are exhausted. 2) Both the concurrent and sequential links between SNAP and UI grew during the recession. 3) As the economy worsened, the lag between UI exhaustion and SNAP take-up declined. After the discussion of data issues, preliminary project results are presented (current as of December 2011). These early results confirm the first hypothesis but show that the sequential connection between the programs is not as large as expected. The second hypothesis is confirmed. The third hypothesis is still being explored.

* Prepared for the 2012 Federal Committee on Statistical Methodology Research Conference, January 10-12, 2012, Washington, DC. **Please be cautious when citing data from this paper.** The results presented are current as of December 2011, but will change. The analysis and opinions expressed are the authors' and do not represent in any way the Economic Research Service, the U.S. Department of Agriculture, or the George Washington University.

2 April 2012

Public assistance programs typically collect a large amount of data on applicants and beneficiaries to use for both program administration and management. Generally such data are narrowly focused, recording information specific to a single program's operation. However, many policy issues involve the interaction of programs and are best informed by information that reaches across program boundaries. Policymakers increasingly realize that linking administrative data can produce such information, but this is complicated by the major challenges inherent to data linkage. These challenges will be referred to here as the "Three C's": custody, confidentiality, and consistency. Though the body of research on the extent and consequences of the overlap of public assistance programs based on administrative data is expanding, these Three C's continue to challenge researchers in their mission to generate valid, policy-relevant results. This paper will explore how the Three C's have been addressed in the *SNAP-UI Data Linkage Project*, a large multi-state analysis of the intersection of the Supplemental Nutrition Assistance Program (SNAP, formerly Food Stamps) and the Unemployment Insurance (UI) program, being coordinated and funded by the Economic Research Service (ERS) at the U.S. Department of Agriculture (USDA).

Traditionally, the overlap between SNAP and Temporary Assistance for Needy Families (TANF) has been an important feature of the nation's safety net and a popular topic for social policy research (cf. Zedlewski 2001, Ziliak et al. 2003, Ratcliffe et al. 2008). However, TANF has played a very limited role in shielding families from the impacts of the recent recession and slow economic recovery (Zedlewski et al. 2011). In the absence of TANF, SNAP has become increasingly vital as a source of assistance for those in need (Pavetti and Rosenbaum 2010). At the same time, UI has proven particularly important in supporting unemployed workers both in terms of population coverage and duration of receipt (Elmendorf 2010). Despite the substantial safety net roles played by the SNAP and UI programs individually, at the beginning of the Great Recession, the overlap between the two appeared slight, at least in available cross-sectional data.[1] Researchers at ERS suspected that this seemingly weak link might understate the total relationship between the programs. When one takes a longitudinal view of the data, it is possible to identify program sequencing—that is, taking up one program after benefits from another have been exhausted. The occurrence of sequencing and the general programmatic interconnection between SNAP and UI might have grown in recent years as a result of changes in SNAP eligibility, the exceptional severity of the downturn, and the very slow economic recovery. This project is designed to improve understanding of the SNAP-UI connection by use of merged administrative records from the two programs. These data allow researchers to take both a longitudinal and cross-sectional view of the SNAP-UI relationship over calendar years 2006 through 2009. Moreover, ERS can look at the UI connection for the SNAP caseload as a whole, as well as for new entrants into SNAP. The new entrants in late 2007 through the end of 2009 might have different characteristics than the general SNAP caseload because they joined the rolls during the Great Recession and slow recovery. For reasons that will be described in the body of this paper, the research is being completed at the state level by seven research teams.

This analysis is important for several reasons. First, the SNAP and UI systems have not historically been very integrated, but greater integration may be useful—and perhaps necessary—in times of economic stress. Because one of the primary goals of SNAP is to provide benefits to all eligible households, the Food and Nutrition Service (FNS)—the USDA agency that oversees the administration of SNAP—has increased outreach efforts to encourage everyone who is eligible to participate in the program. Many UI recipients are likely eligible for SNAP,[2] and sharing data between SNAP and UI systems provides an opportunity for targeted SNAP outreach. Second, when the Congressional Budget Office (CBO) considers the cost of extensions to the UI program, it does not currently consider the cost of entry onto SNAP by people who lose their UI benefits. By understanding the prevalence of movement onto SNAP by former recipients of UI, CBO and other agencies can better capture these costs in budgetary analyses. Third, this project contributes to the body of knowledge on how to complete these types of studies using administrative data.

---

[1] Based on Food and Nutrition Service Quality Control data, which can be accessed at http://hostm142.mathematica-mpr.com/fns/download.htm.

[2] Future research plans to look at the characteristics of unemployed persons who are technically eligible to receive SNAP and UI concurrently. Preliminary analysis shows that a UI recipient receiving the average UI benefit amount in the fourth quarter of 2009 without other household income and without dependents could concurrently receive SNAP in eight (of 51) states. With one dependent, a UI recipient receiving the average benefit without other household income could qualify for SNAP in an additional 30 states. In only two states (Massachusetts and Washington) would such a person need to have more than two dependents in the household to qualify for concurrent SNAP-UI receipt.

ERS has a useful perspective on the process of completing large-scale analyses using administrative data because the agency has been working since early 2010 to coordinate seven independent research teams in seven diverse states. This provides a broad perspective on the challenges unique to acting as the coordinator of such an endeavor, as well as insight into how various states' administrative data systems compare. Unlike the long-established practice of linking administrative data to survey data (cf. Lansing et al. 1961, Sheikh and Mattingly 1981, Assael and Keon 1982, Bound and Krueger 1991, Raijneveld and Stronks 1999, Raijneveld and Stronks 2001, Gundgaard et al. 2007), in this project the state research teams are linking multiple administrative data systems to each other. The primary purpose is to understand the interaction between SNAP and UI, but some states have also incorporated TANF data into the analysis. Though certainly not the only endeavor of its kind (cf. MDRC 2011), there are important lessons to be learned from the unique challenges inherent to the SNAP and UI data systems, particularly the latter. The expansions in the UI program have complicated data management for state agencies and subsequently made it more challenging to analyze the SNAP-UI connection. At the same time, relaxed eligibility rules along with a sluggish economy have caused the SNAP caseload to grow at an unprecedented rate (Eslami et al. 2011). As the preliminary results at the end of this paper will demonstrate, the effort to link these data systems was not in vain; important patterns in the programmatic connection have become apparent, and these can be used to help direct future policy.

This paper is organized as follows. The first section contains an explanation of the nature of administration data. The second section contains a description of the Three C's and the reasons why they pose challenges to researchers on public assistance programs, as well as a discussion of how ERS has sought to overcome these barriers in the *SNAP-UI Data Linkage Project*. The third section outlines the project's hypotheses and presents some initial results. The final section concludes by looking at the policy implications of the process and results and the lessons learned for future research of this type.

Issues with Administrative Data

Administrative records contain a wide range of information that is relevant to program administration and policy. Both means-tested and insurance-based assistance programs necessarily collect and maintain information to determine program eligibility, track participation in complementary programs, administer payments, and provide data for internal auditing. Eligibility data might include individual and family income, assets information, household composition, and—depending on the basis for benefits—health or employment history. Information on participation in complementary programs is important in instances when eligibility for one program is automatic if benefits are received from similar programs; this is known as categorical eligibility. For example, Medicaid files keep track of participation in TANF and the Supplemental Security Income (SSI) program because TANF or SSI participation automatically qualifies a person for Medicaid. Finally, the information necessary for administering payments might include the benefit amount, contact information for beneficiaries, and the part of the program for which the benefits are being administered (e.g., regular or extended benefits for UI). Basic demographic data, such as race/ethnicity, sex, marital status, and birth date also are collected routinely.

Information from administrative records is available for a census of program participants, meaning that using administrative data generally dispenses with research problems of coverage and sampling error, both major concerns in surveys. Moreover, administrative data tend to be timelier than survey data, providing the potential for researchers to complete analyses almost contemporaneously with program administration. Finally, administrative records tend to be less expensive to collect for research than surveys, both to the researchers and the public, since they need to be collected for program administration in any case. The use of administrative data for longitudinal analyses of program participation is relatively new, but administrative sources are generally preferable to surveys for this type of research because they are not subject to recall error and the timing of the sequencing of program usage can be measured much more finely. For example, in the Current Population Survey (CPS), one can only ascertain if both SNAP and UI were received at some point in the same year. Even in the Survey of Income and Program Participation (SIPP), which has much more detailed measures of program participation than the CPS, one can only know if both benefits were received in the same month, but the order of receipt is not necessarily obvious. From administrative data, even one day of co-receipt can be detected, and both the history and future of participation in each program can be known for the period over which records are available.

Administrative data, however, are not without their drawbacks. When merging administrative records with survey data or with other administrative data, errors can emerge from various sources. These error sources include the matching algorithm, differences in the construct of variables between administrative data and the statistical uses to

which they are meant to be put, and the variation in reference time frames among programs (e.g., UI payments occur weekly, but earnings records are maintained quarterly, and many programs issue benefits on a monthly cycle). Moreover, though the collection and analysis of administrative data *can* be nearly immediate, the process of clearing legal hurdles can be more time-consuming than similar processes for survey data, sometimes taking months or years (Gates 2011). Other errors in administrative data can originate from duplicate entries that are difficult to identify, the inclusion of deceased individuals in administrative records (the deceased are rarely accidently captured in surveys), transcription errors, and imputation of values in administrative data.[3] These last two issues are also common in surveys. While information regarding agency action in the administrative record is probably reliable, such as when a check was mailed and for what amount, there are potentially reporting errors in beneficiary and household characteristics, perhaps from misreporting by the beneficiaries themselves.

Historically, administrative data have been used as a "gold standard" against which to evaluate the quality of survey responses (Pedace and Bates 2000, White 2010). In recent analyses, researchers have tried to measure reporting error in administrative data, moving beyond the "gold standard" treatment of these records. Kapteyn and Yeb Ypma (2006) find that administrative data from Sweden produce more bias than survey data on some measures. Abowd and Stinson (2011) compare individuals' earnings records from the Social Security Administration (SSA) with the SIPP and find that earnings from administrative records are somewhat more reliable than SIPP-reported earnings, although the researchers' results are subject to matching error between the data sources. Groen (2011) estimates that reporting differences account for 48 percent of the difference between counts of employees from the Quarterly Census of Employment and from the Wages and the Current Employment Statistics Survey, though from this analysis there is no way determine which source is more accurate.

The Three C's of Administrative Data

Overall, the analyses available to date suggest that administrative records are at least as reliable as survey data and probably more so in various ways. As mentioned, these data also are most appropriate for analyses that require detailed information on the exact timing of program participation. For these reasons, ERS decided to focus on administrative records as the source of information about the SNAP-UI relationship during the recession. This endeavor, however, was not without challenges. This paper proposes a new framework around which challenges in administrative data usage might be organized, called the Three C's, custody, confidentiality, and consistency. These represent three of the major hurdles that researchers must overcome to make use of administrative records.

Custody

Custody is maintained by the entity that collects data and has functional control of their distribution and use. The administrative data for many federally-managed programs are not maintained at the federal level, but rather by the state agencies that interact with program participants. Some federal programs, such as those that fall under the purview of the Social Security Administration, have mechanisms in place whereby all recipient data are transferred to a federal data system on a regular basis. In contrast, the offices that process UI do not send all recipient data to the Department of Labor (DOL). Rather, only aggregate data on total UI beneficiaries, total benefits administered, and total/average wages are shared with the federal government. In SNAP, only a sample of data is shared with FNS. FNS collects aggregate caseload statistics from the states as well as their Quality Control (QC) data, which are compiled from monthly samples that state agencies draw from their administrative SNAP records. FNS selects a QC sample from this database annually to review for SNAP payment errors.[4] Moreover, a report of characteristics of SNAP households based on QC data is released annually, and a public-use version of the QC dataset is made available. However, the QC data are subject to sampling error and suffer from some other biases due to the nature of prospective program budgeting. Given the limited data access to SNAP and UI data on the federal level, in order for ERS to investigate programmatic relationships at the micro level, it was necessary to go directly to state agencies.

[3] Groen (2011) discusses the issue of imputation in administrative data, as well as other threats to administrative data quality.

[4] This is the information that informs reports of the states' SNAP QC error rates, which are used to determine performance bonuses for states with low errors rates or most-improved error rates. Error rate data can be found at http://www.fns.usda.gov/snap/qc/default.htm.

ERS cannot easily negotiate with individual state administrative agencies. Therefore, ERS released a competitive request for proposals for the SNAP-UI project at the beginning of calendar year 2010, seeking "State and academic research organizations with access to SNAP/FSP, UI, and possibly other program records…" (ERS 2010). Ultimately, ERS awarded research grants to three organizations representing seven research teams who had experience negotiating data usage with state SNAP and/or UI agencies. The seven research teams and the states they are investigating are:

- California Institute of Public Policy — California
- University of Missouri — Florida
- Jacob France Institute — Maryland
- Andrew Young School — Georgia
- Chapin Hall Center — Illinois
- Upjohn Institute — Michigan
- Ray Marshall Center — Texas

The last five research organizations above are participating as a consortium of project partners under the aegis of the Jacob France Institute. Most of the researchers involved in this project are familiar with the SNAP or UI agency in their respective states. In various cases, however, the project partners needed to form new relationships with the agency with custody of the other dataset because SNAP and UI program data are rarely maintained by a single agency. In some cases, the SNAP-administering agency communicated with the UI-administering agency on behalf of the research team. In other cases, the research team had to communicate with each agency separately. This not only posed time costs, but was also substantially dependent on the social capital of the research teams. At times, figuring out whom to contact regarding data access, particularly on the UI side, was not straightforward and required substantial networking. Were custody of the data centralized or were the mechanisms by which data could be collected, compiled, approved for research, and disseminated better integrated, the data gathering phase of the project may have been much easier. However, during the course of this project, which officially kicked off in May 2010, not only did state governments face substantial budget challenges, but in four of the seven states the leader of the executive branch changed in fall 2010. The turnover in state agencies as a result of local political dynamics further complicated the process of tracking down the person or office with custody of the necessary data.

Confidentiality

Each of the custodial agencies of administrative data is charged with maintaining confidentiality of the identities of program participants. This is, reasonably, of the highest concern for those who collect and maintain administrative data. When a person answers a survey, it is generally voluntary, there is no substantial monetary payoff, and there is an understanding that the information supplied might be used for research as per the guidelines of the privacy protection notice that respondents receive. On the other hand, administrative data are gathered from participants in social programs in exchange for financial or other needed support, and often participants in these programs do not believe they have a choice but to provide their personal information (Brady et al. 2001). Program participants do not necessarily offer their information with the expectation that it will be used for research purposes, and they are rarely informed of the potential uses of their data (Ibid.). Data linkage is even more sensitive because when information is aggregated, new information about the subject of the linkage (often the individual) is created, and re-identification risk heightens, even for datasets that have been individually scrubbed of personally identifiable information (GAO 2001).

These facts do not preclude the use of the information for research, but they do suggest that more care must be taken in the handling and use of administrative data. Generally, U.S. law permits the use of administrative data for statistical research (Gates 2011). The U.S. Census Bureau in its capacity as a federal statistical agency has a broad legal mandate to use administrative data in lieu of censuses and surveys when possible. The law states:

> (a) The Secretary, whenever he considers it advisable, may call upon any other department, agency, or establishment of the Federal Government, or of the government of the District of Columbia, for information pertinent to the work provided for in this title. (b) The Secretary may acquire, by purchase or otherwise, from states, counties, cities, or other units of government, or their instrumentalities, or from private persons and agencies, such copies of records, reports, and other material as may be required for the efficient and economical conduct of the censuses and surveys provided for in this title. (c) To the

> maximum extent possible and consistent with the kind, timeliness, quality and scope of the statistics required, the Secretary shall acquire and use information available from any source referred to in subsection (a) or (b) of this section instead of conducting direct inquiries. (13 U.S.C. §6)

Most federally-administered programs have their own rules about how the program data can be used. The *Food, Conservation, and Energy Act of 2008* specifies that data on SNAP participants can be disclosed only "to persons directly connected with the administration or enforcement of the provisions of this Act, regulations issued pursuant to this Act, Federal assistance programs, or federally-assisted State programs." Moreover, "the subsequent use of the information by persons described in [the previous] clause [should] only [be] for such administration or enforcement" (Pub. L. No. 110-246, 122 Stat. 1874).

For UI, federal law specifies the confidentiality requirements for wage and benefits data. While strict, the law does specify that:

> (e) Disclosure of confidential [Unemployment Compensation] information to a public official for use in the performance of his or her official duties is permissible. "Performance of official duties" means administration or enforcement of law or the execution of the official responsibilities of a Federal, State, or local elected official. Administration of law includes research related to the law administered by the public official. Execution of official responsibilities does not include solicitation of contributions or expenditures to or on behalf of a candidate for public or political office or a political party. (f) Agent or contractor of public official. Disclosure of confidential [Unemployment Compensation] information to an agent or contractor of a public official to whom disclosure is permissible under paragraph (e) of this section. (20 C.F.R. §603.5)

Moreover, various safeguards are specified in the law that must be taken to ensure that confidentiality is not breached once the information is transferred under this provision.

The most difficult issue in gaining access to administrative data is not the confidentiality rules themselves, but rather the fact that many state officials are not certain about what information is permissible to share with whom for what purpose, given the complexity and both federal- and state-level regulations, as well as program-specific rules. Brady et al. (2001) and Gates (2011) explore this issue in depth. Four of the seven states in the project are members of the Administrative Data Research and Evaluation (ADARE) alliance,[5] which is a partnership of state research institutions that have continuous data sharing agreements with state social service agencies; the researchers use the data for studies of program performance and accountability (UBalt 2011). The Jacob France Institute (JFI) at the University of Baltimore is the coordinator of the ADARE project, and JFI is coordinating a consortium of five state research teams for ERS's SNAP-UI project. These five state research teams, as well as the other two project partners, have substantial experience negotiating data sharing agreements with their state agencies.

Even with this wealth of experience, the process of obtaining signed memoranda of understanding (MOUs) for this project took over a year in some states. This was the most time-consuming aspect of the project, and it led to a need for no-cost extensions for all of the research teams beyond the original one-year grant periods.

Moreover, concerns over confidentiality have meant that five of the seven research teams were not allowed to merge records from the SNAP and UI systems themselves. Rather, employees of the state government completed the match. Though ERS is making an effort to document the exact methodology of the matching process, this has added an extra "black box" to the data quality assessment process. In at least one state, errors in the merge process have been obvious, and have led the research team to devote time and resources to "sleuthing out" the source of the problem.

The final issue occurs because ERS will not gain direct access to the merged data files. Instead, ERS distributed common, detailed analysis shells to the research teams, who completed them and transmitted them back to ERS. The confidentiality protections in some of the MOUs, however, limit the cell sizes that the state research teams can report to ERS. The extent to which this restriction will impact the results remains to be seen.

[5] The ADARE states in this project are Florida, Georgia, Maryland, and Michigan.

Consistency

Consistency refers both to internal consistency, which is a component of data quality generally, and to the consistency of data among different sources. The latter relates to the ability of the data sources to be combined and interpreted in useful ways.

Internal consistency refers to the internal logic of each dataset. Internal consistency might be threatened by data that are inconsistent across time, across similar or related variables, or across observations. An example of data that are inconsistent across time is when variables are repurposed due to a programmatic change. An example of internal inconsistency across similar or related variables is when apparently incorrect benefit amounts are reported, given the wage and other information available in the record. One form of internal inconsistency across observations is missing data. All of these types of inconsistencies are common in administrative data. The problem of inconsistency across time is especially apparent in the UI datasets, where variables were repurposed in several states after UI extensions were implemented in 2008. Without a good, up-to-date data dictionary (which is often hard to come by), interpreting the meaning and use of variables can be complicated at any point, and the difficulty is compounded when the variable definitions or coding structures change without clear indication.

The issue of consistency among data sources relates to how the definitions of variables and the structure of the databases themselves affect the ability of researchers to merge them in useful ways. This is related to the earlier discussion of the general drawbacks of combining multiple administrative data sources or combining administrative data with other types of information, which warns of differences in the construct of variables among sources. An example of this is when earnings are measured on a monthly basis in one database and quarterly in another, and the accounting does not match up between the two (meaning that the sum of three months of earnings within the quarter do not equal the total reported quarterly earnings). This is exactly the type of issue that the researchers have encountered in the *SNAP-UI Data Linkage Project*.

ERS grappled with how to rectify inconsistencies across datasets in the common analysis shells that were distributed to the research teams. One of the first issues confronted was the fact that SNAP is a monthly benefit, UI is a weekly benefit, and UI wage records are maintained on a quarterly basis. Because a week of UI benefits can occur over two calendar months, a common definition of the date of UI receipt was implemented for the purposes of our project: "A payment received in any week of the month is considered to be received for that entire month. For benefit weeks that overlap two months, the payment should be considered to have been received on the Saturday at the end of that week and thus the payment allocated to the month in which that Saturday falls." When quarterly earnings had to be interpreted on a monthly basis, earnings for the entire quarter was delegated to each month in the quarter.

Other struggles originated from the variation in definitions of seemingly basic concepts such as "employment," "earnings," and "dependents" across the two programs. In the analysis shells, constructs such as employment were measured in different ways for the SNAP and UI data to facilitate comparison. For example, from the SNAP files, the research teams were asked to tabulate the number of SNAP recipients who were "[c]urrently residing in a SNAP unit with earned income (from SNAP record)" with the note, "This is based on the presumption that individual earnings and employment status are not reflected in the SNAP administrative data. If individual-level data are available, please also tabulate 'Currently employed (from SNAP record)' and 'Currently not employed (from SNAP record).'" From the UI records, the research teams were asked to tabulate the number of SNAP recipients with "[e]arnings > $100 in current quarter (from UI record)." The $100 threshold for quarterly earnings was chosen to avoid cases in which $1 – $99 of earnings were recorded for a quarter, since this does not likely represent meaningful employment and might be a data error.

Summing Up Data Barriers

In most cases, gaining access to and cleaning the data has taken substantially more time than the analysis itself. The difficulties have varied among states based on state buy-in to the project, political dynamics, the organization of data systems, the interpretation of confidentiality laws, and the interpersonal connections of the researchers. As of fall 2011, a year and a half after the project kick-off, data have been obtained and initial results have been produced from six of the seven states. Serious issues in data quality pose threats to the internal validity of the results of one of the states. The resolution of this problem has been held up by difficulties ascertaining who has custody of the UI data, though progress is being made. The next section will present some preliminary results from five of the seven state research teams.

Project Hypotheses and Preliminary Results

The SNAP-UI project began with three broad hypotheses. The first is that the low rate of concurrent SNAP-UI receipt understates the total connection between SNAP and UI benefits because people tend to take up nutrition benefits only after UI claims are exhausted. The researchers expected that the longitudinal analysis would reveal that program sequencing would account for more than half of the total SNAP-UI connection—that is, the concurrent connection plus the sequential connection. The second is that both the concurrent and sequential links between SNAP and UI grow during recessions. The researchers expected that the growth in the SNAP-UI connection would be especially apparent when looking at *new* SNAP entrants in the late 2007-2009 timeframe (i.e., the recession and recovery). The third is that, as the economy worsened, the lag between UI exhaustion and SNAP take-up declined.

The role of ERS in this project has been to frame the research questions, define the format of analysis, and coordinate the teams so that consistent, comparable final results emerge. The project is organized into three levels of analysis. Level 1 covers the foundational question: What is the intersection of SNAP and UI? This level is divided into two parts, one with a focus on the SNAP-UI connection and one with a SNAP-only focus. ERS developed analysis shells for both parts of Level 1, as referenced earlier. These underwent a collaborative editing process with the research teams before they were finalized. The SNAP-UI analysis shells focus on two groups of SNAP recipients: new entrants and all current recipients. Level 1 is further refined with the Level 1.5 tables, which analyze the basic SNAP-UI connection at the individual, household, and individual-within-household levels of analysis. Level 1.5 also changes the way the UI connection is measured, separating UI receipt from the UI benefit year. Level 2 uses the characteristics categories from the FNS Household Characteristics Reports to break down the groups introduced in Level 1 by demographic characteristics. The tables are segregated by individual-level and household-level analyses; cross-sectional and cohort analyses; and SNAP-UI and SNAP-only analyses. Level 3 offers opportunities for researchers to address issues that are particular to their unique data strengths and the needs of the SNAP or UI programs in their states. This structure is presented visually in Figure 1.

Figure 1: Levels of Data Analysis

Early Results

Though each of the research teams is analyzing its own data, the advantage that ERS has in analysis of the results is the ability to make cross-state comparisons. The following data are preliminary and are subject to change. First, a look at the basics. Figure 2 shows the relative changes in the SNAP caseloads in each state, normalized to the beginning of the recession in December 2007. Though the caseload grew in all states, the growth rates varied substantially. In Illinois, the caseload grew 26 percent between December 2007 and December 2009, while in Florida, it grew 80 percent over that same time period. By October 2010, the number of SNAP recipients in Illinois had increased 34 percent relative to December 2007, while in Florida the number increased by 113 percent over that same time period. Nationally, the number of SNAP recipients grew 42 percent between the beginning of the recession and the end of 2009 and 57 percent by October 2010.

Figure 2: SNAP Receipt in the Seven Project States, Jan. 2005-Oct. 2010

Figure 3 is similar to Figure 2, but it represents the changes in the UI caseloads of the states relative to the number of UI recipients at the beginning of the recession in December 2007. This represents receipt of UI from all programs, including extended and emergency benefits, as well as special programs, such as those for railroad workers and federal employees. Throughout this period, the UI caseload fluctuated substantially. In Michigan, recipiency increased the least, growing 172 percent between December 2007 and December 2009, partially because of the state's already high levels of UI receipt. In contrast, in Florida and Georgia the number of UI recipients grew 399 percent and 463 percent, respectively, between December 2007 and December 2009. Nationally, the number of UI recipients grew 226 percent over this same time period.

Fig UICOMPARE Here

Figure 3: UI Receipt in the Seven Project States, Jan. 2005-Oct. 2010

Besides the variation among the states in the SNAP and UI caseloads, these states represent substantial diversity in geography and size. Table 1 gives some demographic data on the project states. In each demographic category, with the exception of the percentage of the population identifying as American Indian, some of the project states fall above and some below the national statistic.[6] All of these states are larger than the national median. Finally, these states are located in three of the four major Census regions of the United States: West, Midwest, and South.[7]

Figure STATEDEMOS Here

Table 1: Demographics of SNAP-UI Project States

Figure 4 shows the number of persons with concurrent receipt of SNAP and UI benefits as a percentage of the working-age SNAP caseload in five project states based on data gathered in this project. The national rate is from the QC data.[8] Here, "working age" defines persons age 18-64, inclusive, since this is the age range in which unemployed or partially unemployed workers are potentially eligible for UI.

Figure CONCURRENTSNAP Here

Figure 4: Concurrent SNAP-UI Receipt as a Percentage of the Working Age SNAP Caseload

Figure 5 shows the number of persons with concurrent receipt of SNAP and UI benefits as a percentage of the state's UI caseload in the same five states based on the preliminary project results and the national QC estimate. Note that the scale in this figure is twice the scale in Figure 4. Here, the UI caseload is defined as the number of persons in each state who receive UI payment at any time during the month. This is computed by taking the number of UI continuing claims at the beginning of the month in each state and adding the number of new claims in that month.[9]

Figure CONCURRENTUI Here

Figure 5: Concurrent SNAP-UI Receipt as a Percentage of the UI Caseload

A comparison of these figures is revealing. The numerator in both is the same. The denominator in Figure 4 is each state's working-age SNAP caseload, while the denominator in Figure 5 is the monthly state UI caseload. From these figures, it is clear that the proportion of each state's UI recipients that concurrently receives SNAP is much larger than the proportion of each state's SNAP recipients that receives UI. This is reasonable, given that the UI caseload in each state is smaller than the SNAP caseload at all points in time. In both cases, the proportion of concurrent receipt grows over time, but the growth is especially apparent in Figure 4. Figure 5 has substantial variation, mostly because of the variation in the denominator – the UI caseload size – as is evidenced in Figure 3, above. The dramatic drop in program overlap in Texas between May and June 2009 does not have an obvious policy explanation. The Texas research team is exploring the source of this anomaly. In Michigan, the change in the trend in Figure 4 in May 2009 from positive to negative appears to be a combination of declining UI applications and declining rates of monetary UI eligibility. That is, a decreasing number of people had earned enough prior to job-loss to be able to quality for UI benefits. The reasons for these phenomena are not entirely clear.

[6] In this figure, the five-year ACS averages were used to even out some of the turbulence that has occurred in these states as a result of the recession, resulting in a more broadly representative comparison.
[7] Based on geographic regions of the U.S. defined in Census (N.D.)
[8] Data for the other two project states are still being processed and are not ready for presentation as of this writing.
[9] Scott Gibbons at DOL Employment and Training Administration is gratefully acknowledged for his help in constructing the UI caseload numbers for the states.

Figure 6 displays the percentage of current SNAP adults who are UI age-eligible[10] and either currently receive UI or have received UI payments in the recent past in four states. The five graphs are produced on the same scale to allow for easy comparison.

Figure XSECTCOMPARE Here

Figure 6: Percent of Current SNAP Adults Who Are UI Age-Eligible with Prior or Current Unemployment Insurance Benefits

Figure 7 is arranged similarly to Figure 6, but it represents *new* SNAP recipients each month and their UI connections. The main difference between this figure and the previous one is that Figure 6 showed a time series for the SNAP caseload, which is (mostly) the same group of people month to month. In the cohort analysis in Figure 7, each month represents a new group of SNAP recipients.[11] This difference plus the smaller *N* explain the increased volatility of the time series in Figure 7.

Figure COHORTCOMPARE Here

Figure 7: Percent of New SNAP Adults Who Are UI Age-Eligible with Prior or Current Unemployment Insurance Benefits

Some data quality issues in Figure 6 and Figure 7 should be addressed. First, past UI receipt in Florida is not shown in either figure prior to October 2006 due to left-hand data censoring. Second, there are two months of left-hand data censoring in Michigan in Figure 7. Finally, the drop in concurrent SNAP-UI receipt between May and June 2009 in both figures in Texas is part of the same anomaly seen in Figure 4 and is being investigated.

Otherwise, Figure 6 and Figure 7 reveal several important patterns. First, as might be expected, new working age SNAP recipients have a stronger connection with UI than all working age SNAP recipients. The connection among new SNAP recipients among the five states is about 50 percent higher than the connection among all SNAP recipients. This ratio is actually somewhat lower in 2009 than in other years, perhaps because by that time, a large proportion of the UI recipients who were apt to take up SNAP had already done so. The difference in the SNAP-UI connection between new SNAP recipients and all SNAP recipients is largest in Michigan, at 69 percent across the entire timeframe. The difference is lowest in Maryland at 32 percent. Note that the two figures have a difference in scale of eight percentage points on the vertical axis.

Second, and contrary to ERS's initial expectations, concurrent SNAP-UI receipt is substantially more common than sequenced receipt, though this ratio shifts over time. In Figure 7, it is apparent that by the end of 2008 most new SNAP recipients in every state receive a UI payment in the same month that they begin SNAP. On average, in the five states over this timeframe, 68.4 percent of the SNAP-UI connection among all SNAP recipients is concurrent (while the other 31.6 percent is sequenced) and 79.2 percent of the SNAP-UI connection among new SNAP recipients is concurrent (with the other 20.8 percent sequenced). The prevalence of concurrent SNAP-UI receipt relative to all receipt rises over time, from 34.6 percent in January 2006 to 62.0 in December 2009 for all working age SNAP recipients and from 49.0 percent to 71.9 percent over the same time period for new working-age SNAP recipients.[12]

[10] Note that we are saying only that the SNAP individuals are *age-eligible for UI*. For this section of the analysis we do not have information on individuals' wage or non-wage eligibility for UI.

[11] These graphs do not attempt to identify individuals who may have entered SNAP on more than one occasion during the study period.

[12] β_1 for the cross-sectional analysis is 0.0063 ($p<0.001$) with months since December 2005 as the independent variable. β_1 for the cohort analysis is 0.0050 ($p<0.001$) with months since December 2005 as the independent variable. That is, for each month since December 2005, the rate of concurrent receipt increases by 0.50 percentage points per month on average for the entire caseload and increases by 0.63 percentage points per month on average for new recipients.

Third, different states show distinct patterns. Florida has the strongest SNAP-UI connection of these five states for **all** working-age SNAP recipients, with a connection apparent for 12.9 percent of this group at the maximum in November 2009. This value is 3.0 percentage points higher than the average of the other four states. The trend in the prevalence of SNAP-UI overlap in Florida is significantly different than the average trend in the other states.[13] However, Michigan has the strongest SNAP-UI connection of these five states for **new** working age SNAP recipients, with a connection apparent for 20.4 percent of this group, at the maximum in February 2009. This value is 3.4 percentage points higher than the average overall trend in the other four states. The overall trend in the SNAP-UI connection among new SNAP recipients in Michigan is also significantly different from the trend in the other four states.[14] Texas, on the other hand, has the weakest connection, with 8.5 percent of the total working age SNAP caseload and 11.9 percent of its new working age SNAP caseload with a UI connection. These values are 2.5 percentage points and 3.9 percentage points lower than the average of the other four states, respectively. The overall trend in the prevalence of SNAP-UI overlap in Texas is significantly different than the average trend in the three other states for both the cohort and cross-sectional analyses.[15]

Moreover, SNAP recipients in Texas seem to sequence UI and SNAP receipt more than in the other states. That is, the proportion of new working age SNAP recipients who were receiving UI in their first month of SNAP to all working age SNAP recipients with a UI connection is lower than in other states. At the highest point, 67.8 percent of new SNAP recipients with a UI connection were receiving UI in the first month that they entered SNAP in Texas. The overall trend of concurrent receipt of SNAP and UI in Texas is significantly different from the overall average concurrent trend in the other four states.[16] In contrast, in Florida the highest rate is 82.2 percent, in Georgia it is 76.3 percent, in Michigan it is 82.5 percent, and in Maryland it is 86.6 percent. These differences are apparent in Figure 7.

Comments on the Results

These results reveal surprising information: the longitudinal relationship between SNAP and UI is not as strong as the concurrent relationship. The first hypothesis of this project was that the low rate of concurrent SNAP-UI receipt understates the association between SNAP and UI benefits. The idea behind this was that people tend to take up SNAP only after UI claims are exhausted. The researchers expected that program sequencing would account for more than one half of the total SNAP-UI connection. This is not reflected in the data. While consideration of the longitudinal relationship does reveal a stronger SNAP-UI connection, the sequential take-up of the two programs is not as large as the concurrent take-up. In fact, it is about one-third of the total SNAP-UI connection among all SNAP recipients and about one-fifth among new SNAP recipients. This could be attributable to the fact that relatively few people were exhausting UI during this time period.[17] The U.S. Congress has extended Emergency Unemployment Compensation (EUC08) various times from the creation of the program on June 30, 2008 through the writing of this paper, which has allowed person living states with high rates of unemployment to collect UI benefits continually for up to 99 weeks.[18] This means that a person who took up UI anytime after April 2008 in a high unemployment state could have received UI payments continuously through the end of the project focus period.[19] If the person did not claim benefits every week during their benefit year, they could have started UI even earlier than 2008 and still have been eligible for payments through the end of the project period. Because Tiers III

[13] Using a paired t-test, $p<0.001$.
[14] Using a paired t-test, $p<0.001$.
[15] Using a paired t-test, $p<0.001$ for both datasets.
[16] Using a paired t-test, $p<0.001$.
[17] It might also suggest that new SNAP recipients take up SNAP benefits at the same time that they begin UI or at some point while UI benefits are active, rather than only after exhaustion. Further analysis of the data will allow for that hypothesis to be explored.
[18] 99 weeks of benefit receipt is achievable by combining 26 weeks of regular UI benefits with 20 weeks of Tier I EUC08, 14 weeks of Tier II, 13 weeks of Tier III, 6 weeks of Tier IV, and 20 weeks of Extended Benefits (EB). By December 2009, persons living in five of the seven states qualified for up to 99 weeks of UI. Workers in Maryland could only qualify for up to 73 weeks of UI and in Texas only up to 93 weeks, due to the lower unemployment rates in those states.
[19] Though mathematically this should be anytime after January 2008 (99 weeks before the end of 2009 is the beginning of February 2008), given the timing of the UI legislation, a person who became unemployed before May 2008 would have had at least one gap in UI coverage.

and IV of EUC08 were not established until two months before the end of the project focus period, any UI recipient in any of the project states who had exhausted all other benefits and was still unemployed as of November 2009 would have been able to continuously receive UI well into 2010.[20] If the project focus period were extended beyond the expiration of EUC08, the balance between concurrent and sequential SNAP-UI interaction would likely shift.

The second hypothesis was that concurrent and sequential links between SNAP and UI grow during recessions. It was expected that this growth would be especially apparent when looking at cohorts of new SNAP recipients, as opposed to cross-sections of the entire caseload. This hypothesis is supported by the results thus far, at least for the Great Recession. The rate of concurrent receipt for the entire caseload is higher than that measured by FNS's QC data because of the nature of how information is recorded for the QC.[21] The link between SNAP and UI is about 50 percent higher among new SNAP recipients than among all SNAP recipients. Comparisons of ERS's results from survey data are forthcoming, but will likely show a similar pattern.[22]

The third hypothesis is that, as the economy worsened, the lag between UI exhaustion and SNAP take-up declined. This issue is still being explored. It is clear that the likelihood of taking up SNAP and UI in conjunction has increased. However, the change in the lag is unclear and merits further analysis.

Discussion

This paper has presented a discussion of the difficulties associated with using administrative data for research on social assistance programs and described how ERS's *SNAP-UI Data Linkage Project* has dealt with some of these issues. An exposition of the preliminary results has shown that ERS and the research teams have been successful in producing new, meaningful information regarding the relationship between the SNAP and UI programs. Nonetheless, data quality is an ongoing challenge. A key forthcoming component of this project is an in-depth look at how data quality has affected the project processes and results.

As has been shown, the SNAP and UI programs have grown precipitously over the past few years. When the federal extensions to the UI program expire—possibly in 2012—those who have relied on the program during the recession and "jobless" recovery will need a new source of support. Employment is the preferred source of economic support, but in the absence of good work opportunities with sufficient pay, SNAP likely will become an ever more important source of financial assistance. Currently, SNAP is seeing an increasing number of new entrants who have connections to the UI system. The expectation is that as UI expires, the occurrence of new entrants with recent UI attachment will grow substantially. State SNAP agencies can eventually make use of analyses like this one to predict future program demands on the basis of expected exhaustion patterns of UI benefits. However, the only way to measure this is to continue data collection through expiration of extended and emergency UI and beyond.

An extension of this research would require not only further financial support, but it would require several of the state research teams to confront at least two of the Three C's again, as some of the project MOUs were written specifically for the 2006-2009 time frame. An MOU modification may mean further delay. Moreover, at the beginning of 2013, custody of the data might shift as old administrations leave state offices and new ones enter. These new administrations may interpret confidentiality rules differently, and this could further complicate efforts to extend the analyses. As these possibilities indicate, the fact that data linkage and sharing processes are still not standardized means that further research might face high time and monetary costs. Nonetheless, the potential payoffs of using administrative data are quite high. Here, ERS has been able to compile unique information on the social safety net that could have important implications for the future of the SNAP and UI programs.

[20] All project states qualified for EUC08 through Tier III. All of the project states except Maryland and Texas qualified for Tier IV from its creation. All states except Maryland qualified for EB beginning between January and May 2009, depending on the state.

[21] Though a thorough discussion of this topic is beyond the purview of this paper, suffice to say that QC procedures often do not record information about the sampled family unit that would potentially lower SNAP benefits. This includes receipt of UI benefits that would not render the family unit ineligible for SNAP benefits. Thus, QC data may under-represent actual concurrent receipt of SNAP and UI benefits.

[22] A project analyzing the SNAP-UI interaction based on survey data is being led by David Finifter at the College of William and Mary and Mark Prell at the Economic Research Service. Those findings, when they become available, will be informative for comparison with the results of this project.

References

Abowd, John M. and Matha H. Stinson. (2011). *Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Bureau Survey and SSA Administrative Data*. Working Paper CES 11-20. Washington, DC: Center for Economic Studies.

Assael, Henry and John Keon. (1982). Nonsampling vs. Sampling Errors in Survey Research. *Journal of Marketing, 46,* 114-123.

Bound, John and Alan B. Krueger. (1991). The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics, 9(1),* 1-24.

Brady, Henry E., Susan A. Grand, and M. Anne Powell. (2002). Access and Confidentiality Issues with Administrative Data. In Michele Ver Ploeg, Robert A. Moffitt, and Constance F. Citro. *Studies of Welfare Populations: Data Collection and Research Issues*. Washington, DC: National Research Council.

Eslami, Esa, Kai Filion, and Mark Strayer. (2011). *Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2010*. Report No. SNAP-11-CHAR. Washington, DC: Mathematica.

Gates, Gerald W. (2011). How Uncertainty about Privacy and Confidentiality Is Hampering Efforts to More Effectively Use Administrative Records in Producing U.S. National Statistics. *Journal of Privacy and Confidentiality, 3(2),* 3-40.

Groen, Jeffrey A. (2011). *Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures*. Working Paper. Washington, DC: Bureau of Labor Statistics.

Gundgaard, Jens, Ola Ekholm, Ebba Holme Hansen, Niels Kr. Rasmussen. (2007). The Effect of Non-Response on Estimates of Health Care Utilisation: Linking Health Surveys and Registers. *European Journal of Public Health, 18(2),* 189–194.

Kapteyn, Arie and JelmerYeb Ypma. (2006). *Measurement Error and Misclassification*. Working Paper WR-283-1. Santa Monica, CA: RAND.

Lansing, John B, Gerald P. Ginsburg, and Kaisa Braaten. (1961). *True Fact vs. Misstatement: An Investigation of Response Error*. Urbana, IL: The Bureau of Economic and Business Research.

MDRC. (2011). *TANF/SSI Disability Transition Project*. Retrieved 30 November 2011 from http://www.mdrc.org/project_20_95.html.

Pedace, Roberto and Nancy Bates. (2000). Using Administrative Records to Assess Earnings Reporting Error in the Survey of Income and Program Participation. *Journal of Economic and Social Measurement, 20,* 173-192.

Raijneveld, Sijmen A. and Karien Stronks. (1999). The Impact of Response Bias In Estimates of Health Care Utilization In a Metropolitan Area: The Use of Administrative Data. *International Journal of Epidemiology, 28,* 1134-1140.

Raijneveld, Sijmen A. and Karien Stronks. (2001). The Validity of Self-Reported Use of Health Care across Socioeconomic Strata: A Comparison of Survey and Registration Data. *International Journal of Epidemiology, 30,* 1407-1414.

Sheikh, K. and S. Mattingly. (1981). Investigating non-response bias in mail surveys. *Journal of Epidemiology and Community Health, 35,* 293-296.

U.S. Census Bureau (Census). (N.D.). *Census Regions and Divisions of the United States*. Retrieved 30 November 2011 from http://www.census.gov/geo/www/us_regdiv.pdf.

U.S. Economic Research Service (ERS). (2010). *Request for Applications for Competitive Grant Awards to Conduct Economic Research on the Joint Contributions of the Supplemental Nutrition Assistance Program and Unemployment Insurance to the Nation's Social Safety Net*. Retrieved 30 November 2011 from http://www.ers.usda.gov/Publications/AP/AP043/AP043.pdf.

U.S. General Accounting Office (GAO). (2001). *Record Linkage and Privacy: Issues in Creating New Federal Research and Statistical Information*. Report GEO-01-126SP. Washington, DC: GAO.

University of Baltimore (UBalt). (2011). *About ADARE*. Retrieved 30 November 2011 from http://www.ubalt.edu/jfi/adare/about-ADARE.cfm.

White, Michael. (2010). *2003 National Survey of College Graduates Nonresponse Bias Analysis*. Paper presented at the 2010 Joint Statistical Meetings, Vancouver, Canada.

Tables and Figures

Figure 1: Levels of Data Analysis

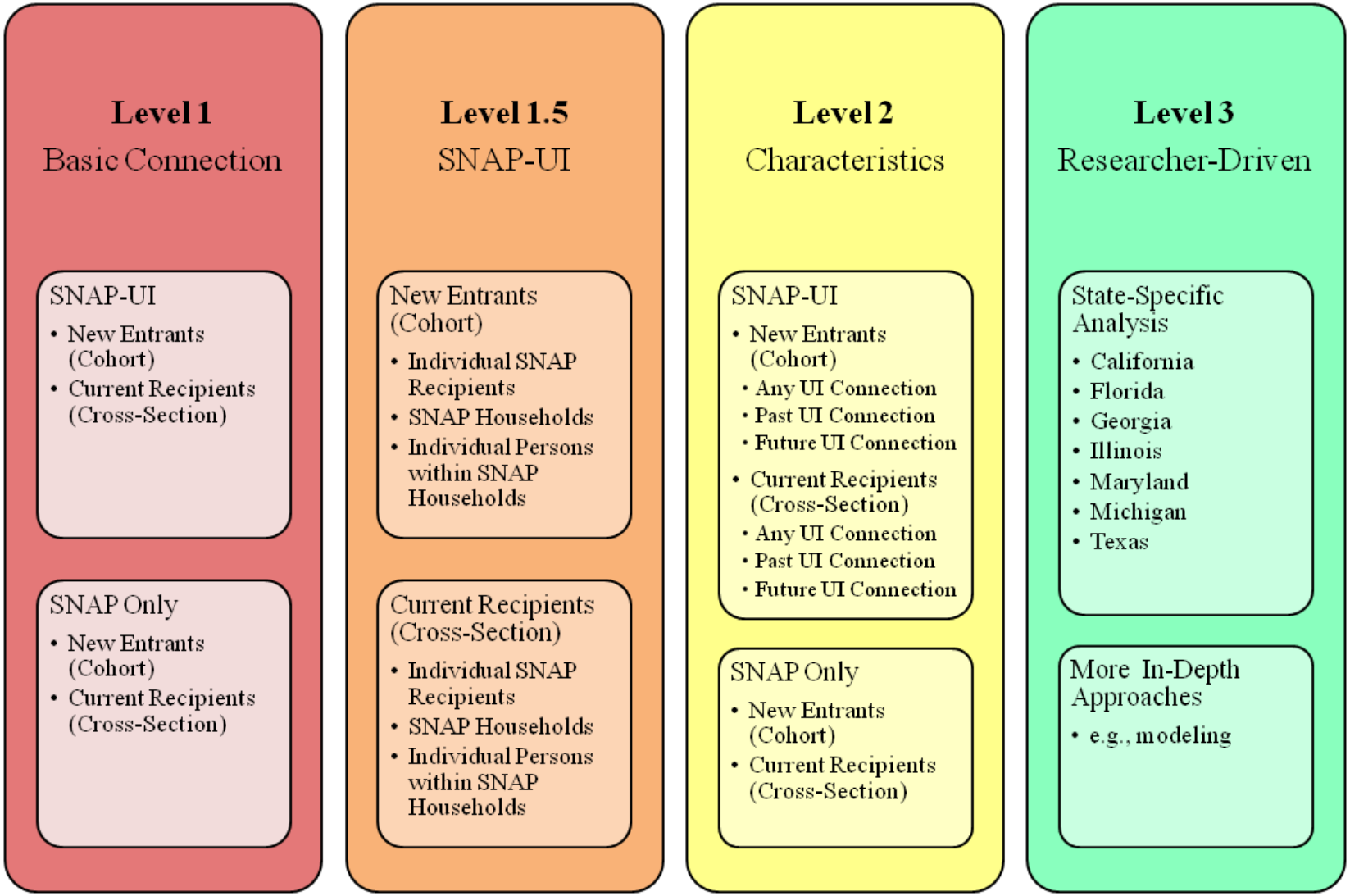

Figure 2: SNAP Receipt in the Seven Project States, Jan. 2005-Oct. 2010

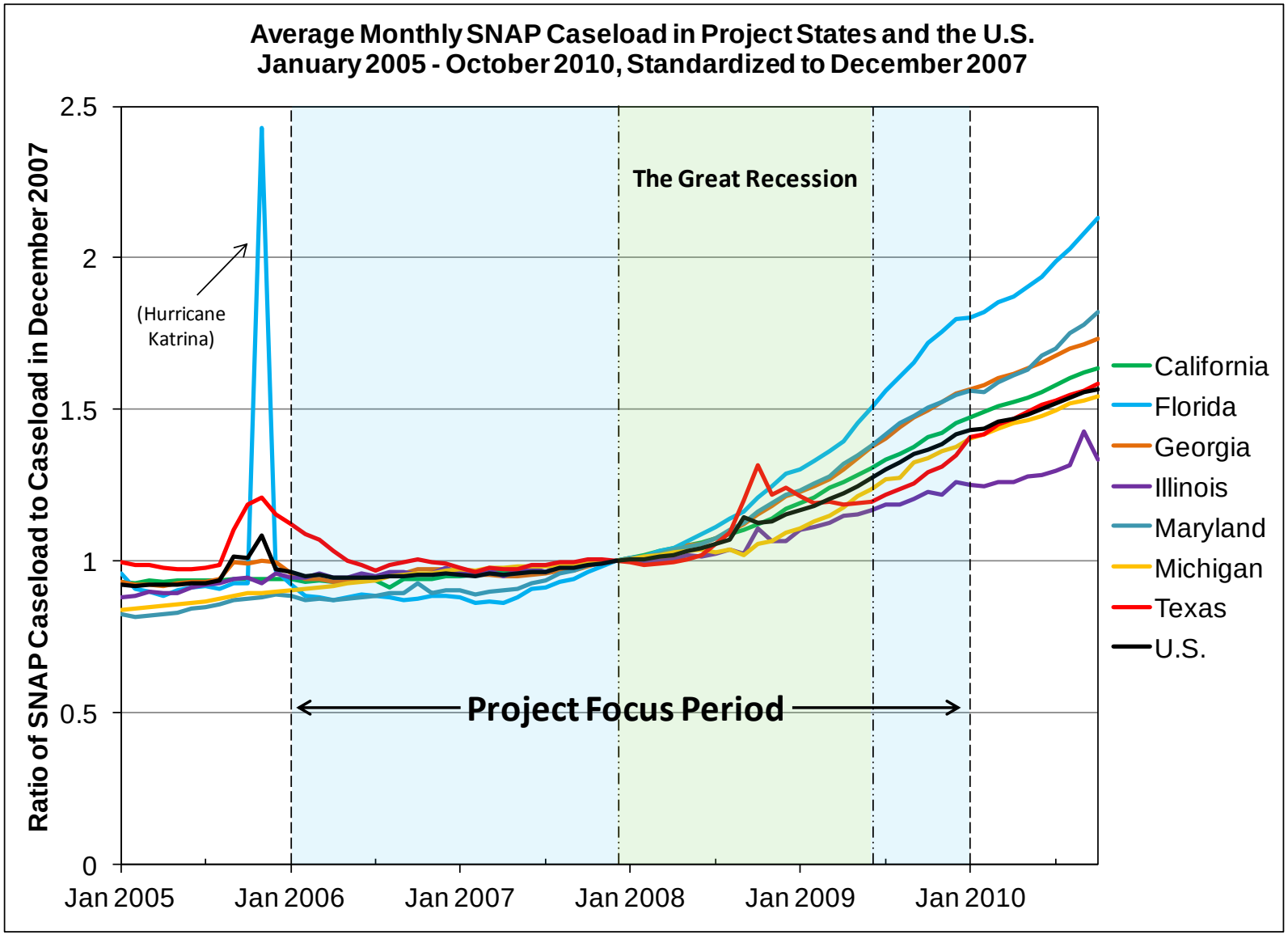

Figure 3: UI Receipt in the Seven Project States, Jan. 2005-Oct. 2010

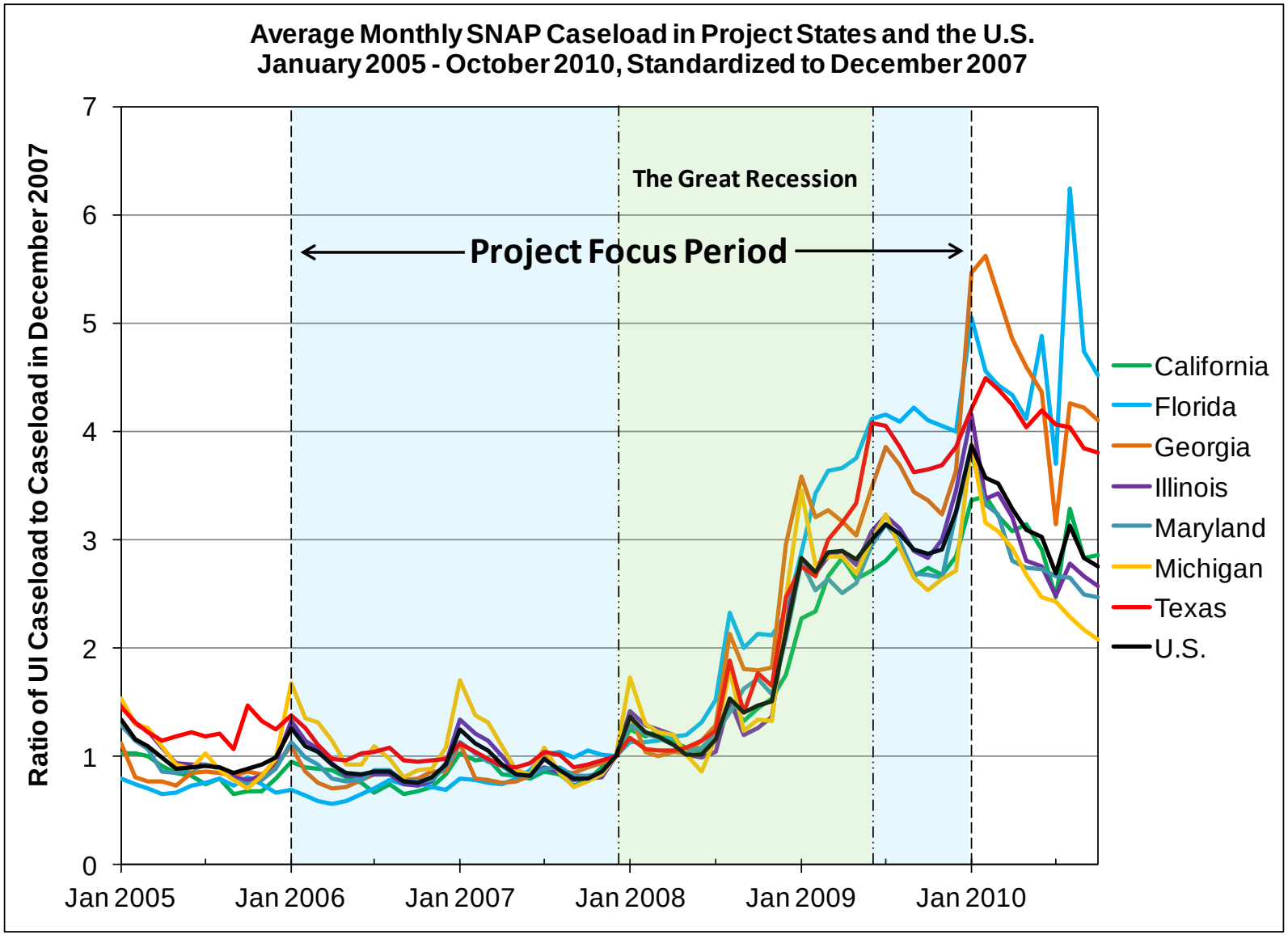

Table 1: Demographics of SNAP-UI Project States

		California	Florida	Georgia	Illinois	Maryland	Michigan	Texas	U.S.
Population		36,308,527	18,222,420	9,497,667	12,785,043	5,637,418	10,039,208	23,819,042	301,461,533
Population Rank (of 51)		1	4	9	5	19	8	2	N/A
Racial Demographics	**White, Non-Hispanic**	42.5%	60.5%	58.4%	65.2%	57.7%	77.5%	47.8%	65.8%
	Black, Non-Hispanic	6.0%	14.8%	29.4%	14.5%	28.5%	13.9%	11.3%	12.1%
	Hispanic or Latino	36.1%	20.6%	7.7%	14.6%	6.6%	4.0%	35.9%	15.1%
	American Indian, Non-Hispanic	0.5%	0.2%	0.2%	0.1%	0.2%	0.5%	0.3%	0.7%
	Asian, Non-Hispanic	12.5%	2.3%	2.9%	4.2%	4.9%	2.4%	3.4%	4.5%
	Other/Multiracial, Non-Hispanic	2.4%	1.6%	1.4%	1.4%	2.0%	1.7%	1.3%	1.9%
Median Household Income (2009$)		$60,392	$47,450	$49,466	$55,222	$69,475	$48,700	$48,199	$51,425
Population in Poverty	**Individuals Below Poverty Line**	12.9%	12.9%	14.6%	12.1%	8.0%	14.1%	16.3%	13.1%
	Individuals Under Age 18 Below Poverty Line	18.0%	18.0%	20.0%	17.0%	10.1%	19.5%	23.4%	18.3%

Source: U.S. Census Bureau, American Community Survey, 5-year estimates, 2005-2009

Table StateDemos, 15 December 2011

Figure 4: Concurrent SNAP-UI Receipt as a Percentage of the Working Age SNAP Caseload

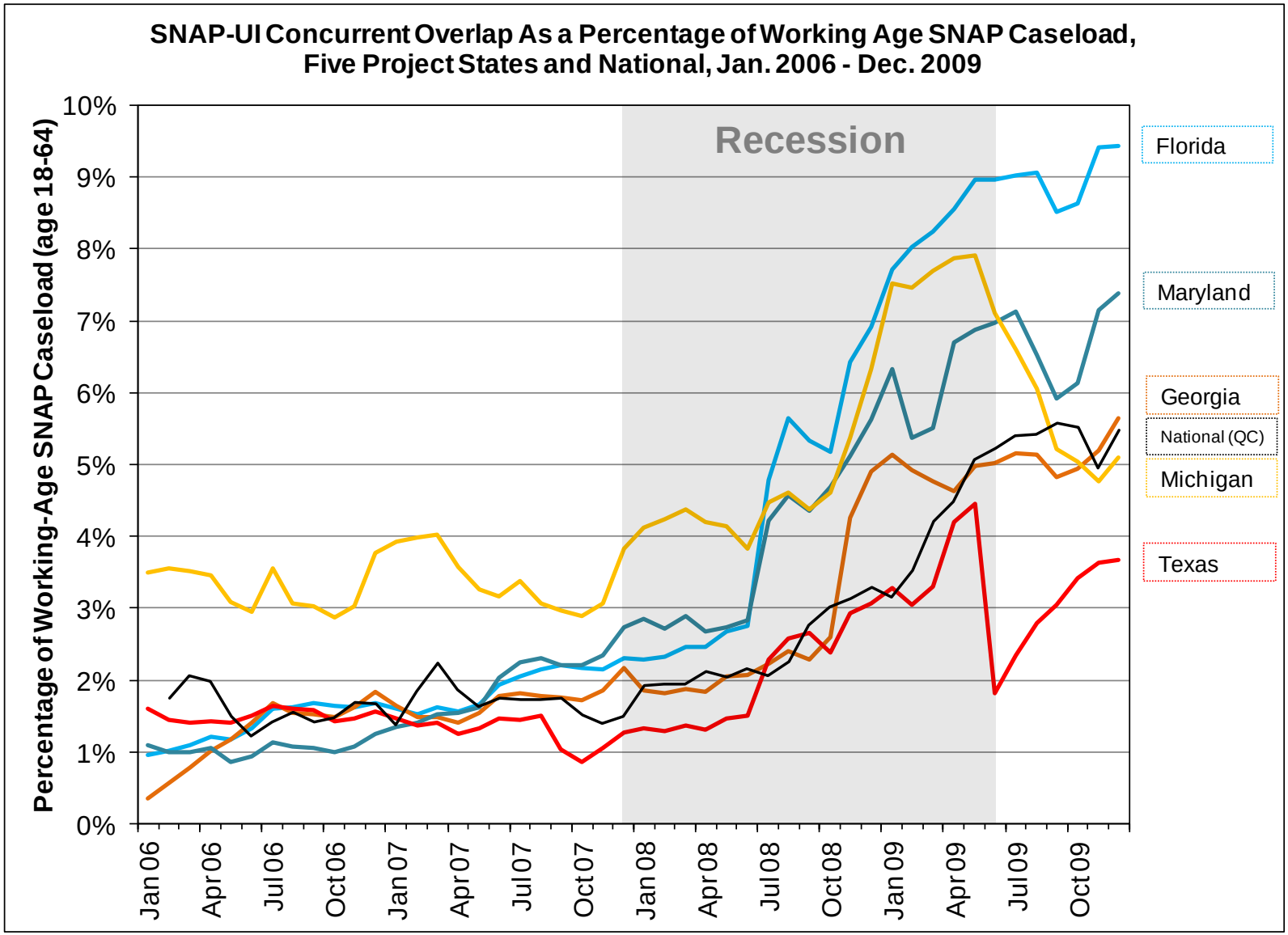

Figure 5: Concurrent SNAP-UI Receipt as a Percentage of the UI Caseload

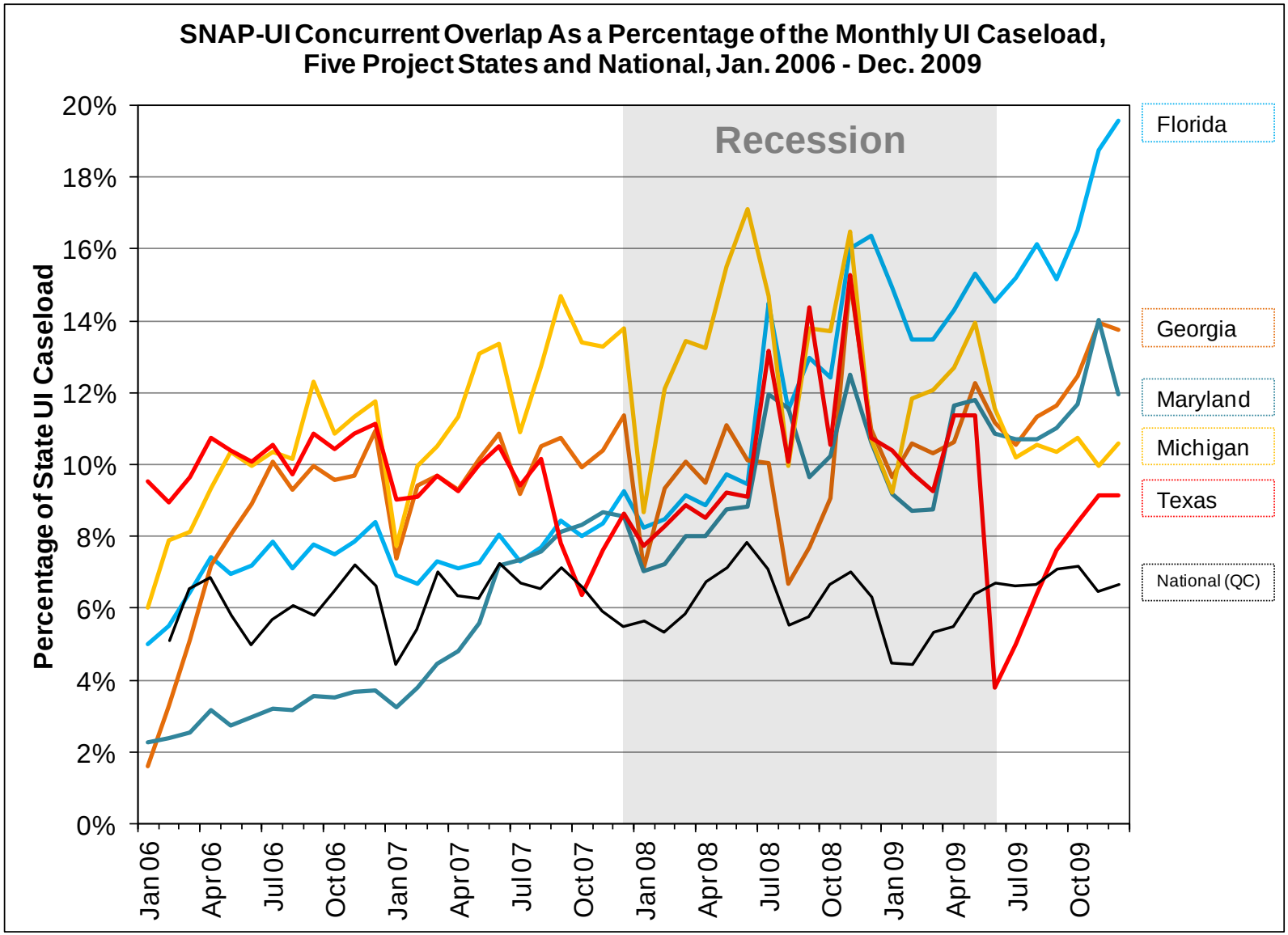

Figure 6: Percent of Current SNAP Adults Who Are UI Age-Eligible with Prior or Current Unemployment Insurance Benefits

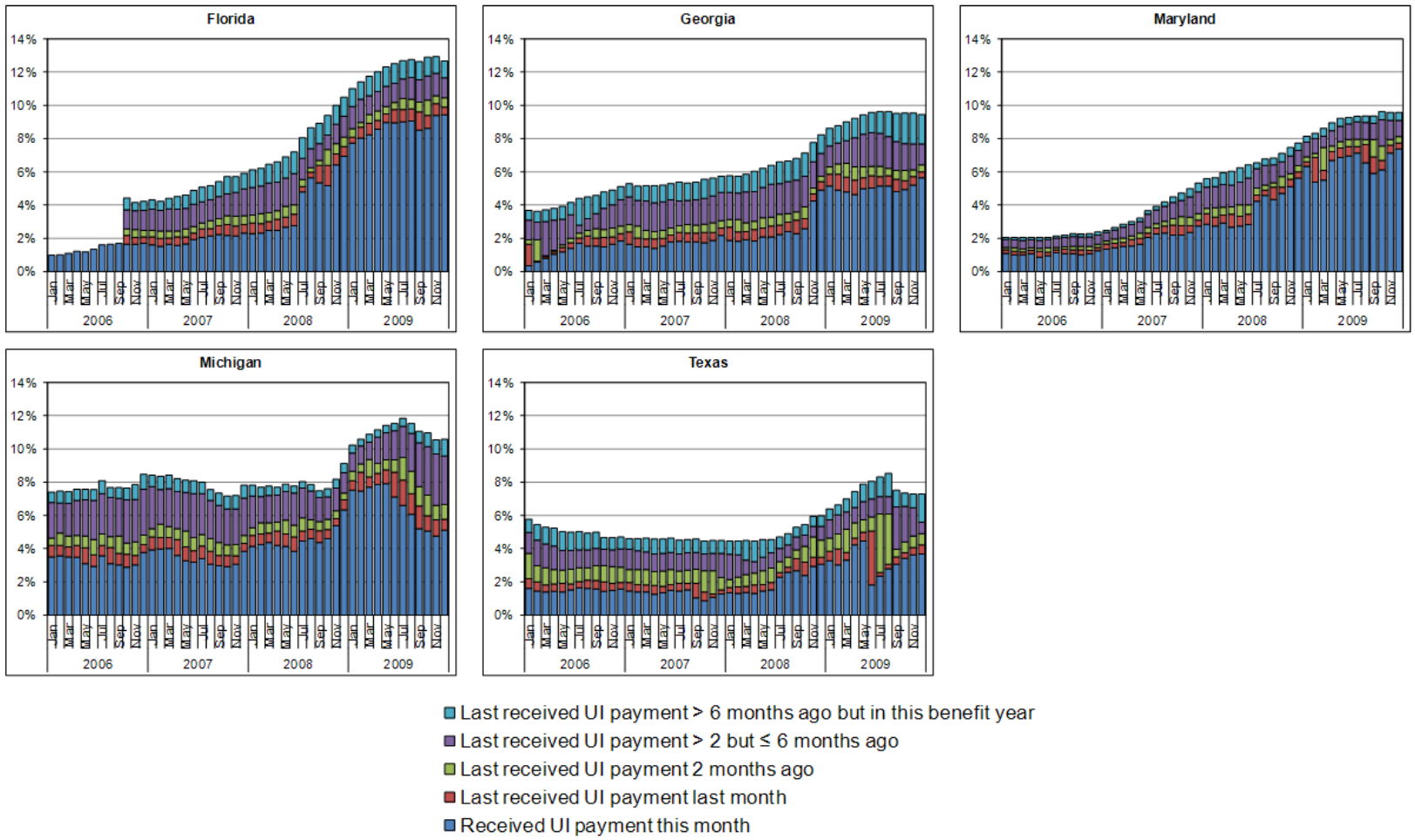

Figure 7: Percent of New SNAP Adults Who Are UI Age-Eligible with Prior or Current Unemployment Insurance Benefits