This is the Final Report for the project, "Analysis of the Current Population Survey Data for Food Security and Hunger Measurement" conducted by Mathematica Policy Research, Inc. (MPR) for the USDA Food and Nutrition Service (FNS), beginning in 1997. The project provided USDA with technical support and statistical estimation work for analyzing the 1996 and 1997 data on food security collected in the U.S. Census Bureau, Current Population Survey (CPS) Food Security Supplement. More broadly, the work examined a number of analytic and empirical issues relevant to analyzing the first three years of CPS food security data available—those for 1995, 1996, and 1997.

It was originally intended that the Final Report would provide the main vehicle for dissemination of the substantive findings on the prevalence of food insecurity based on the 1996 and 1997 data. However, because of the importance of making these results available as early as possible, USDA elected to issue an "Advance Report," thus making the results of the 1996 and 1997 analyses conducted by MPR available before completion of the overall project. In addition, since 1999, a number of publications have become available that present estimates of food insecurity prevalence, as well as discussions of the methods used in computing food security estimates in general. Most important, Andrews et al. (2000) provides a comparative analysis of the annual data for the five-year period 1995 through 1999, while Bickel et al. (2000) provides a how-to guide for measuring food security that incorporates relevant work done prior to that time, including earlier work from the current project. Selected issues in food security are also considered in Ohls et al. 1999.

In light of these developments, USDA suggested that MPR recast this Final Report to focus on several selected topics related to the 1995-1997 data, rather than provide comprehensive treatment of the overall research, much of which has since been incorporated in later publications. The Final Report has been organized around these suggestions.

Among the issues addressed in the report are:

- The stability of the food security measurement scale over time
- Temporal adjustments to the categories or designated ranges of severity on the underlying continuous scale used to classify households by food security status
- Screening issues related to ensuring a strictly comparable analysis sample over the 1995-1997 CPS food security samples
- Alternative imputation strategies for dealing with missing data
- The degree to which household responses to the food security questions are "modal," in the sense that households consistently respond affirmatively to questions involving less severe food insecurity whenever they respond affirmatively to questions involving relatively more severe food insecurity
- The degree to which the estimated parameters of the model used to measure the severity of food insecurity vary across different groups of households, defined by ethnicity and other characteristics
-

The first section below provides background information about the analysis. Subsequent sections summarize findings on each of the above issues.

Background

The analysis in this report is based on a statistical procedure which assigns households to food security status, based on their answers to a series of 18 survey questions. The food security categories used are:

- Food secure
- Food insecure—no hunger
- Food insecure with hunger

The data used for national-level analysis of food security are from annual supplements to the CPS, which is fielded monthly to more than 40,000 U.S. households.

Households are classified by food security status in the analysis, based on a procedure, Rasch modeling, which has a long history in the statistical literature. The first work in applying the Rasch model to food security data was undertaken under an earlier contract let by USDA (Hamilton et al. 1997). The Rasch model, as used in that work, posits that there is a single, one-dimensional attribute among households that indicates food insecurity. The model then uses a set of assumptions and statistical methods to assign "severity levels" to each of a series of 18 survey questions relating to food insecurity and hunger. A continuous food security measure is then assigned to each household in the data set, based on households' replies to the 18 questions. Supplemental procedures developed by Hamilton et al. are then used to translate the continuous scale score into a limited number of discrete food security statuses.

The objectives of the current project were to extend the analysis to 1996 and 1997 data and to address a number of related issues associated with measuring food security over time. Our findings in selected areas are summarized below.

Findings

Stability of the Parameters of the Model Over Time

An important issue in examining the validity of the Rasch modeling approach is whether the model parameter estimates are stable over time. The underlying theory on which the Rasch model is based posits that, if the wording of an item does not change, its estimated level of severity should not change. For example, even if food insecurity became more prevalent over time, a household at a given level of insecurity this year is expected to answer each item the same way a household at that level of insecurity did a year earlier. Due to sampling variability and other factors, such as minor wording changes, we do not expect estimated model parameters to remain exactly the same over time; but a finding of major changes over time would call into question the validity of the model. Particularly problematic would be a finding of important changes in the ordering of the items by severity.

To examine issues of model stability, we estimated the model independently on three CPS data sets (1995-1997), using consistent conventions as to statistical scaling. Some variation across years was found, as expected. In general, however, the estimated parameters of the model were quite stable. Also, the estimated order of severity of the different questions remained largely constant, with the only changes in severity order occurring among questions that were very close to each other on the severity scale in the original estimation work. The conclusion of this component of the research is that the food security model is sufficiently temporally stable to make it a reasonable tool to use in time series analysis.

Adjusting "Cut Points" Used to Classify Households into a Limited Number of Food Security Status Categories

The Rasch model places each household on a continuous numeric food security scale. For purposes of policy analysis, it is also useful to establish numerical "cut points" that assign households to a small number of designated categories which summarize their food security status. To create this categorical measure, Hamilton et al. (1997a and 1997b) specified four categories: food secure, food insecure without hunger, food insecure with moderate hunger, and food insecure with severe hunger. More recently, the latter two categories have usually been collapsed to form a single category, while

additional scale development work has identified a new nested category, food insecure with children's hunger (Nord and Bickel 1999 and 2001).

A key issue that arises in this work is whether it is appropriate to keep the same continuous scale cut points over time, or whether, alternatively, some temporal adjustments may be needed. The analysis of the body of the report concludes that, at least in some situations, it is not optimal to attempt to classify households based on the same cutpoints over time.

While the Rasch model places households on a continuous food security scale, due to certain statistical properties of the model substantial numbers of households tend to be clustered at certain points in the scale. If cut points are held constant, there is a risk that, because of chance statistical variation, the score assigned to one of these clusters of households might accidentally cross one of the cutpoints in a given year, causing considerable instability in estimates of food security prevalence.

Chapter V of the Final Report identifies several technical approaches for avoiding this difficulty. The discussion is based on the principle that a household with a given pattern of survey answers should always be classified into the same food security grouping, independent of when the data are collected.

Screening Household into the Sample in the 1995-1997 Surveys

The food security supplements in the 1995-1997 CPS had two general sections. The first section gathered information about food expenditures, participation in several programs aimed at providing food to needy families (for example, food stamps and school meal programs), and the sufficiency of food eaten during the preceding 12 months. The second section gathered more-detailed information about food insecurity and coping behaviors during the previous 12 months and prior 30 days. Not all households were asked this second set of questions, which includes the questions used to construct the food security scale. In order to minimize respondent

burden, households who, on the basis of earlier questions, appeared to have a high likelihood of being food secure were excluded from the more detailed questions and were assumed to be food secure in the analysis. This prescreening applied to higher income households in all three years, 1995-1997, and in one year, 1996, it was applied to lower-income households as well. Beginning in 1998 and continuing consistently since then, the CPS Food Security Supplement has included a new, less restrictive, pre-screen applied to higher-income households.

To ensure comparability in the analysis samples for the three years, the current research developed a common screen, such that any households giving survey answers that passed this common screen would have been tracked into the food security module in any of the three years. Households that did not pass the common screen were, for purposes of the analysis, treated as if they had not been tracked into the food security module of the survey—essentially, they were assumed to be food secure. Technical details concerning how this common screen was constructed are provided in Appendix B of the Final Report.

While use of the common screen has the desired effect of ensuring consistency in the 1995-1997 analysis samples, it also has the effect of treating as food secure a number of households who, during the survey, gave indications of experiencing food insecurity. Across the three years, use of the common screen was found to result in estimates of the prevalence of food insecurity which are between 1.0 and 1.5 percentage points lower than those that are obtained when the maximum available samples are used in the estimation.

Imputing Missing Data

Most households gave complete answers to the food security questions they were asked in the CPS; however, a limited number did not. Appendix C of the report examines a number of alternative approaches for including households with partially missing data in the analysis. One approach is reliance on the Rasch model itself, which has the capacity to assign food security

scale scores to observations with incomplete data. However, as is noted in Appendix C, in some instances, the determinations made within the model for cases with substantial amounts of missing data may lack face validity. Also, as a practical matter, many researchers may not have easy access to the software needed to implement the model.

An alternative algorithm for dealing with missing data has therefore been developed. Depending on the exact configuration of food security module answers given by the respondent, this alternative algorithm essentially involves imputing missing data items based on either (a) the highest severity item, in terms of level of food security severity, that the respondent answered positively; or (b) the lowest severity item answered negatively.

"Modality" of Household Food Security Response Patterns

The Rasch model implies that many households will exhibit item response patterns that are reasonably "modal" in the sense that if a household answers "yes" to any of the items, it will tend to answer "yes" to the less severe items, then answer "no" to the more severe items. A household that exhibits this pattern exactly—a string of all "yes" answers followed by a string of all "no" answers—is said to be a "modal" household. There is nothing in Rasch theory that predicts that all households will be modal; indeed, the model cannot be estimated if all households are exactly modal. Still, it is of interest in understanding the data to examine the degree of modality present. A large number of strongly nonmodal response patterns could call into question the validity of the model.

Analysis of the 1997 data indicates that most household response patterns tend to be either exactly or approximately modal. Of those households in the 1997 data who gave an affirmative answer to at least one question, approximately 39 percent households provided answer patterns that were exactly modal, while another 36 percent gave sets of answers which had only a single nonmodal response.

Consistency of Estimated Food Security Model Parameters Across Population Subgroups

Essentially, the analysis conducted with the aggregated CPS data sets assumes that different subgroups of the population are similar with regard to how they experience food insecurity. To test this assumption, the Rasch model was estimated separately for subgroups of the population, defined according to (a) race/ethnicity; (b) household composition; (c) metropolitan status; and (d) region of country.

The results of this analysis indicate considerable robustness of the analysis to this kind of disaggregation. In general, estimated severity levels for the individual questions were found to have consistent patterns across different subgroups, and the magnitudes of the parameters do not change substantially.

There is no clear statistical test of how much difference in the estimated subpopulation models would affect confidence in the overall modeling approach. However, the judgment of statistical experts who have used the Rasch model extensively in other contexts is that the findings of the subgroup analyses can reasonably be judged to be highly consistent with one another.

Conclusions

Reflections on the Strengths and Limitations of the Food Security Methodology

We conclude by discussing the strengths and limitations of the use of the Rasch model as a basis for food security measurement. Possible directions for future research are also noted. The food security scale reflects more than 10 years of methodological development by both government and private groups. The use of the Rasch model methodology has made it possible to guide the development of the food security estimates with a thoroughly studied model that has well understood statistical properties. In terms of goodness-of-fit criteria, the mathematical form of the measurement model shows strong correspondence on "fit" to the

empirical data. The approach has undergone extensive review by experts in both the public and the private sector. In general, these experts consider the model an appropriate application of the IRT methodology, and they have viewed the analysis results as reasonable.

Another important strength, as established by the current project, is that the estimated item parameters of the IRT model are robust across time and population subgroups. The values obtained from the 1996 and 1997 data are essentially the same as the original 1995 values. In addition to stability over time, there is stability across subgroups, defined by such characteristics as race/ethnicity, household composition, and region of the country.

Tempering these strengths are a number of limitations which should also be recognized. Most of these, if not all, are a matter of careful interpretation of what the food security measure does and does not do. For example, the CPS indicator questions for food insecurity and hunger and the scale developed from them are designed to provide a household-level measure of the severity of conditions as experienced within U.S. households. This is in line with the conceptual understanding of food insecurity as a condition of deprivation or stress experienced by households in meeting members' basic food needs. However, the experience of hunger as such, which appears only at a more severe stage of food insecurity, is strictly individual. The household classification, "food insecure with hunger" refers to that more severe range where evidence of reduced food intake and hunger has appeared for one or more household members. But this is a collective measure which may apply to all household members, to adult members only, or to as few as one (adult) member.

Second, the basic measure is designed to capture respondents' experiences over the course of a year, while household circumstances can change markedly during such a period. Accordingly, the 12-month measure—designed to provide reliable benchmark and trend figures—may not represent the current situation of given households. Similarly, the "food insecure with hunger" designation can, in principle, result from just one

serious episode during the year, although for most such households evidence of a repeated pattern of reduced (adult) food intakes during the year must be established.

In addition, a number of issues of interpretation flow from the need to have a simple categorical measure as a means of classifying households for purposes of manageable data reporting and monitoring, in addition to the underlying continuous scaled measure. The categorical measure was created to make the scale more accessible to non-technical users and more convenient to users whose needs could be better served by a simple categorical variable than by the detailed continuous measure. The categorical measure as such is straightforward: it represents designated ranges of severity along the continuous scale (i.e., qualitatively differing severity levels of "food insecure"), plus the category of households that either show no evidence of food problems within the CPS data set, and hence can be deemed to be "food secure," or that show only one or two indications of food stress, which is deemed insufficient as evidence to establish confidently their status as "food insecure."

The interpretive problems with the categorical measure stem from at least three sources. First, the designation of appropriate severity ranges, and their exact delineation in operational form based on the available set of indicators, is inherently judgmental and thus leaves room for disagreement.

Second, the Rasch model employs a probabilistic logic in generating the continuous scale measure of severity of household food insecurity; similarly, the corresponding severity-range summary categories share this probabilistic nature. However, the naming conventions adopted for the severity-range categories are determinate in form, which can be misleading.

Thirdly, a misplaced specificity and determinateness can easily be attributed to the individual indicator items as well, causing a misunderstanding of their actual role in the measurement process.

To illustrate this last point, straightforward names adopted for the severity-range categories raise issues of face validity when they seemingly contradict the clear language of particular indicator items embedded within the measurement scale. For instance, it is technically possible for a household to be classified "food insecure with hunger," even though the respondent has answered "no" to the particular question, "In the last 12 months were you ever hungry but didn't eat because you couldn't afford enough food?" In this case, the respondent either must have replied "yes" to a series of increasingly severe indicators of food insufficiency, including at least three items indicating reduced food intake for themselves and/or other adult members of the household, one of which establishes a repeated pattern of such reduced intakes over the year, or they must have replied "yes" to most of the foregoing, plus one or more of the items that are more severe than the explicit hunger question. The categorical measure (and its naming convention) reflects the judgment that, on the balance of this evidence, one or more adult members of the household has, with high probability, experienced resource-constrained hunger sometime during the year. Conversely, the opposite case also can occur: the household can be classified "food insecure without hunger," based on its overall pattern of response and the resulting scale score, even though the respondent has answered "yes" to the explicit hunger question as such.

In creating the scale, a number of steps were taken to minimize the effects of these factors. For instance the numerical cutpoints defining the categories were set to be conservative, in the sense that there must be three answers to questions thought to indicate food insecurity before a household is classified as food insecure, and similarly for the hunger classification. Also, analysis presented in the text of the report indicates that substantial numbers of respondents follow close-to-expected, response patterns, which do not lead to any apparent anomalies in classification. Nevertheless, room for disagreement remains as to what types of answers to the questions should be construed as reflecting the language used in designating the three scale categories.

A possible solution to some of these issues would be to state the category names in more-probabilistic terms, such as "probably food insecure" or "a high likelihood of hunger." This would be in keeping with the probabilistic nature of the underlying model, and it might help ease the concerns of those who are bothered by the anomalies posed by apparently inconsistent patterns of question responses. However, such category name changes might also interfere with the clarity of meaning of the categories themselves, thus reducing their effectiveness.

Overall, it is important to recognize that these limitations have not prevented the food security scale from becoming an important, widely used research and policy tool. Questions to support the scale have been included in an increasing number of national surveys and scale results are frequently cited in the policy process. This evidence suggests that many policy analysts have found the scale to be a valuable tool for measuring an important aspect of material deprivation among America's poor.