Combining Quantitative and Qualitative Data in Welfare Policy Evaluations in the United States

Andrew S. London*
Syracuse University
Email: aslondon@maxwell.syr.edu

Saul Schwartz
Carleton University
Email: saul_schwartz@carleton.ca

Ellen K. Scott
University of Oregon
Email: escott@uoregon.edu

*Earlier versions of this paper were presented on the panel on integrating qualitative and quantitative research methods at the 24th annual APPAM Research Conference, Dallas, Texas, November 7-9, 2002 and at the Q-Squared in Practice Conference on Experiences of Combining Qualitative and Quantitative Methods in Poverty Appraisal, University of Toronto, Toronto, Ontario, Canada, May 15-16, 2004.

Q-Squared Working Paper No. 12
October 2005

Acknowledgements

This paper was written with support from MDRC's Next Generation Project, which is funded by the David and Lucile Packard Foundation, the William T. Grant Foundation, and the MacArthur Foundation. We thank Martha Bonney for editorial assistance.

Introduction

As is the case among researchers working on issues of economic development and poverty in developing countries (Kanbur 2003), there is substantial interest in the promise of mixed methods approaches among researchers evaluating the effects of welfare reform and related policies and programs in the United States. This interest was evident among the participants at a conference entitled "Mixed Methods Research on Economic Conditions, Public Policy, and Family and Child Well-Being," which was hosted in June 2005 by the National Poverty Center at the University of Michigan. Many of the concerns and tensions that have emerged in the discussion of mixed-methods approaches in poverty research in developing countries also exist among poverty researchers working in the United States.

For many years, researchers working on MDRC's Project on Devolution and Urban Change and the Next Generation Project have been attempting to integrate analyses of structured interview data collected in the context of both population-based surveys and experimental evaluations with narrative data collected during lengthy, conversational interviews with welfare-reliant and working poor women. Our goal, always, has been to try to better understand how welfare reform and employment policies affect low-income women, their children, and their families. Partly out of necessity and partly by design, our strategy has generally been to integrate across methods at the data analysis stage and to use the findings from the analysis of one data source to confirm, challenge, extend, complement, supplement, or nuance the findings emerging from the other. Rightly or wrongly, the anthropologists, developmental psychologists, economists, public policy analysts, and sociologists working on our interdisciplinary teams tend to

reduce the distinction between qualitative and quantitative to the difference between data obtained respectively from structured interviews (quantitative) and interviews primarily involving open-ended questions (qualitative). We also tend to think of all attempts to integrate these two types of data as mixed-methods research.

Following the first Q-Squared conference, Kanbur (2003) proposed five key dimensions that can be used to situate studies along the qualitative-quantitative continuum: (1) non-numerical versus numerical information on the population; (2) specific to general population coverage; (3) active to passive population involvement; (4) inductive to deductive inference methodology; and (5) broad social science versus neo-classical economics disciplinary framework. We view our attempts to mix methods in these evaluation studies as attempts to be more broadly social scientific by drawing on both numerical and non-numerical information to generate general and specific information about the focal population and the effects of policies through inductive and deductive approaches to inference. All of the primary studies that we will discuss in this paper were designed as evaluations of policies and programs. As such, they were largely driven by researchers' interests with limited active involvement of members of the population. The qualitative interview studies allow for somewhat more active participation in the interview process, but even in those situations, the prerogatives of the researchers predominate.

In this paper, we review a few of the ways we have used these quantitative and qualitative data in our work and discuss some of the promises and pitfalls of combining qualitative and quantitative research methods and data in poverty appraisal and policy evaluation research. Along with the Q-Squared conference participants represented in

this special issue, Kanbur (2003), Gibson & Duncan (2005), participants in the conference at the National Poverty Center conference, and a growing number of other researchers in the United States, we believe there are synergies to be gained from the careful combination of mixed research methods. Drawing on our own experience and the reports of others, we identify several factors that we think will enhance the value of combining qualitative and quantitative methods and data in poverty appraisal research and policy evaluations. However, we also caution that all methods and data have their limitations, and the goals of quantitative and qualitative researchers sometimes legitimately differ. Researchers must be careful to specify the types of questions that these different types of research methodologies and data can and cannot answer, and the kinds of contributions they can and cannot make. To fail to do so risks raising expectations too high.

Data Sources

The population-based survey data and the qualitative interview data referenced in this paper were collected between 1997 and 2001 under the auspices of MDRC's Project on Devolution and Urban Change.[1] Urban Change is a multi-city, multi-component study of the effects of the 1996 welfare reforms in the United States on welfare administration, caseload dynamics, poor women and their children, community-based social service organizations, and urban neighborhoods. Urban Change had several components that were directly linked: the administrative records component, which involved analysis of the universe of cases in each city over a 10-year period; a neighborhood indicators component, which focused on census-tract level change over the same 10-year period; and a survey component, which involved the collection of two waves of data from a

random sample of women selected from the administrative records. The qualitative interview studies were designed to be nested loosely within but not directly linked to other components of Urban Change. The qualitative interview component was also designed to articulate with an implementation component, which focused on changes in welfare offices, and an institutional component, which focused on the effects of welfare reform on community-based social service agencies. The offices and agencies that were studied in the latter two components were located in the same "neighborhoods" (i.e., clusters of census tracts with particular characteristics) in which the women in the qualitative interview samples lived. Thus, the overall design of Urban Change included comparative city and neighborhood studies nested within a larger, quasi-experimental analysis of administrative records and neighborhood change, and a population-based, longitudinal survey.

The first wave of the Urban Change survey involved in-person interviews in 1998-1999 with women who had been single mothers receiving welfare and/or food stamp benefits in May 1995 and who were living in neighborhoods (census tracts) characterized by high rates of poverty (30 percent or more of households) or welfare receipt (20 percent or more of households). Using data from administrative records, women who met these criteria were randomly selected, with a sample size targeted to be 1,000 completed interviews per site. In-person interviews were completed in 1998-1999 with 3,960 women (78.6 percent of those sampled three years earlier). Follow-up interviews with the same women were completed in 2001.

The longitudinal, qualitative interview component of Urban Change involved 3 or 4 semi-structured, in-person interviews conducted between 1997 and 2001 with

approximately 40 women per site. These women lived in neighborhoods of concentrated poverty similar to the neighborhoods in which the survey respondents lived; however, no woman participated in both the survey and longitudinal, qualitative interview studies. Racial/ethnic criteria were also used to select respondents for the qualitative studies. In Cleveland and Philadelphia, the cities represented most consistently in the reports discussed below, all of the women in two neighborhoods per city were African American, while in one neighborhood per city they were white.

The qualitative interview samples were not directly linked to or embedded within the sampling frame for the survey component because the interdisciplinary team of researchers working on the qualitative interview component believed it was critical to recruit women through means that were completely independent of the welfare office. Thus, the research teams recruited women in the community from selected census tracts that had the same poverty and welfare receipt rates as the census tracts from which the survey sample was drawn. As a result, the qualitative and quantitative components of Urban Change are largely independent of one another even though they were designed in tandem, were fielded over the same time period, and included women from the same population. Despite the fact that there was no explicit linkage between the survey and qualitative interview components of Urban Change, we have tried to integrate them at the data analysis stage to add breadth and depth to our answers to specific questions. The structured survey and open-ended, qualitative interviews covered similar topics; however, the qualitative interviews yield richer, narrative data about how families are coping and what they are experiencing, which arguably is an important basis for policy evaluation and formulation.

Under the auspices of MDRC's Next Generation Project, we have also used the qualitative interview data from the Cleveland and Philadelphia sites of Urban Change to inform analyses of structured interview data from a set of experimental evaluations of the impacts of welfare-to-work programs. These experiments were conducted to evaluate the effects on low-income families of programs designed primarily to increase parental employment, although some also had income supplementation components. All of the experiments were started in the early to mid-1990s under waivers of the federal rules governing Aid to Families with Dependent Children (AFDC) granted to individual states. In each study, the heads of low-income households in a given location — most of them welfare recipients — were randomly assigned to a program group or a control group. Program group members were eligible for the services and subject to the requirements of the program under study, while the control group in most sites remained subject to the eligibility requirements and benefits of the existing AFDC system.[2]

These experimental data provide strong evidence of causal impacts, but often do not allow us to see within the black box to surmise how the policy or program might have had its impact. Although the qualitative interview data were not collected from the same participants or in the same period, they were collected from women who were subject in a later period to many of the same policies as the women in these experiments. Thus, we used the qualitative interview data to help us think about processes that might have been affected by such policies, to nuance the conclusions emerging from the experiments, and to suggest hypotheses that could be tested.

Combining Survey and Qualitative Interview Data

Even though the survey and qualitative interview samples in the Urban Change study were not linked, having complementary survey and qualitative interview data has been beneficial. The qualitative data have reminded us to consider the possibility of mismeasurement in population-based structured interview studies due to heterogeneity in the meanings people attribute to concepts, social desirability bias, the format of the interview, and the social context in which people are living their lives. The qualitative data are useful for scratching beneath the surface of prevalence estimates to suggest some of the dynamic cognitive and behavioral processes that underlie static, point-in-time survey estimates. Below, we discuss several examples that illustrate this set of benefits.

Measuring Food Insecurity

Polit, London & Martinez (2000) combined data from the 1998 Urban Change survey and the baseline qualitative interview studies conducted in Cleveland and Philadelphia to examine food security and hunger. Food security in the survey was measured with the U.S. Department of Agriculture's Household Food Security Scale (HFSS). The HFSS consists of 18 items that can be used to classify respondents into one of four categories: food secure, food insecure without hunger, food insecure with moderate hunger, and food insecure with severe hunger (Carlson, Andrews & Bickel 1999).[3] The scale has been judged to be reliable and valid for both individual and population uses (Frongillo 1999), and has been administered by the U.S. Bureau of the Census's Current Population Survey since 1995. The women in the qualitative study samples were not administered the HFSS at baseline; however, they were queried about their food expenditures, use of Food Stamps, WIC, and other publicly-funded food

programs, food deprivation, and the use of emergency food services. These data were used to classify women into one of the above-mentioned food security categories. About half of the women in the survey and qualitative interview samples respectively were classified as food secure.

Although the survey and qualitative interview samples were not linked and the strategies for measuring food security differed, having both data sources available provided important insights that allowed these researchers to raise questions about the completeness of the HFSS as a measure of food security and highlight some of the limitations of self-reported survey data as measures of qualitative constructs such as "food security." Scales used in surveys often consist of items allowing only fixed categorical responses and researchers often fail to appreciate the heterogenous meanings that people ascribe to those categories. In-depth interviews point out some of this heterogeneity and suggest that analysts should have a more nuanced sense of how much confidence should be places in the survey responses and what the scale does and does not measure.

By design the HFSS measures whether families are food secure, but it does not address the effort that is expended in attaining food security. Carlson, Andrews & Bickel (1999: 512S) note that the HFSS was developed as a unidimensional scale using factor analysis and that "with one major exception" nearly all the items fit. The one general type of item that did not fit the model was "indicators of coping strategies that a food-insecure or at-risk household might use to improve its food supply from emergency sources, such as obtaining food from a food bank or borrowing money for food." Such "coping items" were explicitly excluded from the scale because they did not fit the

conceptual and empirical model that guided the development of the scale and were viewed by the scale developers as representing a domain that was distinct from "food security."

Most of the mothers in the qualitative interview samples who were classified as food secure expended considerable energy in piecing together numerous strategies to make sure that there was adequate food for themselves and their children. Women who were food secure talked a lot about managing their paperwork with the welfare office to maintain their Food Stamps, shopping carefully by sometimes purchasing day-old bread and reduced-price meats, and routinely relying on friends or relatives for meals or loans that enabled them to feed their families. The qualitative data suggests that the label "food secure" masks considerable variation in the efforts families expend to achieve food security and in the precariousness of their ability to maintain this status.

Researchers studying the impact of increased maternal employment resulting from welfare reform, especially when such employment does not appreciably increase income, should be concerned about the effort required to attain any level of food security, as well as the level of food security itself. Additional survey questions might capture women's efforts to feed their families but, as always, space on surveys is limited. At a minimum, survey analysts should recognize that the level of food security as indicated by this unidimensional scale is an incomplete measure that does not provide an indication of the effort it takes to put food on the table, the opportunity costs associated with that effort, or the precariousness of the status. We note that this is a traditional way in which in-depth qualitative interviews can inform the development of quantitative survey questions and the design of survey questionnaires.

The ethnographic data also revealed other issues of relevance to the analysis and interpretation of the HFSS categories. In keeping with the conceptual definition of food security as having adequate access to food "without resorting to emergency food supplies" (Anderson 1990), all of the women in the qualitative studies who used food banks were classified by the coders as food insecure. However, it is possible that these women might have classified themselves as food secure, partially as a result of their strategic and planned use of the food bank resources available in their communities. This suggests that there might be some misclassification in the HFSS resulting from implicit normative assumptions by researchers about what the use of food banks means. Such assumptions may not map onto the normative views of low-income mothers who regularly rely on food banks to achieve a desired level of food security for their families.

Polit, London & Martinez (2000) noted that there were hints in the qualitative data that the women were proud of their ability to feed their children. They might not always be able to pay their utility, rent, and other bills, but they made sure there was adequate food on the table. There were also hints in the data that some women feared that if their children were not adequately fed, the children would be taken away from them. If there is pride associated with maintaining adequate food, and shame or fear in not being able to do so, it is possible that these concerns would lead some women to avoid endorsing survey items that are indicative of food insecurity. These biases would plague not only the HFSS but also any additional survey and qualitative interview efforts to monitor the process by which poor women attain whatever food security they have.

Contextualizing Assessments of Satisfaction with Living Standards

In a second example that continues to develop the themes from above, having both quantitative and qualitative data on women's assessments of their standard of living provided nuance to our understanding of what might otherwise have been a perplexingly unclear survey finding. Although half of the women in the Urban Change survey sample were living below the federal poverty threshold in 2001 and four out of five were near poor (i.e., within 185 percent of the poverty threshold), three out of four respondents reported that they were very satisfied or satisfied with their standard of living (Brock et al. 2002).

We asked women in the qualitative study in Cleveland to rate their standard of living and discuss why they felt as they did about it. Although the samples were not linked, what we learned from the analysis of the qualitative interview data provided insights that helped us make better sense of the survey estimate. The qualitative interview data indicated that women sometimes framed their relatively positive assessments of their standard of living in relation to their perception that many others were worse off than they were, or that they were better off than they had been previously. For example, several women said that they would not use food banks because they are only for people "who really need it." By framing their situations in such terms, women could maintain their self-esteem and optimism while resisting demoralization and the sense that they were somehow failing. This represents, we believe, a sort of cognitive coping strategy that may have affected how some of these women responded to the survey questions.

Sometimes, when women's assessments of their standard of living were negative, it was because financial constraints made it difficult for them to fulfill important social role obligations, such as being a good mother and provider. For example, Cindy[4], a white mother of two children lived with many housing problems in a neighborhood she considered unsafe, experienced unmet dental care needs and utility shut-offs, and had to skimp on her own food so that her children could eat. Cindy described herself as being "unhappy" about her standard of living because it made her feel "bad as a mom." She said:

> When I got my income tax back, I went out and I bought Kim, my oldest girl, clothes. She had not had new clothes in over a year and a half, or shoes. I mean, you know, and that hurts me as a mom…It hurts me that they would have to wait for a windfall like that, just to get what would be considered basic day-to-day clothes. I mean, my idea of basic day-to-day clothes is a pair of jeans for each day and stuff like…I mean, I'm not talking about going out to London for the weekend or something like that…And maybe able to go out…like being able to let them do the things they want to do, simple things like going roller skating, have extra money to do that sort of stuff. When you don't have it, you can't, you know, it hurts."

The narrative accounts we obtained from women living in some of the most disadvantaged neighborhoods in Cleveland provided context for the generally high estimates of satisfaction with living standards we obtained from respondents in both the survey and qualitative interview components. In these neighborhoods of concentrated poverty, the reference group that women use for making their assessments of their own

circumstances include families who they perceive to be worse off than they are. Additionally, poor women's coping strategies may include denial and other cognitive strategies that allow them to maintain their dignity and sustain their considerable efforts to meet their families' needs. Researchers using such subjective appraisals must be mindful of the ways that the context of respondents' lives shapes how they respond to questions in both survey and qualitative interview studies.

Under-Estimating Ongoing Domestic Violence

The notion that the context of people's lives shapes how they respond to questions, whether they are asked in structured survey questionnaires or open-ended interviews, was brought home to us forcefully in our work on domestic violence. In this instance, we see the potential for the under-estimation of domestic violence, which can be pieced together from two different research reports. The first study reported baseline survey and baseline qualitative interview findings from the four Urban Change sites (Polit, London & Martinez 2001). The qualitative interview data were used to illustrate the types of abuse women faced, but not to raise questions about the survey estimates. The second study reported longitudinal, qualitative data from the Cleveland site of the Urban Change project (Scott, London & Myers 2002a).

It is well documented that poor women are at disproportionate risk for domestic violence, and that many women on welfare have a history of abuse (Raphael & Tolman 1997; Scott, London & Myers 2002b; Tolman & Raphael 2000). The lifetime prevalence rate for welfare-reliant women is estimated to be 50 to 75 percent, with a past-year prevalence of 15 to 20 percent. Polit, London & Martinez (2001) reported that 8.8 percent of women in the Urban Change sample were hit, slapped, or kicked in the year

prior to the 1998 survey. Results from our longitudinal ethnographic study suggest that this may under-estimate the extent of recent or ongoing violence in this population.

In our qualitative interview sample of 38 women in Cleveland, approximately 60 percent of the women disclosed domestic violence (Scott, London & Myers 2002b). However, three women who were being abused during the course of the study did not disclose the abuse to the interviewer until after the abuse had ended. Despite very high levels of rapport between the interviewers, these three women responded "no" to all questions on domestic violence on a self-administered questionnaire, which was the same one that was used in the Urban Change survey, and to open-ended questions in the qualitative interview that took place the year prior to the qualitative interview in which the woman disclosed the abuse (i.e., these women reported no abuse in the second-year qualitative interview and disclosed it in the third-year qualitative interview). In each case, when they disclosed that they had been abused, the women indicated it had been going on at the time of the prior interview and sometimes throughout the entire course of the study. We illustrate this with the case of Maria about whom we have written elsewhere (see Scott, London & Myers 2002a).

Maria had been physically abused by her boyfriend, the father of her two youngest children, for eight years, until February 2000 when he moved out of the house. Maria said she was beaten at least once a month during that time, sometimes resulting in trips to the hospital for stitches. The interviewer perceived herself to have good rapport with Maria, and this was demonstrated in the many intimate details Maria had shared about her life with this interviewer. The interviewer reported being shocked when the

abuse was disclosed. Maria had had other opportunities in prior interviews to talk about the abuse, but all of them occurred while she was still living with the batterer.

There may be a number of reasons why Maria chose not to disclose the abuse until it had stopped, despite the high level of rapport between her and the interviewer. She may have been afraid her boyfriend would retaliate if he found out she had disclosed the abuse. She may have felt ashamed of the fact that she was being beaten, and somehow that shame was lessened when it had ended. Finally, she may have been surviving the abuse by denying either the fact of the violence or the extent of the violence. Denial becomes much more difficult if one acknowledges the violence to someone else.

Our experience with having three cases of domestic violence disclosed two years into a study involving extensive contact between interviewers and respondents surprised us. We think these data raise important methodological and policy concerns. Given that most studies cited in the literature on welfare reform and domestic violence involve single contacts, or at least significantly less contact with subjects than we had in the Urban Change study, and given that most welfare caseworkers asking questions to identify cases of domestic violence will likely meet with recipients for much shorter periods, we have reason to be very concerned about whether we are capturing the extent of the problem either in our research or in our attempts to provide public assistance. This is particularly true for those women who are currently being abused. Those are the women about whom we should be most worried and for whom we should seek the most flexible policies. If the shame, denial, and fear battered women typically experience keep some women from disclosing an ongoing history of domestic violence to a trusted

interviewer with whom they have been talking intimately for two years, why do we think this is going to be easy to identify in a welfare office in a half-hour meeting when uncertain benefits might be at stake? This poses an extraordinary conundrum for policy-makers. It also suggests that researchers in both developed and developing country settings who are studying domestic violence need to build trust through long-term contacts in order to monitor how changes in the social contexts of women's lives affect their willingness to report that they have experienced violence.

Combining Experimental and Qualitative Interview Data

In a recent paper, Gibson & Duncan (2005) argue that there are potentially large benefits to be gained from efforts to combine qualitative and experimental evidence. Although we did not have in place the two design features that they identify as contributing to the synergy they experienced in the New Hope Demonstration Project (i.e., an embedded and linked, randomly drawn sample for the ethnography and having the same analysts conduct both qualitative and quantitative analyses), we have also had experiences in which this combination of methodologies was beneficial. In this second section of the paper, we turn our attention to two examples in which we fruitfully combined qualitative and experimental data analyses in MDRC's Next Generation Project. We begin with a discussion of the combination of qualitative and experimental findings in relation to young children's well-being in the context of maternal welfare-to-work transitions, and then we turn our attention to the same issue with respect to the well-being of adolescents.

<u>Young Children's Well-Being</u>

Morris, Scott & London (2005) attempted to further our understanding of how welfare and work policies affect young children by comparing, contrasting, and synthesizing the results from two separate Next Generation reports. They sought to interrogate and integrate the conclusions drawn from these studies in order to better understand the mechanisms through which welfare reform policies targeted at adults, most of whom are mothers, impact young children's lives. One of the initial studies synthesized findings from eight experiments (Morris et al. 2001) and the other examined longitudinal, qualitative interview data from Cleveland and Philadelphia (Scott, Edin, London & Kissane 2004).

Findings from the experimental studies revealed that impacts on children's well-being differed depending on the policy approach and subsequent changes in parents' economic outcomes. Programs that increased parents' employment and earnings benefited young children; they exhibited better school achievement, reduced behavior problems, increased positive social behavior, and improved health. By contrast, programs that simply moved mothers into employment (without increasing earnings) did little to benefit young children. Because experimental impacts are derived from the *difference* in functioning between children and parents in the program and control groups, these findings may reflect the fact that receiving welfare (as more parents in the control group are doing) and working (as more parents in the program group are doing) are equally stressful for parents. The advantages that employment may bring to parents may be outweighed by the stresses of balancing work and family.

The majority of respondents in the qualitative interview sample, which included women from some of the most disadvantaged neighborhoods in Cleveland and

Philadelphia, experienced neither stable work nor sustained increases in income in the first year of the study. They struggled to make ends meet with income levels hovering around the federal poverty level. Juggling stressful jobs, irregular or part-time hours, and often temporary employment, mothers worried a great deal about their children's well-being.[5] Exhaustion, disrupted routines, and increased absence from the home due to employment meant that some mothers reported their children were doing worse, exhibiting diminished school performance and increased problem behavior.

Consistent with the experimental results, the qualitative interviews also suggested that when women moved from welfare to work *and* their incomes increased, they felt their children and families benefited. In addition to increased employment and income, two other factors also appeared to be critical to these families' success: (1) they had stable employment, and (2) they felt their child care arrangements were reliable and trustworthy, which most of the women defined as kin care. Because these two factors tended to coincide with work-related increases in income among the women in our sample, it was difficult to disentangle the relative contributions of maternal employment, the income increase that accompanied the transition to employment, and employment and child care stability to the improvements these women said were occurring in their families' and their children's lives.

As a consequence of their examination of the intersections of the experimental and qualitative results, Morris, Scott & London (2005) argued that these two lines of research helped them develop a more-nuanced understanding about how families and young children fare as single parents make the transition to employment and the conditions that aid families in coping with this transition. The qualitative data clearly

suggest that when considering employment, income, and family well-being, researchers must consider a host of interacting and possibly countervailing factors. Income increases must be substantial and not offset by work expenditures, employment and income must be stable, the psychosocial and financial benefits of maternal employment for children must not be outweighed by maternal exhaustion and absence, and children must be in reliable care that is trusted by their mothers. Proponents of changes in welfare policy had suggested that one way maternal employment may benefit children is that it would increase the regularity of family routines. These findings suggest that employment itself does not necessarily lead to such regularity. But employment increases that are accompanied by increases in income, perhaps because they are accompanied by increases in stable employment and child care, are more likely to result in such benefits for family life, and, in turn, children's development.

The qualitative findings also point to the role of employment stability as an important component of the benefits of increased employment and income for young children. Interestingly, employment stability may be responsible for some of the benefits of the experimental programs that increase both employment and income. Experimental findings suggest that earnings supplementation programs increase employment stability as they increase the proportion of families working. While some of these programs increased the number of jobs or spells of employment experienced by program group members (indicating a greater number of employment transitions), they all increased the length of employment spells. It appears that at least some of the increase in employment generated by these programs was in stable employment. As indicated by the qualitative

research, this may have added some stability to these families' lives and led to some of the benefits for children.

The qualitative findings also suggest that stable and reliable child care is a key intervening mechanism in creating the benefits of increased employment and income (see also Scott, London & Hurst 2005). Interestingly, the quantitative studies also find evidence that programs that increase employment and income may increase formal child care and participation in after-school programs. More specifically, in three of the four earnings supplement programs, mothers in the program groups were more likely to enroll their children in formal child-care programs or after-school programs and extracurricular activities than were mothers in the control group. These findings stand in contrast to those regarding impacts of these earnings supplementation programs on measures of parenting behavior. These outcomes were not much affected by these programs.

Together, these findings point to the benefits to children when welfare-reliant parents increase both their employment *and* their income. While income in these programs increased as a result of income supplements, other mechanisms, such as increases in minimum wages and providing supports for education might also have effects on income. Longitudinal, qualitative interview findings suggest that increases in employment and income may benefit families in part because they are accompanied by stability of employment and child care arrangements. By contrast, few, and perhaps countervailing, effects were observed when parents increased their employment but not their income, or when income increases were not enough to move them substantially above the poverty threshold.

Adolescent Well-Being

A second example of the benefits derived from having complementary qualitative and experimental evidence emerged in the Next Generation team's work on adolescent well-being (Gennetian et al. 2002, 2004). Members of the research team worked independently on analyses of the quantitative and qualitative data, but regularly held conference calls to discuss emerging findings. Case studies, developed from the longitudinal, qualitative interview data from Cleveland and Philadelphia, were sent as they were completed to the other team members who were working on the experimental analyses.

Consistent with the emerging experimental findings, the case studies revealed instances of adolescents who were faring poorly as their mothers went to work. The case studies suggested that one mechanism by which adolescent school performance and other aspects of well-being might be compromised was that adolescents were called upon to provide care to younger siblings. For example, Tina, a 35-year-old African American mother of six living in Philadelphia, described the negative impact of her welfare-to-work transition on her oldest child, Tamara. Tina said that because of her work schedule, Tamara had to be responsible for waiting in the morning with the younger children until the van that took them to their day care center arrived. The van typically ran late, which made Tamara 20 to 30 minutes late for school on a regular basis. Tina said:

> She's late every day for school, every day. And what the school says to me is…they gotta do what they, what's their policy. She's gotta stay after school, do her detention…or she'll lose her credit out of her, out of that morning class 'cause she didn't get there on time. So, she feels sad and I feel bad because I gotta be at

work at 7:00. She can't be at school by 7:00 – she can't. We all can't be at the same place at the same time.

Although Tina felt tremendous guilt for imposing this child care burden on her oldest daughter, she felt she had no choice. Tamara's timely school attendance meant less for the family's immediate well-being than Tina's timely attendance at work. So, it was Tina's adolescent daughter who bore the brunt of this burden in the short term and, potentially, in the longer term.

Other women told similar stories, which led the group that was engaged in analyzing the experimental data to hypothesize that the negative effects they were observing might be more pronounced in those families where there were younger children who might need care provided by an adolescent sibling. Quantitative results indicated support for this hypothesis. The experimental analyses indicate that programs that increased employment and had negative effects on adolescents' school outcomes were the same programs that increased adolescents' home responsibilities.

Most of the experimental studies did not collect detailed information about older children's responsibilities at home. However, three studies did collect information about child care responsibilities for the subset of families that had elementary school-aged children. In these studies, adolescents were taking care of their siblings in 14 to 28 percent of the families, with the adolescent being the primary source of care in 7 to 19 percent of cases. The majority of adolescents who provided sibling care were female (50 to 80 percent) and ranged in age from 14 to 18 years. Adolescents provided care to younger siblings up to about 10 hours per week. In the two programs that increased maternal employment and participation in employment-related activities and negatively

affected adolescents' school outcomes, adolescents were more likely to be caring for elementary school aged siblings, suggesting that they did take on some additional responsibility at home. Adolescents in the third program, which did not adversely affect adolescent school outcomes, were no more likely than their control group counterparts to care for younger siblings. Taken together, the complementary evidence from the experiments and the qualitative interview study provides a compelling case for the role of increased home responsibilities and child care as an important mediator in the relationship between adult welfare and work policies and adolescent school outcomes.

Discussion

As noted at the outset of this paper, there appears to be increasing interest in mixed-methods, qualitative-quantitative research among poverty researchers in various contexts (Gibson & Duncan 2005; Kanbur 2003; see also the other papers in this special issue). There are many approaches to doing this work and substantial accumulated wisdom. Numerous recent analyses address issues related to mixed-methods research (e.g., Jessor, Colby & Shweder 1996), including the recently-published *Handbook of Mixed Methods in Social & Behavioral Research* (Tashakkori & Teddle 2003). While a full discussion of approaches to mixed-methods research is beyond the scope of the current analysis (e.g., we do not discuss the sequencing of methods, which is discussed in other papers in the special issue and in the *Handbook* referenced above), we believe our accumulating experience suggests several important lessons about best practices and directions for the future.

Consistent with the examples and arguments put forward by Gibson & Duncan (2005), we believe it is important that there be tighter integration of quantitative and

qualitative components of studies than we have been able to achieve in the Urban Change and Next Generation projects. While our analyses have yielded valuable insights, as discussed above, the fact that there is no formal linkage between the quantitative and qualitative samples in any of the studies constrains full integration. We think there are a number of points where linkage can (and arguably should) be made that will enhance the ability of analysts to integrate the data and findings.

First, there should be greater emphasis on designing and funding studies as mixed-methods studies so that research questions and instruments can be coordinated from the outset and fielded in tandem. As Gibson & Ducan (2005) note, funding constraints meant that the qualitative component of the New Hope evaluation was not fielded at the beginning of the experiment, which compromised to some extent the integration and analyses they were able to conduct.

Second, we believe that, whenever feasible, the samples should be embedded and formally linked to allow for better integration of results. This applies both to non-experimental survey research and to experiments. In combination with fielding mixed-methods studies from the outset, such a linkage has benefits for analysis and ensures that the time frame and location for both types of data are the same. The two examples described above, which combine experimental and qualitative data, combine data from different settings and time periods (i.e., different policy contexts), which necessarily weakens the formality of the integration. The Urban Change survey and qualitative interviews were collected in the same locations and periods, from women drawn from the same population, but the samples were not linked, which again compromises the extent of the integration.

Third, we agree with Gibson and Duncan (2005) that having the same analysts working on the analysis of the quantitative and qualitative data is beneficial and should be pursued whenever time, funds, expertise, and goodwill for data sharing exist. Like the analysts on the New Hope evaluation, Polit, London & Martinez (2000, 2001) conducted both quantitative and qualitative analyses themselves and found the cross-fertilization across data sources generated new ideas and insights that might not have otherwise arisen.

A fourth suggestion does not directly emerge from our experiences, and is perhaps the most novel and potentially controversial of our suggestions. However, we believe it follows logically from the considerations outlined above. We believe that there is utility in exploring and developing methods for conducting mixed-methods interviews, mixing closed-ended surveys with in-depth interviews. Although there are various issues that need to be worked out and tested, we believe the integration of funding, design, data collection, sampling, and analytic personnel holds the most promise for maximizing the contributions of mixed-methods research in the context of policy evaluation.

Overall, our experience suggests that mixed-methods research and the integration of quantitative and qualitative research methodologies and data holds great promise for enhancing the evaluation of policies and programs. We believe that having high quality qualitative data is important even when high quality quantitative data are collected, as in the case of the Urban Change survey and the experimental evaluations. We also believe it is important for quantitative researchers to pay attention to variability as well as central tendencies in their data and impacts in their surveys, and to pay more attention than they do to the nature of the scales that they use. The qualitative data remind us that there are

different pathways to similar ends, as well as difference in experiences and appraisals of seemingly similar circumstances. Additionally, the qualitative data in their own right, but also in relation to the quantitative data, encourage researchers to consider non-monetary aspects of economic well-being, poverty, poverty appraisal, and policy evaluation more systematically.

Although our paper argues for the benefits of mixed-methods research, we close with a caution. We must be careful not to raise expectations too high. Doing mixed-methods research – designing the study, raising the money for it, collecting the data, and analyzing it – are difficult and require different kinds of initiatives than those that are currently pursued, funded, and staffed. Training personnel will require effort, and data collection and processing is differentially expensive and time-consuming. There may be questions that one method can answer well that the other cannot, and if resources are prioritized toward one set of questions (and method) over the other, it is not clear that mixed-methods research would be the best approach in that case. Moreover, inductive analyses of qualitative data have the potential to yield valuable and unanticipated insights. However, the time involved for such analyses may be at odds with both the deductive, a priori framing of research questions and the time frames for producing results that can drive much policy and program analysis. Recognizing the strengths and limitations of these methodologies, and making sure that they are not overly compromised in our efforts to integrate them will be important to prevent overselling the promise or underselling the pitfalls associated with undertaking mixed-methods research.

References

Anderson, S. A. (Ed.) (1990). Core indicators of nutritional status for difficult-to-sample populations. *Journal of Nutrition*, 120, 1559-1600.

Brock, T., Coulton, C., London, A., Polit, D., Richburg-Hayes, L., Scott, E., & Verma, N. (2002). *Welfare reform in Cleveland: Implementation, effects, and experiences of poor families and neighborhoods*. New York: MDRC.

Carlson, S. J., Andrews, M. S., & Bickel, G. W. (1999). Measuring food insecurity and hunger in the United States: Development of a national benchmark and prevalence estimates. *Journal of Nutrition*, 129, 510S-516S.

Frongillo, E. A. (1999). Validation of measures of food insecurity and hunger. *Journal of Nutrition*, 129, 506S-509S.

Gennetian, L. A., Duncan, G. J., Knox, V. W., Vargas, W. G., Clark-Kauffman, E., & London, A. S. (2002). *How welfare and work policies for parents affect adolescents: A synthesis of the research*. New York: MDRC.

Gennetian, L. A., Duncan, G. J., Knox, V. W., Vargas, W. G., Clark-Kauffman, E., & London, A. S. (2004). How welfare policies can affect adolescents. *Journal of Research on Adolescents* 14: 399-423.

Gibson, C., & Duncan, G. J. (2005). Qualitative/quantitative synergies in a random-assignment program evaluation. Pp. 283-315 in Thomas Weisner (Ed.), *Discovering successful pathways in children's development: New methods in the study of childhood and family life*. Chicago: University of Chicago Press.

Jessor, R., Colby, A., & Shweder, R. A. (Eds.) (1996). Ethnography and human development: Context and meaning in social inquiry. Chicago: University of Chicago Press.

Kanbur, R. (2003). *Q-Squared: Qualitative and quantitative methods of poverty appraisal*. New Delhi: Permanent Black.

London, A. S., Scott, E. K., Edin, K., & Hunter, V. (2004). Welfare reform, work-family tradeoffs, and child well-being. *Family Relations*, 53(2), 148-158.

Morris, P. A., Huston, A. C., Duncan, G. J., Crosby, D. A., & Bos, J. M. (2001). *How welfare and work policies affect children: A synthesis of research*. New York: MDRC.

Morris, P. A., Scott, E. K., & London, A. S. (2005). Effects on children of parents' transitions from welfare to employment: Integrating quantitative and qualitative research. Pp. 87-116 in Jill Duerr Berrick & Bruce Fuller (Eds.), *Good parents or good workers? How policy shapes families' daily lives*. New York: Palgrave Macmillon.

Polit, D. F., London, A. S., & Martinez, J. M. (2000). *Food security and hunger in poor, mother-headed families in four U.S. cities*. New York: MDRC.

Polit, D. F., London, A. S., & Martinez, J. M. (2001). *The health of poor urban women: Findings from the project on devolution and urban change*. New York: MDRC.

Quint, J., Edin, K., Buck, M. L., Fink, B., Padilla, Y. C., Simmons-Hewitt, O., & Valmont, M. E. (1999). *Big cities and welfare reform: Early implementation and ethnographic findings from the Project on Devolution and Urban Change*. New York: MDRC.

Raphael, J, & Tolman, R. M . (1997). *Trapped by poverty/Trapped by abuse*. Ann Arbor: University of Michigan, Research Development Center on Poverty, Risk, and Mental Health.

Scott, E. K., Edin, K., London, A. S., & Mazelis, J. M. (2001). My children come first: Welfare-reliant women's post-TANF views of work-family trade-offs and marriage. Pp. 132-153 (Chapter 8) in G. J. Duncan and P. L. Chase-Lansdale (Eds.), *For better and for worse: Welfare reform and the well-being of children and families*. New York: The Russell Sage Foundation.

Scott, E. K., Edin, K., London, A. S., & Kissane, R. J. (2004). Unstable work, unstable income: Implications for family well-being in the era of time-limited welfare. *Journal of Poverty* 8(1): 61-88.

Scott, E. K., London, A. S., & Hurst, A. (2005). Out of their hands: Patching together care for children when parents move from welfare to work." *Journal of Marriage and the Family* 67: 369-385

Scott, E. K., London, A. S., & Myers, N. A. (2002a). Living with violence: Women's reliance on abusive men in their transitions from welfare to work. Pp. 302-316 (Chapter 14) in N. Gerstel, D. Clawson, and R. Zussman (Eds.), *Families at work: Expanding the bounds*. Nashville: Vanderbilt University Press.

Scott, E. K., London, A. S., & Myers, N. A. (2002b). Dangerous dependencies: Domestic violence in the context of welfare reform." *Gender & Society* 16 (6): 878-897.

Tashakkori, A., & Teddle, C. (eds.) (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks: Sage Publications.

Tolman, R. M, & Raphael, J. (2000). A review of research on welfare and domestic violence. *Journal of Social Issues,* 56(4), 655-681.

Endnotes

[1] See Quint, Edin, Buck, Fink, Padilla, Simmons-Hewitt et al. (1999) for additional details about the Urban Change study.

[2] One of the experiments involved a Canadian welfare reform demonstration program and therefore had a somewhat different institutional structure. Additional details about these experimental evaluations are available in Morris, Huston, Duncan, Crosby & Bos (2001) and Gennetian, Duncan, Knox, Vargas, Clark-Kauffman & London (2002, 2004).

[3] The 18 items in the HFSS ask respondents to indicate whether, as a result of inadequate financial resources, they experienced any of the following in the prior 12 months (Carlson, Andrews & Bickel 1999: 513S). The items are arrayed in order of severity: worried food would run out; food bought didn't last; unable to afford balanced meals; child(ren) fed few, low-cost foods; adult(s) cut size or skipped meals; couldn't feed child(ren) balanced meals; respondent ate less than felt s/he should; adult(s) cut, skipped meals in 3+ months; child(ren) not eating enough; respondent hungry but didn't eat; respondent lost weight; cut size of child(ren)'s meals; adult(s) not eat for whole day; child(ren) hungry; adult(s) not eat for whole day in 3+ months; child(ren) skipped meal; child(ren) skipped meal in 3+ months; child(ren) not eat for whole day.

[4] All names used in this paper are pseudonyms.

[5] See also Scott, Edin, London & Mazelis (2001) and London, Scott, Edin & Hunter (2004).