

The FGA’s “First of Its Kind Study” Should Have Been the Last: An Evaluation Note for Pre-Post Conservatives

Peter (the Citizen for Evaluation Accountability) Germanis¹

July 4, 2017

Writing in *Forbes*, Nic Horton, Jonathan Ingram, and Josh Archambault of the Foundation for Government Accountability (FGA) touted the findings of a recent “study” – one they characterize as “the first of its kind” – that examined the effects of a “work requirement” in Kansas for able-bodied adults without dependents receiving benefits from the Supplemental Nutrition Assistance Program (SNAP).² Specifically, SNAP limits able-bodied adults aged 18 to 49 without children to three months of benefits in any 36-month period, unless they are employed or in a work or training program for at least 20 hours a week. States do not have to offer these individuals a work or training program slot, and most do not, so this provision is more appropriately seen as a time limit rather than a “work requirement.”³

The authors of the FGA report summarize their results as follows:

New research from the Foundation for Government Accountability examines the results of Kansas’ welfare reforms and the findings are simply staggering. Work requirements have led to more employment, higher incomes, and less poverty. And instead of drawing millions of dollars in taxpayer-funded welfare benefits, these able-bodied adults are now contributing tens of millions of dollars to the local economy.

Some of the specific “findings” reported in the study include:

- “Nearly 60 percent found employment within a year of leaving food stamps.”
- “After leaving food stamps, incomes increased 127 percent.”
- “Within a year of leaving food stamps, the number of able-bodied adults living in poverty dropped significantly and roughly half of those working climbed out of poverty entirely.

What made this study the “first of its kind”? Apparently, the authors believe that tracking the employment and earnings outcomes of those affected by the policy is something new. They explain:

The Brownback administration also set in motion a first-of-its-kind tracking system to monitor the state’s success in moving Kansans from welfare to work. Over the course of two years, the state’s Department of Labor and its Department for Children and Families tracked earnings and employment for nearly 41,000 able-bodied adults leaving food stamps after the work requirements went into effect.

There are hundreds of evaluations that track the earnings and employment of those subject to various welfare reforms – what makes the FGA’s study somewhat unique (though unfortunately not the first of its kind)? The FGA has no comparison or control group that could serve as a counterfactual! As a result, its findings are not credible – at all! The FGA confuses data collection with evaluation. The authors of the study simply assume that any change in the

outcomes they observed were the direct result of the “work requirement.” Even worse, they make unfounded claims that may mislead and misguide policymakers.

This paper is divided into four parts. The first section describes some of the basic threats to the validity of a simple pre-post study and explains why the FGA methodology is likely to dramatically overstate the earnings gains of those subject to the SNAP “work requirement.” The second section provides an example that shows the dangers of the FGA approach – a case in which proponents of a program boasted great success on the basis of a pre-post study only to find that there were no effects when compared to a true counterfactual. The third section describes a number of other shortcomings of the report beyond having a weak counterfactual. The fourth section describes a number of statements made throughout the FGA report that are misleading or speculative. If the authors of the FGA report (and other “pre-post conservatives”) want to be taken seriously, they need to address the methodological weaknesses of their studies (or the studies they cite as examples of success) and be careful in how they use and interpret data.

I. Threats to Causal Validity

There are a wide range of evaluation methods that can be used to assess the effects of a program or policy. Some examine changes in a program group over time (e.g., pre-post and interrupted time series), some compare a program group to a “comparison group” of similar groups or individuals (with various statistical adjustments to try to make the comparison group as comparable as possible to the program group), and some compare a program group to a randomly assigned “control group.” The FGA study is most like a pre-post study, as the authors make claims of program effects based on a comparison of a baseline quarter (pre) to a subsequent quarter (post).

Every evaluation approach is subject to various threats to internal and external validity, but simple pre-post studies are particularly problematic for assessing the impact of employment-related interventions. Internal validity refers to the extent to which a causal conclusion can be made based on a study, ruling out other possible explanatory factors. External validity refers to the degree to which the findings of a study can be generalized to other populations, settings, times, and programs. The FGA study is weak on both, though the focus of this critique is on its weak internal validity.

Pre-post studies are subject to a number of threats to internal validity, including maturation (i.e., where changes in the outcomes of individuals are caused by the normal passage of time), history (i.e., where events unrelated to an intervention affect the outcome of interest), and secular trends (i.e., where outcomes are the result of a societal trends, such as changing economic conditions). Of the three, the threat from secular trends is probably the most important confounding factor, as the economy in Kansas improved over the observation period. For example, the state’s unemployment rate dropped from 5.0 percent in October 2013 to 4.3 percent in June 2015.

The most serious threat to FGA study, however, may be “regression to the mean,” in that the initial observation in a pre-post study may be at an extreme point relative to an individual’s or group’s normal trajectory. In the economics literature, this is a problem known as “Ashenfelter’s

Dip,”⁴ where the average earnings of program participants in a training or work-related program just prior to program entry is at a low point. Howard Bloom of MDRC explains:

... previous studies of employment and training programs have identified a “pre-program dip” whereby average earnings in the year before participants enter a program are well below their preceding trend. This phenomenon has been the subject of extensive debate, and there is still no agreement about the extent to which it represents a temporary aberration (due to illness, a business failure, or other “bad luck”), or the extent to which it represents the onset of a permanent decline in earnings (which happens to economically displaced workers).

Individuals are probably more likely to enroll in a job-training program when they have just experienced an unusually bad period than when they have just experienced an unusually good period. In other words, their motivation to enroll is probably negatively correlated with random fluctuations in their earnings prospects. If so, then their average earnings in the next period are likely to increase regardless of whether or not they enroll in a program. This phenomenon is a statistical artifact called regression to the mean.⁵

This same concern would be true for able-bodied adults without dependents who enroll in SNAP – many are likely to have just lost a job or had a temporary set-back and their use of SNAP is relatively temporary. Thus, even in the absence of a “work requirement,” they would begin to work anyway. (A more detailed discussion of the dynamics of SNAP receipt is provided in section IV below.)

The best way to rule out this threat to validity is a randomized control trial, whereby able-bodied adults without dependents are randomly assigned to the “work requirement” group or to a control group that is not subject to the requirement. The findings from random assignment experiments are considered the most credible, because the program and control groups are alike and subject to the same external conditions, with the only difference being the intervention itself. Thus, any difference in outcomes between the groups can be attributed to the intervention – the SNAP “work requirement.” In particular, the threat from “regression to the mean” would be overcome because both program and control groups would be alike in terms of their employment and earnings histories.

The frequency of work among this population has been noted by Dorothy Rosenbaum and Ed Bolen of the Center for Budget and Policy Priorities (CBPP) in their critique of the FGA’s reports on this topic. They note:

At least a quarter of households with childless adults work while receiving SNAP, and about 75 percent work in the year before or the year after receiving SNAP. Limited education, lack of training, and a sporadic work history make it difficult to compete for anything other than low-skill, low-wage jobs that do not lift them out of poverty.⁶

They conclude:

The reports claiming that terminating SNAP after three months for this population in Kansas and Maine resulted in increased work rates misleadingly present only what happened after the time limit took effect. The reports fail to consider that many SNAP recipients already worked, or would work soon even without the time limit. Looking at the same population before the time limit took effect reveals similar work rates. While challenges exist in getting unemployed childless adults back into the workforce, there's no evidence that the time limit itself has a significant effect on work.⁷

This conclusion is absolutely right! This does not mean that work requirements for SNAP able-bodied adults without dependents don't induce some employment and earnings gains, but it is implausible that taking away \$150 to \$180 a month in SNAP benefits would induce the kinds of employment and earnings gains suggested by the FGA, particularly if the "work requirement" provided little actual help in finding a job.⁸ (As discussed below, a major shortcoming of the study is the failure to examine any aspect of the implementation of the "work requirement.")

In short, without a credible counterfactual, this report cannot be taken seriously and it, along with most other pre-post studies, should not be considered in a serious policy debate.

II. The Importance of a Counterfactual

The FGA authors claim to have a "first of its kind study." A simple pre-post comparison of outcomes is easy to do and sadly has been all too common. For example, Jonathan Freedman, a Pulitzer prize-winning journalist, wrote favorably about one early intervention program in the 1990s:

The Comprehensive Child Development Program (CCDP) has an unwieldy name but a single purpose – to end welfare dependency by promoting family self-sufficiency. It really works. When one center in Battleboro, Vermont, in 1990, participating families' average earned income was \$9,358. Four years later, average earnings more than double to \$19,643. The Vermont program slashed welfare dependency in half: 68.5 percent of the families subsisted on welfare in 1990; only 37 percent in 1994.⁹

Like the FGA, these claims are based on a simple pre-post comparison. And, like the FGA, Freedman advocated for extending the model across the nation. Fortunately, the U.S. Department of Health and Human Services contracted with Abt Associates to conduct a rigorous, random assignment evaluation of the program in 21 projects across the country.¹⁰

The Abt researchers did just what the FGA did – they tracked the employment and earnings (and many other outcomes) of those in the CCDP group. Within two years, the percent of mothers employed in a quarter nearly doubled – from about 18 percent to about 33 percent. So, was Jonathan Freedman right? Did this program really reduce dependency and promote self-sufficiency? Sadly, the answer is no. The reason we know this is because the researchers also had a control group that formed the basis for a comparison. The Abt report explains:

...mothers in the control group found employment and earned more money, the percentage of control group families receiving AFDC and Food Stamps decreased... This

pattern of findings tells us that ... control group families cannot be assumed to be static or unchanging.¹¹

So, what is the lesson?

These findings point out the need for a randomly assigned control group. Data collected only on CCDP families would have given the misleading impression that the observed improvements in the lives of low-income families were attributable to participation in the program. When we see that the same types of improvements happen for control group families, we realize that we are observing normal changes in the lives of families – changes that cannot be attributed to CCDP.¹²

The FGA’s simplistic evaluation approach could easily be used to support a wide range of policies because outcomes change over time for reasons other than an intervention – hence, the need for a credible counterfactual. For example, suppose there were a program that targeted unemployed SNAP families with no income, aside from their SNAP benefits, and provided them \$500 a month in cash. If these families were tracked for a year and 50 percent became employed and had earnings, would this be result of the cash payment? Obviously, not – but this is the same flawed approach used by the FGA in assessing the impact of the SNAP “work requirement.”

III. Methodological Limitations and Misleading Data Comparisons

Aside from lacking a credible counterfactual, the FGA report presents data in an incomplete and unconventional manner – many of their findings are either misleading or meaningless.

Misleading employment claims. The FGA report asserts, “The data shows that the less time these able-bodied adults spend on welfare, the quicker they can get back into the workforce once they are freed from welfare and the more money they will make. These Kansans are discovering new lives of independence and self-sufficiency that, in some cases, they haven’t known for more than two decades.” To support this claim, they add, “Nearly 60 percent found employment within a year of leaving food stamps.”

The employment rate is an important outcome in any welfare reform-related evaluation. Most random assignment evaluations show the employment rate for a baseline quarter (i.e., the quarter of random assignment), and then each quarter thereafter. This shows the status of program participants at a particular point in time and relative to a control group so as to permit accurate estimates of employment impacts. The authors of the FGA report did not do this. Instead, they simply reported the number “ever-employed” through a particular quarter – not the employment rate for a particular quarter.

The first three columns of Table 1 below come from Table 7 of the FGA report. The first column shows the time period and the second column shows the number of able-bodied adults with dependents exiting food stamps in December 2013 (i.e., receiving benefits in that month, but not the next), who have a “record of employment since Q4 2013” (i.e., the baseline quarter). The third column reflects a *cumulative* employment rate, that is, the percent of leavers ever-employed at any time since the fourth quarter of 2013. (The term “cumulative” is added because

an employment rate presented this way does not show the percent of leavers employed in any one quarter, but rather the percent employed at any time since a baseline period – the fourth quarter of 2013). The *cumulative* employment rate in the fourth quarter of 2014 is 58.9 percent. This forms the basis of the claim that “nearly 60 percent found employment within a year of leaving food stamps.”

The FGA’s characterization of the employment rate leaves the impression that the SNAP “work requirement” caused the increase, because the report provides no reference point. The authors don’t include the baseline employment rate, which would have been the fourth quarter of 2013. For a pre-post study, this is extremely sloppy and misleading. The fact that nearly 60 percent of leavers “found” employment would be impressive if none were employed during the baseline quarter (i.e., the fourth quarter of 2013). In addition, instead of reporting quarterly employment rates, the authors use an employment rate based on any “record of employment since Q4 2013.” Comparing a quarterly employment rate for the baseline (or “pre”) period to an “ever-employed” over the course of a year employment rate (or “post”) period would be comparing apples to oranges. Even with a simple pre-post study, what is needed is a comparison of the quarterly employment rate in the baseline quarter to a quarterly employment rate to a subsequent quarter.

Fortunately, a quarterly employment rate can be calculated from the FGA report, because it provides data on the average quarterly wages of “all disenrollees” and as well as those who are “current workers.” The quarterly employment rate can be computed by dividing the second figure by the first figure and is shown in the fourth column of Table 1 below. This calculation shows that 38.9 percent of able-bodied adults without dependents who exited SNAP in December 2013 were employed during the baseline quarter.

Table 1: Comparing Employment Rates: The FGA’s Big Exaggeration			
	Record of employment since Q4 2013	Cumulative Employment Rate	Quarterly Employment Rate (NOT reported by FGA)
Q3 2013			36.4%
Q4 2013	?	?	38.9%
Q1 2014	4,920	38.8%	38.4%
Q2 2014	6,217	48.5%	42.1%
Q3 2014	7,012	54.8%	41.8%
Q4 2014	7,537	58.9%	42.1%
Q1 2015	7,870	61.5%	
Q2 2015	8,252	64.4%	

Source: Table 7 in Jonathan Ingram and Nic Horton, *The Power of Work: How Kansas’ Welfare Reform Is Lifting Americans Out of Poverty* (The Foundation for Government Accountability, February 16, 2016), available [here](#). The last column is derived from Table 8 of the FGA report.

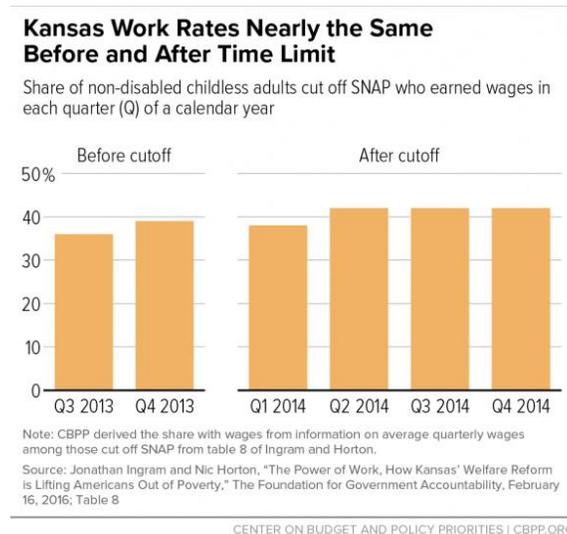
When quarter-by-quarter employment rates are examined, it is obvious that the FGA’s claims are exaggerated. By the fourth quarter, only 42.1 percent of those who exited SNAP in December 2013 were employed. So, even if one believes a simple pre-post study is appropriate, the proper comparison would be 38.9 percent at baseline to 42.1 percent at the time of follow-up a year later – an impact of 3.2 percentage points (8 percent). Does this really justify a claim like, “These

Kansans are discovering new lives of independence and self-sufficiency that, in some cases, they haven't known for more than two decades”?

Dorothy Rosenbaum and Ed Bolen of the CBPP first noted this problem in the FGA “analysis” in December 2016, stating:

... work rates before and after the time limit were very similar... Almost 40 percent of those whose SNAP was cut off already worked in each of the last two quarters before the time limit returned (the third and fourth quarters of 2013). The share working each quarter in the year after the time limit was implemented rose slightly, to just over 40 percent. This modest increase could be explained by two factors: (1) low-wage workers are more likely to apply for and participate in SNAP when they lose a job or their incomes drop, so they often experience improvements in the future as their employment situation improves; and (2) Kansas' economy was improving over 2014, so a slightly larger share of recipients may have been able to find jobs or higher pay. The time limit does not appear to have dramatically affected work rates for the group subject to it.

The figure below from the CBPP critique illustrates this point.



The authors of the FGA report had this data and they should have presented it. Their depiction of an increasing “ever-employed” rate is highly misleading and is not evidence of the impact of the SNAP “work requirement.”

Misleading income and poverty claims. The FGA report also has bold claims about income gains and reductions in poverty. For example, it states, “After leaving food stamps, incomes increased 127 percent.” To support this claim, the report has a figure that shows the “Average annualized wages of able-bodied adults without dependents exiting December 2013, by Quarter.” The figure is based on the following data:

Table 2: Average Annualized Wages of Able-bodied Adults Without Dependents	
	Average annualized wages
FGA's Baseline (Q3 of 2013)	\$2,453
Real Baseline (Q4 of 2013)	\$3,584
Q1 2014	\$3,856
Q2 2014	\$4,837
Q3 2014	\$5,196
Q4 2014	\$5,562

Source: Derived from Table 8 of Jonathan Ingram and Nic Horton, *The Power of Work: How Kansas' Welfare Reform Is Lifting Americans Out of Poverty* (The Foundation for Government Accountability, February 16, 2016), available [here](#).

These data come from Table 8 of the FGA report, where the authors simply multiply the “Average quarterly wages of all disenrollees” by four.¹³ The claim of an increase in “income” of 127 percent is based on the comparison of the FGA’s reported baseline figure of \$2,453 to the fourth quarter of 2014 figure of \$5,562. Even if a pre-post approach were indicative of causal effects (and as noted above, it is not a credible approach for this type of intervention), the authors arguably used the wrong baseline quarter. The \$2,453 figure is based on the third quarter of 2013 – reflecting a period 4 to 6 months before the cut-off. The more appropriate baseline quarter is actually the fourth quarter and if those earnings had been used the baseline would have been \$3,584. Then, the increase is just a 44 percent. It is also noteworthy that this group experienced a significant increase in earnings between the third and fourth quarters of 2013 – from \$2,453 to \$3,856, or a 57 percent increase.

It is possible that the FGA authors chose the third quarter of 2013 as the baseline, because the three-month time limit started in October 2013. So, one reason for the gain in “average annualized wages” is the anticipatory effect of the cut-off. However, if this was their intention, then they should have collected data on those able-bodied adults without dependents receiving SNAP benefits in October 2013 – instead, the focus of their report is on those exiting December 2013. This suggests that the more appropriate baseline quarter is the fourth quarter of 2013. Moreover, in a figure showing annualized incomes over time (p. 16 of the FGA report), the authors choose the fourth quarter of 2013 as the baseline quarter for purposes of their “average annual benefits” (see Table 10b of the FGA report).¹⁴

The FGA suggests that it is comparing “incomes,” but its figure is limited to earnings. The report includes a figure that compares the baseline “income” reflecting annualized earnings plus an annualized SNAP benefit to annualized earnings in the third quarter of 2014 (\$5,562). This figure shows the FGA’s baseline earnings as \$2,453 and the annualized SNAP benefit as \$2,139, for a total annualized income of \$4,592. This suggests a gain of \$970. Even if one believes this is a valid approach for determining causation, as noted above, the FGA used the wrong quarter for its “baseline” earnings. Using the fourth quarter of 2013 for earnings and the December 2013 annualized SNAP benefit would have raised the baseline “income” to \$5,723 – \$161 higher than the fourth quarter of 2014. (Using this approach, the income losses would have been even greater in earlier quarters of 2014.)

Even this approach understates the potential loss in income for many able-bodied adults without dependents because it is an average. As noted in Table 1 above, in any one quarter of 2014, about 60 percent of those who exited in December 2013 had no reported earnings. If they also lost their SNAP benefits, the income loss would have been much greater for them than the average loss. Yet the FGA report ignores this group altogether. Instead, the authors of the report make incredibly irresponsible claims like, “These reforms immediately freed nearly 13,000 Kansans from welfare on December 31, 2013.”

The failure to examine implementation. Any serious evaluation would have included a process study to assess how the state had implemented the “work requirement” – to examine the disability determination process, the activities that were offered, the reasons for participation or non-participation, and perhaps most important to follow-up with a sample of those who exited to find out how they were coping – particularly for the majority of exiters who do not seem to have employment in any given quarter.

Meaningless comparisons. A significant share of the FGA report compares SNAP “enrollees,” i.e., those receiving SNAP, before and after the “work requirement” is imposed. For example, the FGA report claims, “After work requirements, food stamp enrollees are twice as likely to work.” This implies that the work requirement *increased* the work effort of those affected by the policy. In fact, it is an artifact of the FGA’s flawed methodological approach. Before and after comparisons of employment outcomes, conditional on receiving SNAP benefits, are meaningless, because the “work requirement” changed the composition of those receiving benefits and thus the groups are not comparable.

If this were a random assignment evaluation, the basis of comparison would be the entire treatment group to the entire control group, regardless of whether individuals continued to receive SNAP. Limiting the comparisons to those receiving SNAP benefits would undermine the comparability of the groups, particularly when one group is much less likely to receive benefits than the other.

Rosenbaum and Bolen explain:

These claims, which the report makes across a range of measures, are misleading because the childless adults who remained as SNAP participants after the time limit went into effect were significantly different from those who participated before because of the policy change. The state cut off SNAP those participants who were not working at least 20 hours a week, so the work rates, average earnings, and other characteristics of those who remained SNAP participants after the return of the time limit were better, not because those individuals became better off, but because they were better off to begin with and were the only ones still eligible and participating in SNAP.

Those who may still participate in SNAP are more likely to have earnings and, as a result, lower SNAP benefits and appear better off on a range of other characteristics. In fact, the number of childless adult SNAP recipients working at least 20 hours a week, and thus the only non-exempt childless adult SNAP recipients eligible for the program, dropped modestly in the year after the time limit took effect.

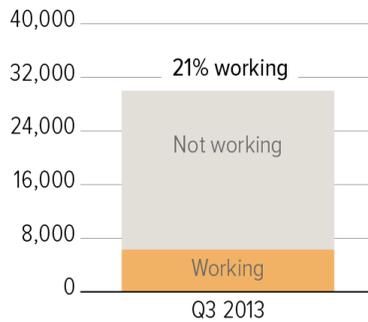
As an example of this misleading representation, consider the authors’ assertion that, “[p]rior to restoring work requirements, just 21 percent of childless adults on food stamps were working at all. Two-fifths were working less than 20 hours per week. But since work requirements have gone back into effect, that employment rate has risen to nearly 43 percent.” The change was driven by a drop in the number of SNAP recipients who are childless adults subject to the time limit, not an increase in the number of recipients who are working. The number of such SNAP recipients who were working fell by more than 40 percent (from 6,300 to 3,600), as those who were working less than 20 hours a week were cut off, while the total number of non-disabled childless adults receiving SNAP dropped by more than 70 percent (from almost 30,000 to 8,500).¹⁵

Another illustration from the CBPP’s report highlights this problem.

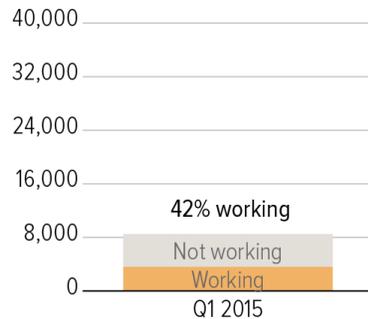
Kansas SNAP Benefit Cutoff Did Not Boost Work

Non-disabled childless adult SNAP participants before and after January 2014 cutoff of those not working 20+ hours per week

Number of such SNAP participants who worked fell...



...but the work rate rose only because there were fewer such SNAP participants overall



Source: Jonathan Ingram and Nic Horton, “The Power of Work, How Kansas’ Welfare Reform is Lifting Americans Out of Poverty,” The Foundation for Government Accountability, February 16, 2016; Tables 1 - 3

IV. Misleading Statements

The FGA's report is replete with misleading statements and unsupported conclusions. This response addresses just some of the claims made by the FGA, followed by a "PC Response." ("PC" is short for "Peter the Citizen.")

The FGA: "If that were not bad enough, food stamps have also become the gateway to long-term government dependency. According to the latest data, the typical enrollee now stays on food stamps for an average of more than eight years."

PC Response: This statement is highly misleading in the context of a SNAP report about able-bodied adults without dependents. The authors cite a study sponsored by the U.S. Department of Agriculture that examined the dynamics of SNAP participation from 2008 to 2012.¹⁶ This study does show that the average enrollee, *at a point in time*, received SNAP benefits for eight years. However, the FGA report leaves out several important facts.

First, the finding is based on *all* SNAP recipients, including the elderly, disabled, and others who may not be able to work. The result is not generalizable to able-bodied adults without dependents. That same study shows that nearly half of completed spells among nonelderly/nondisabled childless adults are two years or less – a far cry from the eight-year spell length cited by the authors of the FGA report.¹⁷ Second, the survey covered the 2008 to 2012 period – five to nine years ago – when the nation was in the midst of a severe economic downturn and jobs were not readily available.

Third, the findings in the USDA study examined "completed spell lengths" for "SNAP participants in spells underway in December 2008." In other words, the study identifies all those receiving SNAP in a given month and then reports the total length of the spell, including periods prior to the start of the survey up to the point they leave (or when the survey ends). This approach significantly exaggerates the duration of SNAP receipt of the "typical enrollee" over the course of a year. A classic study by Mary Jo Bane and David Ellwood illustrates this point, examining the dynamics of welfare receipt under the Aid to Families with Dependent Children (AFDC) program:

- For those going on AFDC, most spells are short-term, lasting two years or less, while fewer than one-sixth can be thought of as long-term, with recipients spending eight or more continuous years on the program.
- At any point in time, half of all AFDC recipients are in the midst of long-term spells.

These findings may seem paradoxical, but they explain, using an example of hospitalization spells to illustrate the point:

Consider the situation in a hypothetical hospital complex. Most of the persons admitted in any year will require only a very short spell of hospitalization. But a few of the newly admitted patients are chronically ill and will have extended stays in the hospital. If we

ask what proportion of all admissions are people who are chronically ill, the answer is relatively few. On the other hand, if we ask what fraction of the hospital's beds or equivalently what proportion of the patients in the hospital at any one time are chronically ill, the answer is much larger. The reason is simple. Although the chronically ill account for only a small fraction of all admissions, because they stay so long they end up being a sizable part of the population in the hospital and they consume a sizable chunk of the hospital's beds and other resources.¹⁸

This very same issue has to be considered in looking at SNAP dynamics and is clearly explained in the USDA report:

... in the cross-sectional sample, we miss many short spells that occur within the same panel period – they are likely to end before or to begin after our sample month. Longer spells, however, are more likely to include our sample month. For this reason, the longer spells are more heavily represented in the cross-sectional sample than in the entry cohort sample.¹⁹

This study also shows spell lengths for those who are “new entrants.” Nearly 80 percent of nonelderly/nondisabled childless adults who started a spell exited the program within two years; indeed, over a third exited within 4 months. The median spell length was 8 months.²⁰

Understanding these and other data issues is essential in specifying the size and scope of a policy problem. In this case the FGA authors present a grossly misleading exaggeration of the extent of long-term dependency of the “typical enrollee” as it pertains to a study about the SNAP “work requirement” for adults without dependent children.

The failure to understand this dynamic raises another limitation of the FGA study. Even if its findings were believable (and they are not), they would be limited to a cohort consisting mainly of longer-term recipients. This cohort misses many who use SNAP for relatively short periods of time during a year, but who nevertheless would be subject to the “work requirement.” Remember, the authors state that the state is tracking 41,000 able-bodied adults without dependents who left TANF – but most of the findings reported are for the 12,807 that left in December 2013. The results for those 30,000+ leavers not reported here may be very different. This limited generalizability is a problem of external validity.

The FGA: “This welfare explosion not only siphons resources away from other critical needs – it traps more Americans in poverty and robs them of the hope of a better life.”

PC Response: Are the FGA authors seriously suggesting that paying able-bodied adults without dependents a grant less than \$200 a month is so enticing that these individuals are “trapped” in poverty? If so, they should cite some empirical data instead of making unsupported and speculative claims. Many of these adults face barriers to employment, such as homelessness, physical and mental limitations, limited education, and criminal records – they are likely to poor regardless of whether they receive SNAP and irrespective of the SNAP “work requirement.” Indeed, even the FGA report indicates that 88 percent of able-bodied adults exiting SNAP in December 2013 were poor in the quarter of exit and 79 percent were poor four quarters later

(see Table 9 of the FGA report). This hardly justifies a title like the one the FGA used for its report, *The Power of Work: How Kansas' Welfare Reform Is Lifting Americans Out of Poverty*. It should be clear, even for those who believe in simplistic pre-post comparisons, that the SNAP work requirement is not the magical solution suggested by the FGA. Addressing the nation's poverty problems requires a testing a wide range of policy options and subjecting them to rigorous evaluation.

The FGA: “State leaders know that work changes lives. The state-led, work-first welfare reforms of the 1990s moved millions of Americans back into the labor force, spurring greater economic growth. Welfare caseloads plummeted, employment rose, and poverty rates dropped, particularly among the most at-risk populations.”

PC Response: The FGA cites two studies, neither of which provides definitive evidence to support these claims, particularly for SNAP able-bodied adults without dependents – the supposed focus of the FGA's “analysis.” Indeed, the primary focus of both studies is about the role of AFDC/TANF and neither isolates the impact of “welfare reform,” much less work requirements, from other possible confounding factors. Most of the 1990s were characterized by a period of strong economic growth and expansions in aid to the working poor, particularly the Earned Income Tax Credit (EITC) and Medicaid and related coverage.

The impact of “state-led, work-first welfare reforms” is best examined by looking at rigorous evaluations of such programs. Researchers at RAND prepared a comprehensive synthesis of the impact of dozens of state welfare reform programs – including mandatory, welfare-to-work programs – on welfare caseloads, child poverty, and a range of other outcomes.²¹ The random assignment evaluations they reviewed examined programs in the very period when caseloads and poverty fell rapidly nationally. While most reform programs showed declines in welfare receipt, and some showed reductions in poverty, the magnitude of the impacts was considerably smaller than suggested by the simple trends in national data highlighted by the FGA authors. This is because the control group also benefitted from a strong economy and increased aid to the working poor.

It is also important to understand that TANF is not really “welfare reform,” but a form of revenue sharing. While the 1996 law sent a strong message about work requirements and time limits neither of those provisions was implemented as intended. Indeed, if the law had an effect on the outcomes the FGA authors suggest, then it would be likely the fact that TANF's early years were characterized by a massive windfall in federal funding. This is because the block grant was based on funding levels in the early to mid-1990s, when caseloads were at historic highs. Almost immediately, most states received 20 to 30 percent more in federal funding than they would have had without TANF. Meanwhile, TANF actually weakened work requirements. Because of the caseload reduction credit, by 2000, 31 states had a 0 percent target. When that wasn't enough, states could take advantage of an array of loopholes created by the legislation itself. For an extensive discussion about the failure of TANF's work requirements, see:

- “TANF Work Requirements: An Epic Fail,” in *TANF is Broken! It's Time to Reform “Welfare Reform” (And Fix the Problems, Not Treat their Symptoms)*, July 25, 2015 draft, available [here](#).

- “The Failure of TANF Work Requirements: A *Much Needed* Tutorial for the Heritage Foundation and the American Enterprise Institute,” August 12, 2016, available [here](#).
- “The Failure of TANF Work Requirements in 2015: The Need for ‘A *Much Better* Way,’” December 20, 2016, available [here](#).

In any event, for purposes of assessing SNAP’s “work requirement” for able-bodied adults without dependents, TANF’s history is irrelevant. The programs have different funding structures, different work requirements, and different populations. The TANF experience is not generalizable to the SNAP “work requirement.”

The FGA: “This new evidence provides policymakers with an opportunity to rethink how they approach welfare. Reformers must refocus their anti-poverty efforts on freeing people from welfare completely instead of simply reforming the welfare experience itself. Policymakers across the country should take a page from Kansas, restore the working class, and give real hope to millions trapped in lives of dependency and poverty.”

PC Response: The FGA report provides absolutely no credible evidence regarding the impact of the SNAP “work requirement” for able-bodied adults without dependents. If anything, a more careful examination of the data suggests that most of the adults that left SNAP did so without employment and were worse off. Even without a credible counterfactual, the quarterly increase in employment rates is small.

Sadly, some policymakers appear to take the FGA’s report seriously. For example, a press release by the Kansas Department of Children and Families states:

A new report from the Foundation for Government Accountability (FGA), focuses on welfare reforms in Kansas and how they have positively impacted low-income residents who previously relied on food assistance. Among some of the findings, before work requirements, 93 percent of able-bodied adults without dependent children (ABAWDs) on welfare were in poverty. Most were in severe poverty and not working at all. Within one year of leaving welfare, their incomes had more than doubled, an increase of 127 percent on average.

“Our staff worked closely with FGA to supply data,” said Kansas Department for Children and Families (DCF) Secretary Phyllis Gilmore. “We are pleased to see the information shared turned into a meaningful report that demonstrates for Kansas and the rest of the nation that employment is the most effective path out of poverty. And our reforms, led by Governor Sam Brownback, are empowering Kansans to achieve self-reliance.”²²

The Kansas press release follows the FGA pattern of cherry-picking statistics to generate a misleading appearance of success. For example, the statement notes that 93 percent of able-bodied adults without dependents were poor before “work requirements” and that their incomes increased 127 percent afterwards. Leaving aside that both figures are wrong because the FGA

used the wrong baseline quarter, the Kansas press release fails to note other important statistics, namely that the poverty rate one year after imposition of the “work requirement” was still 79 percent, that nearly 60 percent of leavers did not have an employment record in the fourth quarter of exiting, and that even among workers the poverty rate was 51 percent.

The FGA’s report on the SNAP “work requirement” uses a weak methodology and the conclusions stemming from the report are wrong, misleading, and/or exaggerated. Worse, the authors show no concern for those who lost their benefits and remained in poverty. There is no attempt to examine the implementation of the “work requirement” to assess the accuracy of disability determinations, the availability of employment-related activities, or to follow-up with a sample of those individuals who cut-off to assess their experiences. What the FGA report should really do is to lead policymakers to “rethink” the use of simplistic pre-post “studies” as evidence of causation.

Conclusion

Evaluation can be an important tool for policymakers when done right. Indeed, this was a central element in President Reagan’s “waiver-based” approach to “welfare reform. In 1987, he started encouraging states to use existing authority to conduct welfare reform experiments – through waivers of AFDC’s rigid rules (and, to a lesser extent, food stamp and Medicaid rules due to more limited waiver authorities for those programs). This approach had strict accountability provisions, most notably cost neutrality (not block grants) and a requirement for a rigorous evaluation.

During the early years of this waiver process, there was an ongoing debate about what constitutes a “rigorous” evaluation and a range of approaches was considered, including statistical models and various comparison group designs.²³ A simple pre-post model, like the one used in the FGA report, was rejected outright, because it was deemed to be unreliable. The early experiences in several states using comparison groups and statistical models suggested that even these approaches were imperfect, even though in practice they captured many more factors than a simple pre-post study. In the end, we looked to the “gold standard” of evaluation – random assignment. As noted above, the findings from random assignment experiments are considered the most credible, because the experimental and control groups are alike and subject to the same external conditions, with the only difference being the intervention itself.

President Reagan’s evidence-based approach was continued by President Bush and President Clinton. It provided rigorous evidence, including many examples of state experiments that increased employment and earnings, and also reduced welfare dependency and poverty.²⁴ TANF ended this approach to welfare reform and now there is little credible evidence about the effects of state welfare reform efforts. In recent years, conservatives have increasingly relied on simplistic pre-post comparisons to make claims about the effectiveness of various welfare reform efforts and justify their policy prescriptions. The FGA report is one such example. This approach is not a reliable basis for making policy – conservatives must again recommit themselves to rigorous evaluations to find out what really works and what doesn’t.

¹ The views in this document reflect my own as a citizen and do not reflect the views of any organization I am now or have ever been affiliated with. By way of background, I am a conservative and have worked on welfare issues for the Heritage Foundation, the American Enterprise Institute, and the White House under both President Reagan and President George H.W. Bush. This paper assumes the reader has a basic understanding of the TANF program, but for those readers who want more context and background, see Peter Germanis, *TANF is Broken! It's Time to Reform "Welfare Reform" (And Fix the Problems, Not Treat their Symptoms)*, July 25, 2015 draft, available at: <http://mlwiseman.com/wp-content/uploads/2013/09/TANF-is-Broken.072515.pdf>.

² Nic Horton, Jonathan Ingram, and Josh Archambault, "First-Of-Its-Kind Study Shows The Power Of Work," *Forbes*, February 19, 2016, available at: <https://thefga.org/wp-content/uploads/2016/12/Forbes-First-Of-Its-Kind-Study-Shows-The-Power-Of-Work.pdf>. The study they refer to is: Jonathan Ingram and Nic Horton, *The Power of Work: How Kansas' Welfare Reform Is Lifting Americans Out of Poverty* (The Foundation for Government Accountability, February 16, 2016), available at: <https://thefga.org/download/PowerOfWork-KansasWelfareReform.pdf>.

³ To maintain consistency with the terminology in the FGA report, I will refer to the provision as a "work requirement," but use quotation marks to emphasize how it has been implemented in practice.

⁴ Orley Ashenfelter, "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 1978, vol. 40, no. 1, pp. 47-57.

⁵ Howard S. Bloom, "Building a Convincing Test of a Public Housing Employment Program Using Non-Experimental Methods: Planning for the Jobs-Plus Demonstration," MDRC Working Papers on Research Methodology, October 1996, available at: https://www.mdrc.org/sites/default/files/full_32.pdf.

⁶ Dorothy Rosenbaum and Ed Bolen, "SNAP Reports Present Misleading Findings on Impact of Three-Month Time Limit," Center on Budget and Policy Priorities, December 14, 2016, available at: <http://www.cbpp.org/research/food-assistance/snap-reports-present-misleading-findings-on-impact-of-three-month-time>.

⁷ Dorothy Rosenbaum and Ed Bolen, "SNAP Reports Present Misleading Findings on Impact of Three-Month Time Limit," Center on Budget and Policy Priorities, December 14, 2016, available at: <http://www.cbpp.org/research/food-assistance/snap-reports-present-misleading-findings-on-impact-of-three-month-time>.

⁸ The authors also note a 75 percent decrease in the SNAP ABAWDs caseload. Much of the caseload decline is likely due to the "work requirement," because the change was so large and so quick. However, some of this decline would have undoubtedly occurred anyway due to the improving economy. For example, according to Table 5a of the FGA report, the number of able-bodied adults without dependents declined from 30,121 in June 2013 to 25,913 in December 2013. In addition, whereas the employment and poverty statistics cited by the FGA authors reflect able-bodied adults without dependents leaving in December 2013 the caseload statistics include this group as well as any "entry effects."

⁹ Jonathan Freedman, "Early Head Start" and a related project proposal for "developing a television documentary with a national scope and focus based on the Comprehensive Child Development Program (CCDP)."

¹⁰ Robert G. St.Pierre, Jean I. Layzer, Barbara D. Goodson, and Lawrence S. Bernstein, *National Impact Evaluation of the Comprehensive Child Development Program* (Washington, DC: Abt Associates Inc., June 1997), available at: https://www.acf.hhs.gov/sites/default/files/opre/ccdp_fullreport.pdf.

¹¹ Robert G. St.Pierre, Jean I. Layzer, Barbara D. Goodson, and Lawrence S. Bernstein, *National Impact Evaluation of the Comprehensive Child Development Program* (Washington, DC: Abt Associates Inc., June 1997), p. EX-7, available at: https://www.acf.hhs.gov/sites/default/files/opre/ccdp_fullreport.pdf.

¹² Robert G. St.Pierre, Jean I. Layzer, Barbara D. Goodson, and Lawrence S. Bernstein, *National Impact Evaluation of the Comprehensive Child Development Program* (Washington, DC: Abt Associates Inc., June 1997), p. EX-7, available at: https://www.acf.hhs.gov/sites/default/files/opre/ccdp_fullreport.pdf.

¹³ This is another odd adjustment – the authors of the report could have compared changes in quarterly earnings, as the percent increase would be the same. The problem with "annualizing" earnings is that many of those employed in a particular quarter may not be employed during the year and/or have the same earnings each quarter.

¹⁴ In the figure on p. 16 showing annualized wages and food stamp benefits, they use the third quarter of 2013 for earnings and the fourth quarter of 2013 for SNAP benefits, suggesting a confused approach.

¹⁵ Dorothy Rosenbaum and Ed Bolen, "SNAP Reports Present Misleading Findings on Impact of Three-Month Time Limit," Center on Budget and Policy Priorities, December 14, 2016, available at:

<http://www.cbpp.org/research/food-assistance/snap-reports-present-misleading-findings-on-impact-of-three-month-time>.

¹⁶ Joshua Leftin, Nancy Wemmerus, James Mabli, Thomas Godfrey, and Stephen Tordella, *Dynamics of Supplemental Nutrition Assistance Program Participation from 2008 to 2012* (Arlington, VA: Decision Demographics, December 2014), available at: <https://fns-prod.azureedge.net/sites/default/files/ops/Dynamics2008-2012.pdf>.

¹⁷ Joshua Leftin, Nancy Wemmerus, James Mabli, Thomas Godfrey, and Stephen Tordella, *Dynamics of Supplemental Nutrition Assistance Program Participation from 2008 to 2012* (Arlington, VA: Decision Demographics, December 2014), p.91, available at: <https://fns-prod.azureedge.net/sites/default/files/ops/Dynamics2008-2012.pdf>.

¹⁸ Mary Jo Bane and David T. Ellwood, "The Dynamics of Dependence: The Routes to Self-Sufficiency," prepared for the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (Cambridge, MA: Urban Systems Research and Engineering, Inc., June 1983).

¹⁹ Joshua Leftin, Nancy Wemmerus, James Mabli, Thomas Godfrey, and Stephen Tordella, *Dynamics of Supplemental Nutrition Assistance Program Participation from 2008 to 2012* (Arlington, VA: Decision Demographics, December 2014), p. 76, available at: <https://fns-prod.azureedge.net/sites/default/files/ops/Dynamics2008-2012.pdf>.

²⁰ Joshua Leftin, Nancy Wemmerus, James Mabli, Thomas Godfrey, and Stephen Tordella, *Dynamics of Supplemental Nutrition Assistance Program Participation from 2008 to 2012* (Arlington, VA: Decision Demographics, December 2014), p. 71, available at: <https://fns-prod.azureedge.net/sites/default/files/ops/Dynamics2008-2012.pdf>.

²¹ Jeffrey Grogger, Lynn A. Karoly, and Jacob Alex Klerman, *Consequences of Welfare Reform: A Research Synthesis* (Santa Monica, CA: July 2002), available at: http://www.acf.hhs.gov/programs/opre/welfare_employ/res_synthsis/reports/consequences_of_wr/rand_report.pdf.

²² Kansas Department for Children and Families, "New Report Details Success of Kansas Employment Policies," February 16, 2016, available at: <http://www.dcf.ks.gov/Newsroom/Pages/02-16-2016.aspx>.

²³ See Judith M. Gueron and Howard Rolston, *Fighting for Reliable Evidence* (New York, NY: Russell Sage Foundation, June 2013) and Michael E. Fishman and Daniel H. Weinberg, "The Role of Evaluation in State Welfare Reform Waiver Demonstrations," in *Evaluating Welfare and Training Programs*, edited by Charles Manski and Irv Garfinkel, (Harvard University Press, January 1992), pp. 115-142.

²⁴ Jeffrey Grogger, Lynn A. Karoly, and Jacob Alex Klerman, *Consequences of Welfare Reform: A Research Synthesis* (Santa Monica, CA: July 2002), http://www.acf.hhs.gov/programs/opre/welfare_employ/res_synthsis/reports/consequences_of_wr/rand_report.pdf.