Journal of Public Economics 96 (2012) 151–162

Contents lists available at SciVerse ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Work incentives and the Food Stamp Program☆

Hilary Williamson Hoynes [a,c], Diane Whitmore Schanzenbach [b,c,*]

[a] Department of Economics, University of California, Davis, CA, United States
[b] School of Education and Social Policy, Northwestern University, IL, United States
[c] NBER, MA, United States

ARTICLE INFO

Article history:
Received 12 July 2010
Received in revised form 15 August 2011
Accepted 24 August 2011
Available online 8 September 2011

JEL classification:
H31
J22

Keywords:
Welfare policy
Labor supply

ABSTRACT

Labor supply theory makes strong predictions about how the introduction or expansion of a social welfare program impacts work effort. Although there is a large literature on the work incentive effects of AFDC and the EITC, relatively little is known about the work incentive effects of the Food Stamp Program and none of the existing literature is based on quasi-experimental methods. We use the cross-county introduction of the program in the 1960s and 1970s to estimate the impact of the program on the extensive and intensive margins of labor supply, earnings, and family cash income. Consistent with theory, we find reductions in employment and hours worked when food stamps are introduced. The reductions are concentrated among families headed by single woman.

1. Introduction

A central question in public finance, one that has generated decades of research, is how tax and transfer programs affect labor supply. The early literature typically estimated cross sectional models thereby ignoring the endogeneity of net wages, net income, and selection into transfer programs. Concerns about biases from these cross sectional approaches led to methodological innovations, in particular structural modeling in the presence of kinked budget constraints.[1] Structural approaches had their own limitations, including concerns about sensitivity to choice of utility function, stochastic assumptions, and so on. The "new public finance" approach followed, with a reliance on using policy-induced variation and quasi-experimental methods to analyze impacts on labor supply. Quasi-experimental approaches have been used to analyze the impact of wide range of policies on labor supply such as federal taxes, the Earned Income Tax Credit, Medicaid, social security, and welfare reform.

Our paper contributes to the literature on taxes, transfers and labor supply by using a quasi-experimental approach to estimate the impact of the Food Stamp Program (FSP) on labor supply. The FSP is a federal means-tested program, providing benefits to buy food for families who are income and asset eligible. Importantly, while the primary goal of the Food Stamp Program is to increase food consumption, Hoynes and Schanzenbach (2009) show that most households are inframarginal and thus food stamp benefits can be treated as an income transfer.

Food stamp benefits are *the* fundamental safety net in the U.S., being the only public assistance program that is available to all family types (most programs are targeted on female-headed households, children, or the elderly). In fact, food stamps is the largest U.S. cash or near-cash means-tested transfer program with spending in 2009 of 50 billion dollars compared to 30 billion for TANF and 40 billion for the federal EITC.[2] The importance of the FSP program is particularly apparent in the current great recession, where more than 1 in 9 persons is receiving food stamps.

A central challenge for the empirical food stamp literature is that the program is federal and exhibits no variation across states, which is an approach commonly used in the quasi-experimental literature.[3]

☆ We are grateful to Bob Schoeni and Donna Nordquist for help with the PSID. Alan Barreca and Rachel Henry Currans-Sheehan provided excellent research assistance. Schanzenbach thanks the Joint Center for Poverty Research USDA Food Assistance and Nutrition Research Innovation and Development Grants in Economics Program and the Population Research Center at the University of Chicago for generous financial support.

* Corresponding author at: School of Education and Social Policy, Northwestern University, IL, United States.
E-mail addresses: hwhoynes@ucdavis.edu (H.W. Hoynes), dws@northwestern.edu (D.W. Schanzenbach).

[1] For example, innovating papers that use structural models applied to kinked budget constraints include Burtless and Hausman (1978) analyzing the negative income tax, Hausman (1980) on the federal income tax, and Moffitt (1983) on AFDC.

[2] Food stamp program statistics are available at http://www.fns.usda.gov/pd/SNAPsummary.htm; AFDC/TANF statistics are available at http://www.acf.hhs.gov/programs/ofa/data-reports/index.htm, and EITC program data is available at http://www.taxpolicycenter.org/TaxFacts/index.cfm.

[3] There was some cross-state variation in eligibility standards in the earliest years until federal standards were adopted in the January 1971 amendments to the Food Stamp Act.

0047-2727/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpubeco.2011.08.006

Further, the universal nature of the program means there are no ineligible groups to serve as controls, which is another common approach in the quasi-experimental literature. Instead, the typical food stamp study in some way compares recipients to nonrecipients leading to a possible bias if there is selection into program participation (Currie, 2003). The small existing literature on the labor supply effects of the Food Stamp Program uses structural estimation with little attention to exogenous variation in the program.[4]

In this paper, we take a very different approach to estimating the labor supply effects of the FSP using the introduction of the program as it was phased in across U.S. counties over a relatively gradual period. We utilize the natural experiment afforded by the nationwide rollout of the modern Food Stamp Program during the 1960s and early 1970s. Our identification strategy uses the sharp timing of the county-by-county rollout of the FSP, which was initially constrained by congressional funding authorizations (and ultimately became available in all counties by 1975). While the existing literature limits attention to hours worked, we examine the impacts of the program on labor force participation, annual hours, earnings, and total family cash income. Further, reflecting the universal eligibility in the FSP, we examine impacts on all families, including married couples and families headed by single women. Our "program introduction" research design has the appeal of relying on non-marginal changes in incentives faced by consumers.[5]

Safety net programs, such as AFDC, TANF and food stamps, are designed to insure a basic level of consumption in low-income families. Consequently, benefits in traditional income support programs feature a guarantee—a benefit level if the family has no income. As earnings or income increases, benefits are reduced leading to an implicit tax rate on earnings (called the benefit reduction rate or BRR). Benefits in the FSP also take this form; for example in 2010 a family of three has a food stamp guarantee of $526 per month and the benefit is phased-out using a benefit reduction rate of 30%.[6] Notably, the benefit reduction rate in the Food Stamp Program is lower than the rate under the old AFDC program or most states' TANF programs.[7]

As is well known, a family's labor supply response to the income transfer program may partially offset the income and consumption enhancing goals of the program. The guarantee produces an income effect and the benefit reduction rate reduces the net wage leading to an income and substitution effect. Standard static labor supply theory predicts that the program will reduce labor supply on both the extensive (employment) and intensive (hours conditional on work) margins. As a result, it may cost more than $1 in income support payments to increase a low-income family's available cash and near-cash resources by $1.

We use data from the Panel Study of Income Dynamics (PSID) from 1968 to 1978 to examine the impact of the FSP on labor supply, earnings, and income. We employ a difference-in-difference model where the treatment is at the county level, with controls for county and year-fixed effects and state linear time trends. In this model, identification requires that there are no contemporaneous county-level trends that are correlated with food stamp introduction and family economic outcomes. We also estimate a triple-difference model that uses variation across subgroups with varying propensities of being affected by food stamps. Our results are robust to adding controls for possible confounders and event study models further support the validity of the research design.

Overall, our results indicate that recipients of the FSP transfer behave as the theory predicts. Although we find no significant impacts on the overall sample, this may be expected given the relatively low participation rates in the population at large. When we limit to families headed by a single woman—a group much more likely to participate in the program—we find a significant reduction of 183 h worked per year (an intent-to-treat estimate) which given the group's program participation rate implies a treatment-on-the-treated estimate of −505 h/year. Our triple-difference estimates show a significant reduction in the employment rate with a treatment-on-the-treated estimate of a 24 to 27 percentage point reduction. We find no significant impacts of the FSP on earnings or family income, though the estimates are imprecise.

The remainder of the paper is as follows. Section 2 summarizes the prior literatures and Section 3 provides a history and summary of the Food Stamp Program. Section 4 describes the expected effects of the FSP on labor supply and Section 5 describes the data. Section 6 presents the results and Section 7 discusses the results in the context of the existing literature. Section 8 concludes.

2. Prior literature

While there is a sizable literature examining the impacts of the Food Stamp Program on family consumption, nutrition, and family well-being, there is little research examining its impacts on labor supply. The prior literature, which is based mostly on structural estimation, finds modest impacts of the FSP on labor supply.

The prior studies of the effect of FSP on labor supply include Fraker and Moffitt (1988), Hagstrom (1996), and Keane and Moffitt (1998). Fraker and Moffitt (1988) use structural models and kinked budget constraints to estimate the impact of food stamps on labor supply for a sample of female heads of household (families headed by a single women). They specify a utility function and model the choice of hours of work (zero, part time, full time) and participation in the AFDC and food stamp programs. Fraker and Moffitt find that the FSP reduces hours of work by participants by 1 h per week among food stamp recipients, or since mean weekly hours worked for food stamp participants is about 9.5, a 9% reduction. Keane and Moffitt (1998) extend this paper, also looking at female-headed households, by simultaneously modeling AFDC, food stamps, Medicaid and housing benefits. They find larger elasticities than Fraker and Moffitt, but do not report simulations for the effect of the overall FSP on labor supply. Hagstrom (1996) estimates the impact of the FSP on married couples' labor supply using a multinomial logit model. He reports small impacts of changes in the food stamp guarantee and BRR on labor supply but does not report simulations for the total effect of the FSP on labor supply. Notably, all of these studies use cross sectional models where food stamp benefits do not vary across families.[8] Moffitt (2002) summarizes the literature by concluding "the Food Stamp Program has little effect on work disincentives." In this paper, we find labor supply effects that are larger than the prior literature.

The vast literature on labor supply effects of a variety of income transfer programs is well summarized elsewhere (Moffitt, 1992; Hoynes, 1997; Moffitt, 2002). The empirical literature on the AFDC program may also be useful here because there is substantial overlap in participation between the two programs. The general findings from the literature on the AFDC program (which provided cash income

[4] These studies include Fraker and Moffitt (1988), Hagstrom (1996), and Keane and Moffitt (1998) and are discussed below.

[5] This "program introduction" research design has been taken in recent analyses of other social programs such as Head Start (Ludwig and Miller, 2007), Medicare (Finkelstein and McKnight, 2008), Title I (Cascio et al., 2010), and in our own prior work on the food stamp program (Almond et al., 2011, Hoynes and Schanzenbach, 2009).

[6] There are additional deductions for dependent care, child support, medical costs, high housing costs, and 20% of earned income.

[7] Up until 1967, the benefit reduction rate in the AFDC program was 100%. It was reduced to 67% in 1967, then increased again to 100% in 1981. After federal welfare reform, and the conversion to TANF, there is substantial variation across states in the programs' BRR.

[8] The papers use data from a single year, so the food stamp schedule varies only by family size (which is a source of variation not used in these papers). The *effective* benefit will vary by state since AFDC income is included in a family's countable income (and AFDC payments vary by state).

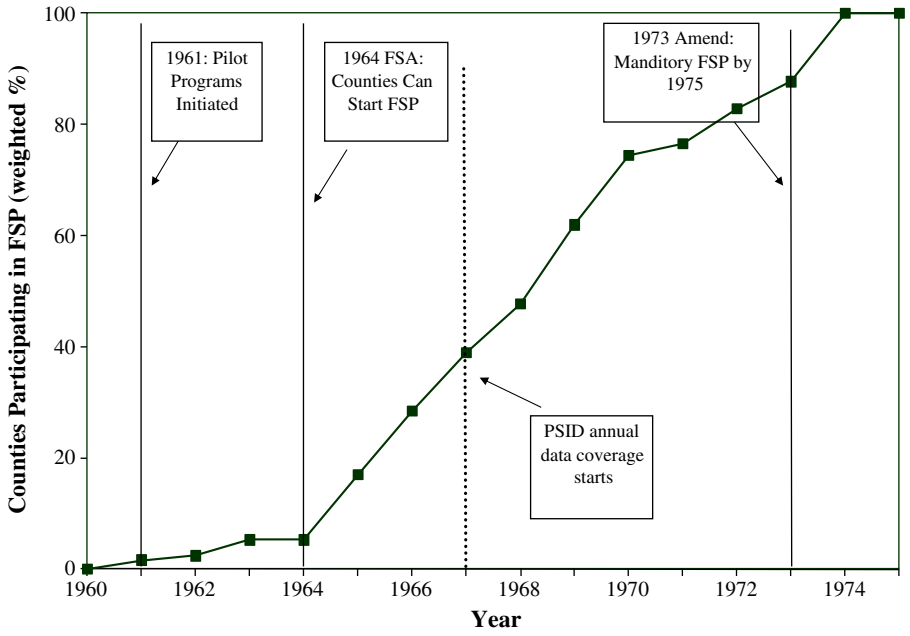

Fig. 1. Cumulative percent of counties with Food Stamp Program, 1960–1975. Source: Author's tabulations of county FSP start dates. Counties are weighted by their 1960 populations.

support) are that AFDC reduces labor supply among program participants by 10 to 50%. In particular, the influential paper by Moffitt (1983) finds that annual hours worked by AFDC recipients are 546 lower per year because of the program. This translates into a reduction of 208 h/year among female-headed households as a whole. The earlier work by Hausman (1981) finds even larger effects, a reduction of 1024 h/year for AFDC recipients. These estimates suggest that income support programs can have large work disincentive effects. Compared to the Food Stamp Program, AFDC benefits and the benefit reduction rate are higher, thus we would expect the work disincentive effects to be lower in the Food Stamp Program.

3. Introduction of the Food Stamp Program

President Kennedy's first executive order was to introduce the modern Food Stamp Program by establishing eight county-level programs. The number of pilot programs grew to 43 by 1963. The pilot programs were seen as a great success, and were credited for improving diets of low-income families while also strengthening markets for farm commodities (Johnson, 1964). Lyndon Johnson expanded the program and made it permanent when he signed the Food Stamp Act on August 31, 1964. The Act gave local areas the authority to introduce a federally funded FSP in their area. Counties introduced the program at a steady rate over the next decade, until Amendments to the Food Stamp Act in 1973 required all remaining counties to adopt the program by 1975.

Fig. 1 summarizes the overall pattern of FSP introduction. The figure plots the percent of counties offering FSP, where the counties are weighted by their 1970 population. Note this is not the food stamp caseload, but represents the percent of the national population that lived in an area offering a FSP. The figure shows that there was a long ramp up period between 1964 and 1975, leading to the eventual universal coverage of the FSP. For example in 1967 (our first year of income data from our main data source—the PSID) a bit less than half of the population lived in counties with FSP; by 1972 coverage rose to over 80% of the population. According to Berry (1984), funding limits were a major factor in the timing that counties moved off of the waiting list and were able to start up their program: "The program was quite in demand, as congressmen wanted to reap the good will and publicity that accompanied the opening of a new project. At this time there was always a long waiting list of counties that wanted to join the program. Only funding controlled the growth of the program as it expanded" (pp. 36–37).

In this paper we rely on variation across counties in the timing of the original introduction of the FSP to isolate the impact of the program on labor supply behavior. Consequently the causal identification of the impact of FSP relies on the exogeneity of the county implementation of the program. An important starting point is to recognize that counties applied for funds to start a local Food Stamp Program and were funded subject to current allocations by Congress. Notably, the county application was voluntary.

In the early 1960s, some counties provided food aid to the poor via the Commodity Distribution Program (CDP). The Federal government would purchase surplus commodities to support farmers and distribute a portion of them directly to low-income families. The CDP was criticized not only for being inefficient relative to normal market channels for distributing food, but also because the timing of the goods distribution was irregular and only a narrow variety of commodities were available.[9] Further, at the time the Food Stamp Act was signed, only about one-third of families living in poverty were participating in the CDP program. Part of this was likely due to the fact that many counties did not offer the program.[10] In general, the CDP was preferred by agricultural interests and the FSP was preferred by advocates for the poor.

Because of the 10-year ramp up to countrywide availability of the FSP, we can exploit that variation as a natural experiment of exposure to the program. For this research design to be a valid approach to studying the labor supply impact of the FSP, though, counties' FSP start dates much be exogenous to other underlying county-level

[9] The most frequently available commodities were flour, cornmeal, rice, dried milk, peanut butter and rolled wheat (Citizens' Board of Inquiry, 1968).

[10] Because the introduction of FSP replaced the existing CDP in some counties, we can think of the estimated impacts as a lower bound. We have not been able to construct a consistent time series for county participation in the CDP, so we are unable to use this information in the empirical analysis.

trends in labor supply.[11] To explore the exogeneity of FSP adoption, we take several approaches. Below we show event study analyses that document the absence of pre-trends in our outcomes variables and show a sharp (though imprecisely estimated) change in measures of labor supply after the FSP is introduced. In addition, here we explore what county-level characteristics predict when counties adopt the program.

Given the politics of the FSP relative to the CDP (Kotz, 1969; Berry, 1984), we expect that northern, urban counties with large poor populations would likely adopt the FSP relatively earlier, while Southern counties and those with strong agricultural interests in general are likely to adopt the program later. To measure this, we take county-level characteristics measured in 1960—that is, before even the introduction of the pilot programs—and use them to predict the timing of a county's eventual adoption of FSP.[12] The dependent variable in this analysis is the month and year of the county's food stamp start date expressed as an index equal to 1 in January 1961. The independent variables include the percent of the 1960 population that lives in an urban area, is black, is less than age 5, is age 65 or older, has income less than $3000 (in 1959$), the percent of the county's land used for farming, and log of the county's total population (constructed from the 1960 Census of Population and Census of Agriculture). All regressions are weighted by the 1960 county population.[13]

The results are presented in Appendix Table 2 and include specifications with and without state fixed effects, and with and without the early pilot counties (which were clearly nonrandom). We find that counties that have a larger overall population and a higher share that is poor, black, very young, or elderly implement FSP earlier. Consistent with the political history of the program, counties with more land used for farming implement the program later. We also find that the relationship between start date and county characteristics is weaker (in absolute value) in Southern counties. While these results show statistically significant impacts of county characteristics in predicting the timing of the introduction of FSP, the quantitative importance of these characteristics is small and most of the variation remains unexplained. To control for possible differences in trends across counties that may be spuriously correlated with the county treatment effect, we include interactions of the 1960 pre-treatment county characteristics with time trends in all of our models (as in Acemoglu et al., 2004 and Hoynes and Schanzenbach, 2009). Note that the main effects of the 1960 county characteristics are absorbed by the county fixed effects. The results are little impacted by the inclusion of these trends. These findings are consistent with Berry's characterization that the exact timing of FSP introduction across counties was driven largely by fiscal constraints and not the lack of desire by counties to introduce the program.

The FSP was introduced around the same time as other programs that were part of the federal "war on poverty." For example, this period included the introduction of Medicaid, Medicare, Head Start, and the Supplemental Nutrition Program for Women, Infants and Children (WIC), and saw expansions of AFDC, Social Security, and disability insurance programs. If these other programs mainly varied at the state level, then our controls for state linear time trends or state-year fixed effects should absorb their impacts. Nonetheless, to control for possible coincident expansions of other programs we also include annual measures of county per-capita transfer payments for cash income support, medical care, and retirement and disability programs (see data section below).

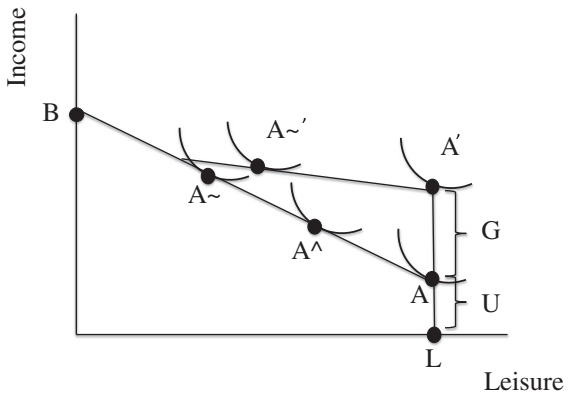

Fig. 2. Income–leisure response to the introduction of food stamps.

4. Labor supply predictions of food stamp introduction

Food stamp benefits have the structure of a traditional income support program, with a guaranteed income benefit that is reduced with family income at the legislated benefit reduction rate. Today, food stamp benefits (recently renamed Supplemental Nutrition Assistance Program or SNAP benefits) are paid via electronic debit card that can be swiped at the checkout line, and can be used to purchase most grocery store food goods. Recipients are allotted a benefit amount B equal to the difference between the federally defined maximum benefit level for a given family size (i.e. G, the guarantee amount) and the amount that the family is deemed to be able to afford to pay for food on its own according to the benefits formula (essentially 30% of cash income, less some deductions). During the time period studied here, the program was slightly different. Through 1978, there was a "purchase requirement" which allowed recipients to buy G dollars worth of food stamp coupons for a price P set by a federal schedule and capped to be no higher than 30% of income Y.[14] The difference between G and P is called the "bonus coupon amount" and is analogous to today's benefit level. During this period, benefits were paid out via coupons that were slightly smaller than dollar bills and could be used to purchase almost all food goods at grocery stores.[15]

We illustrate the labor-leisure tradeoff with and without food stamps in Fig. 2. Like other means-tested programs, the FSP alters the household's labor-leisure tradeoff increasing after tax and transfer income at earnings up to the breakeven point. In particular, the food stamp benefit is largest at zero hours of work, and benefits are reduced as income and earnings are increased leading to an implicit tax rate on earned income. The benefit reduction rate in the Food Stamp Program is 30%, which is lower than other means-tested transfer programs (e.g. 67 or 100% in AFDC, the precursor to TANF).

In Fig. 2, the x axis measures the amount of leisure consumed, and the y axis measures total income including the cash value of in-kind transfers through the Food Stamp Program.[16] Prior to the

[11] In addition, the interpretation of our results will change if the introduction of FSP has general equilibrium labor market effects in the county. Even though the overall participation rate is relatively low, if the employment of food stamp participants is concentrated in certain occupations or industries, the negative labor supply effects may lead to an increase in wages (and hence possible labor supply) of nonrecipients.

[12] This is a replication of the same test conducted in our prior work (Hoynes and Schanzenbach, 2009; Almond et al., 2011).

[13] We drop observations from Alaska throughout the paper due to inconsistencies in mapping FSP service areas to standard county FIPS codes. Here we also drop the (few) counties where the percent of land used for farming is greater than 100 percent, and very small counties with population less than 1000 because of missing data.

[14] In practice, the price P increased by $3 for every additional $10 in income price and averaged about 25% of income for a family of 4. See Appendix Table 1 for the 1969 price schedule.

[15] Benefits cannot be used for hot food intended for immediate consumption, vitamins, alcohol, pet food or paper products. During the time period studied in this paper, they also could not be used to purchase food that was "obviously imported".

[16] For simplicity we model food stamp benefits as cash in the amount of the face value of the coupons, and assume there are no other welfare programs in place. We ignore the purchase requirement and model the program based on the bonus coupon amount. We also ignore that F is flat across small ranges of incomes. Adding any of these to the model complicates the graphs but does not change the prediction or the intuition.

introduction of the FSP, the budget constraint is a straight line with a slope equal to the individual's wage W. The individual has a certain amount of unearned income (U), and the budget constraint is represented by the line BAL. The simple static labor supply model states that an individual maximizes her utility subject to this budget constraint, and assuming a positive labor supply choice, chooses some combination of consumption of goods and leisure at points illustrated for consumers with different preferences by $A^{\sim}$ and $A^{\wedge}$. If her offer wage is below her reservation wage (the slope of the indifference curve at zero hours of work) then it will be optimal to remain out of the labor force, as illustrated by point A (at maximum leisure choice L, or hours = 0).

The introduction of the FSP alters the budget constraint to line $BA'L$ by adding non-labor income G, and rotating the slope of the budget constraint to $W(1-t)$ where t is the tax rate on benefits as income increases (0.3 during this time period). For the individual supplying zero hours of work and consuming only leisure, consumption opportunities increase by the FSP "guarantee" amount G.

As is well known in the analysis of traditional income support programs, this combination of a guaranteed income and benefit reduction rate leads unambiguously to predictions of reductions in the intensive and extensive margins of labor supply. In this case, both the income effect of the benefit as well as the income and substitution effect from the benefit reduction rate leads, unambiguously, to a predicted decline in employment (extensive margin), hours worked (intensive margin), and (if wages are fixed) earnings. In addition, family *cash* income (which as measured does not include food stamp benefits) would also be predicted to fall. (Of course, family total after transfer income including food stamps is likely to increase.)

Referring back to Fig. 2, our representative individual who was, prior to the introduction of the Food Stamp Program, in the labor force and consuming at point $A^{\sim}$, is predicted to increase their leisure (reduce their hours worked) choosing a consumption bundle $A^{\sim}$'. Alternatively, it is possible that the combination of the negative income and substitution effects can push them out of labor market to point A'.

As discussed above, the FSP is run by the U.S. Department of Agriculture, and throughout most of its history the benefit and eligibility criteria have been uniform across states. Thus, at a point in time, the guarantee and benefit reduction rate are constant across all eligible families of a given family size.[17] The prior literature examines the period after the FSP was expanded to a national program. Without variation across states that is often utilized in analysis of government programs, the literature relies on structural estimation where the emphasis is on identifying parameters of the utility function. Despite no variation in the FSP program parameters G and t, by estimating the parameters of the utility function in the presence of the budget constraint, the studies simulate the effect of marginal changes to the FSP (change in G or t) or out-of-sample predictions of the total effect of the FSP. Instead, we use the rollout of the FSP in a quasi-experimental difference-in-difference design whereby we compare counties that have implemented the FSP to other counties who have not yet implanted. We measure whether family labor supply responds to the introduction of the program as predicted by economic theory. Notably, this is a non-marginal calculation; instead we identify the labor supply with and without the program in place.

5. Data

In order to utilize the county-level variation in FSP rollout, we require a dataset that covers as much of the rollout period as possible (1963–1975) and provides information on county of residence. The Panel Study of Income Dynamics (PSID) is a panel data set that began in 1968 with a sample of about 5000 households. Subsequently all members and descendants of the original survey families were reinterviewed annually. The original 1968 sample consists of two subsamples: a nationally representative subsample of 3000 households (Survey Research Center subsample) and a subsample of 1900 households selected from an existing sample of low-income and minority populations (Survey of Economic Opportunity subsample). To adjust for this nonrandom composition, the PSID includes weights designed to eliminate biases attributable to attrition, and to the oversampling of low-income groups. All analyses use the weights provided by the PSID.

The central focus of the PSID is labor market and demographic variables, containing substantial detail on income, employment, and family composition. From this, we measure the annual hours worked and annual earnings over the previous year for both the household head and spouse (if applicable), whether the family reports receiving any food stamp benefits, and total annual family cash income (not including the value of food stamps). We use data from interview years 1968 to 1978.[18] We stop the sample in 1978 because of a significant change in the structure of food stamp payments that begins in 1979.[19]

We limit the sample to include families with a family head that is less than 65. We exclude elderly families because they have lower food stamp takeup rates and are less attached to the labor market. Our nonelderly headed sample consists of 48,168 family-year observations. In some specifications, we limit the sample to nonelderly heads with a high school education or less (37,474 observations), families headed by a single women or "female-headed households with children" (7280 observations), or nonwhite, female-headed households with children (5464 observations) to represent samples with a higher FSP participation rates. Note, unlike other means-tested transfer programs, the FSP is available to married and unmarried families alike as long as they are income eligible. Reflecting their lower family incomes, however, female-headed households have significantly higher eligibility and participation rates compared to their married counterparts.

Appendix Table 3 presents descriptive statistics for the main analysis samples. All dollar amounts are in 2005 dollars. Over 90% of heads report working at some point in the previous year in the overall sample, with 1947 average annual hours. Female-headed households report much lower rates of employment (71% overall, and 62% for nonwhite female heads) and hours worked (1068 and 864, respectively).

The public-use release of the PSID includes only state level identifiers for each family-year. Through special arrangement, we have obtained county-level identifiers for each family in each year. We merge the PSID data using these county identifiers with three additional data sets of county variables. First, the key treatment or policy variable is the month and year that each county implemented a Food Stamp Program, which comes from USDA annual reports on county food stamp caseloads (U.S. Department of Agriculture, various years). We code the FSP policy variable equal to one if the respondent's county of residence has the program by January of that

[16] For simplicity we model food stamp benefits as cash in the amount of the face value of the coupons, and assume there are no other welfare programs in place. We ignore the purchase requirement and model the program based on the bonus coupon amount. We also ignore that F is flat across small ranges of incomes. Adding any of these to the model complicates the graphs but does not change the prediction or the intuition.

[17] The food stamp guarantee, G, varies by size of family.

[18] In 1968, respondents were not asked about food stamps directly, but were instead asked whether they received "free food, clothing or food stamps worth more than $50" in the previous year. In addition, in 1973 no information on food stamps was collected. As a result, we drop 1968 and 1973 for regressions where receipt of food stamps is a dependent variable.

[19] Most notably, the purchase requirement was eliminated. As a result, participants were simply given stamps worth the bonus coupon amount B instead of having to purchase the entire G amount of food stamps at a discounted price.

year.[20] Second, the 1960 City and County Data Book—which compiles data from the 1960 Census of Population and Census of Agriculture—is used to measure economic, demographic, and agricultural variables for the counties pre-treatment (before FSP is rolled out) period. In particular, we use: the percent of the 1960 population that lives in an urban area, is black, is less than 5, is 65 or over, has income less than $3000 (1959$), the percent of land in the county that is farmland, and log of the county population. Finally, we use Bureau of Economic Analysis, Regional Economic Information System (REIS) data to construct annual, county real per-capita income and government transfers to individuals, including cash public assistance benefits (Aid to Families with Dependent Children AFDC, Supplemental Security Income SSI, and General Assistance), medical spending (Medicare and Military health care), and cash retirement and disability payments for each county-year.

6. Results

6.1. Difference-in-difference approach

We begin by estimating a difference-in-difference model using the 1968–78 PSID. This compares labor supply measures across counties and over time relative to when the FSP was introduced. Specifically, we estimate the following model:

$$y_{ict} = \alpha + \delta FSP_{ct} + X_{it}\beta + \sigma CB60_c * t + \gamma REIS_{ct} + \eta_c + \lambda_t + \mu_{st} + \varepsilon_{ict} \quad (1)$$

where y_{ict} is the outcome variable (such as head's employment status or annual hours worked) for family i living in county c in year t. FSP_{ct} is an indicator variable equal to one if there is a Food Stamp Program in county c at time t. X_{it} are family demographic characteristics (urban residence, education and race of head, female-headed household indicator, and state unemployment rate), $CB60_c$ are 1960 county characteristics (interacted with linear time), and $REIS_{ct}$ are county-level per-capita income transfer program data, η_c and λ_t are county- and year-fixed effects respectively, and μ_{st} are either state-specific linear time trends or state-by-year fixed effects.

As described in Section 4 above, we include pre-treatment county characteristics (*CB60*) interacted with linear time trends to control for the observable determinants of county food stamp adoption. The variables in *CB60* include the percent of land in farming and the percent of population black, urban, age less than five, age greater than 65 and with income less than $3000. We also include per-capita county income transfers ($REIS_{ct}$) for (1) retirement and disability programs, (2) medical care (Medicare, Medicare, and military health care), and (3) cash public assistance (AFDC, SSI, and general assistance) to control for coincident introduction or expansion of other programs that are not swept out by state linear time trends or state × year fixed effects. All estimates are weighted using the PSID family weight, and standard errors are clustered at the county level.

Results for the full sample are presented in Table 1. Each cell provides estimates from a separate regression and only the coefficient on the FSP availability variable is reported (standard errors in parenthesis). Columns (1)–(2) cover all nonelderly families. Panel A presents the "first stage" impact of program introduction on an indicator variable for whether the family reports any receipt of food stamps in that year.[21] We find that about 3% of the overall sample reports receiving food stamps when the program is in place. Two specifications are reported for each subsample—the odd columns use state linear time trends and the even columns use state-by-year fixed effects. State-by-year fixed effects may do a better job accounting for concurrent expansion in other welfare programs. The results are largely unchanged when state-by-year effects are included.[22]

Table 1
Impacts of food stamp introduction on labor supply and family income, by group.

	All nonelderly households		Nonelderly, head educ<=12	
	(1)	(2)	(3)	(4)
A. Any food stamps = 1				
County FSP implemented	0.037	0.041	0.051	0.060
	(0.007)***	(0.008)***	(0.009)***	(0.010)***
Number of observations	39,607	39,607	30,889	30,889
B. Head any work = 1				
County FSP implemented	0.000	0.010	0.006	0.019
	(0.005)	(0.008)	(0.007)	(0.010)*
Dependent variable mean	0.926	0.926	0.904	0.904
C. Head annual hours				
County FSP implemented	8	35	16	36
	(20)	(25)	(24)	(31)
Dependent variable mean	1947	1947	1879	1879
D. Head annual earnings				
County FSP implemented	270	−445	−32	−219
	(729)	(960)	(643)	(966)
Dependent variable mean	41,742	41,742	34,600	34,600
E. Log (family income)				
County FSP implemented	0.004	−0.003	−0.001	−0.008
	(0.015)	(0.020)	(0.017)	(0.023)
Number of observations	48,148	48,148	37,447	37,447
1960 cty vars × linear time	X	X	X	X
Year and county fixed effects	X	X	X	X
Per-capita cty transfers	X	X	X	X
State × linear time	X	X	X	X
State × year FE		X		X

Notes: Each parameter is from a separate regression of the outcome variable on a dummy variable equal to 1 if the county-year observation had a Food Stamp Program in place by January of that year. The sample includes nonelderly PSID families using interview years 1968–1978. Observations from Alaska are dropped because of missing data on Food Stamp Program start dates. All outcome variables correspond to annual measures taken as of the interview (in spring of the interview year). Demographic controls include dummies for education, number of children, number of adults, race, urban location and state unemployment rate. 1960 county variables include log of population, percent of land in farming, percent of population black, urban, age<5, age>65 and with income less than $3000, each interacted with a linear time trend. Per capita county transfer income comes from the BEA REIS and includes measures for public assistance (AFDC, General Assistance), medical care (Medicare, Medicaid, military), and retirement and disability benefits. Estimates are weighted using the PSID weight and clustered on county. Standard errors are in parenthesis and ***, **, and * indicate that the estimates are statistically significant at the 1%, 5% and 10% levels, respectively.

In panel B, we report the mean of the dependent variable. We find no impact on the extensive margin of employment for the overall sample.

In panel C, we estimate the impact on head's annual hours of employment, including zeros for those that are never employed in the year. Any measured effect, therefore, can come about from a change in the extensive or intensive margins of work. Again we find no impact of the introduction of the FSP in the overall sample. Next, panel D presents the impact on head's annual earnings (including zeros). Finally, the dependent variable in panel E is log of family income.[23] We find no impact on any of these outcomes in the overall sample.

[20] This is measured for the same year as is reported on in the annual labor supply and income data. That is, since the 1968 survey asks about labor supply and income during calendar year 1967, the corresponding FSP variable for that year reflects availability of the program in their county in 1967.

[21] Recall that because of missing data we drop data for 1968 and 1973 when analyzing food stamp participation. Thus the number of observations is lower in Panel A; we use all available data for the other panels.

[22] Though we are using it as a repeated cross sectional data, the PSID is a panel dataset. When we include dynastic family fixed effects, the results are qualitatively similar (available on request).

[23] We estimate earnings in levels in order to include the zeros, but since there are few (7 observations) zeros in family income we use log income.

Table 2
Impacts of food stamp introduction on labor supply and family income, by group.

	Female heads		Nonwhite female heads	
	(1)	(2)	(3)	(4)
A. Any food stamps = 1				
County FSP implemented	0.223	0.214	0.377	0.280
	(0.047)***	(0.043)***	(0.075)***	(0.065)***
Number of observations	5681	5681	4264	4264
B. Head any work = 1				
County FSP implemented	−0.043	−0.040	−0.048	−0.097
	(0.037)	(0.050)	(0.041)	(0.055)*
Dependent variable mean	0.707	0.071	0.605	0.615
Treatment on treated	*−0.118*	*−0.109*	*−0.088*	*−0.179*
95% CI of ToT	[−0.320,0.084]	[−0.383,0.165]	[−0.236,0.058]	[−0.378,0.021]
C. Head annual hours				
County FSP implemented	−183	−238	−158	−283
	(77)**	(97)**	(74)**	(88)***
Dependent variable mean	1068	1068	864	864
Treatment on treated	*−505*	*−658*	*−289*	*−518*
95% CI of ToT	[-923,-88]	[-1186,-130]	[-559,-20]	[-838,-198]
D. Head annual earnings				
County FSP implemented	−533	−1065	−1193	−3169
	(1112)	(1329)	(986)	(1317)**
Dependent variable mean	14,194	14,194	10,022	10,022
Treatment on treated	*−1472*	*−2943*	*−2190*	*−5816*
95% CI of ToT	[−7520,4577]	[−10169,4284]	[−5768,1389]	[-10596,-1035]
E. Log(family income)				
County FSP implemented	−0.046	−0.096	−0.036	0.015
	(0.051)	(0.060)	(0.064)	(0.086)
Treatment on treated	*−0.128*	*−0.265*	*−0.067*	*0.027*
95% CI of ToT	[−0.403,0.147]	[−0.594,0.063]	[−0.301,0.167]	[−0.286,0.339]
Number of observations [panels B–E]	6890	6890	5175	5175
1960 cty vars × linear time	X	X	X	X
Year and county fixed effects	X	X	X	X
Per capita cty transfers	X	X	X	X
State × linear time	X	X	X	X
State × year FE		X		X

Notes: Each parameter is from a separate regression of the outcome variable on a dummy variable equal to 1 if the county-year observation had a Food Stamp Program in place by January of that year. The sample includes nonelderly PSID families using interview years 1968–1978. Observations from Alaska are dropped because of missing data on Food Stamp Program start dates. All outcome variables correspond to annual measures taken as of the interview (in spring of the interview year). Demographic controls include dummies for education, number of children, number of adults, race, urban location and state unemployment rate. 1960 county variables include log of population, percent of land in farming, percent of population black, urban, age <5, age >65 and with income less than $3000, each interacted with a linear time trend. Per capita county transfer income comes from the BEA REIS and includes measures for public assistance (AFDC, General Assistance), medical care (Medicare, Medicaid, military), and retirement and disability benefits. Estimates are weighted using the PSID weight and clustered on county. Standard errors are in parenthesis and ***, **, and * indicate that the estimates are statistically significant at the 1%, 5% and 10% levels, respectively.

Our identification strategy is based on the introduction of a county-level FSP program, but only the low-income residents of treated counties who are eligible to enroll (and, further, who actually do enroll in the program) should be directly impacted by the program. As shown in Table 3, Food Stamp Program participation varies widely with education, race and family type.[24] To account for this, we limit the overall sample to subgroups that are most likely to be impacted by the program (and similarly, in Table 4 we show placebo regressions limited to subgroups that are unlikely to be impacted by the program).

Columns (3)–(4) in Table 1 limit the sample to nonelderly families with a head who has 12 or fewer years of education. The first stage results in Panel A show that 5 to 6% of families report receiving food stamp benefits after the program is introduced in their county. There is no evidence that head's work effort declines in response to the program, though. In fact when state-by-year fixed effects are included, head's work force participation the prior year actually increases by a marginally statistically significant 2 percentage points. Annual hours remain unchanged. The coefficients on head's annual earnings and family income are negative, but never statistically significant.

As shown in Table 2, the highest food stamp participation rates occur for families with children headed by a single woman (female-headed

Table 3
Food stamp participation rates by demographic group.

	Education of head		
	Less than HS	High school grad	More than HS
White heads			
Single with children	0.41	0.14	0.07
Married with children	0.11	0.05	0.01
Single, no children	0.11	0.04	0.03
Married, no children	0.03	0.00	0.01
Nonwhite heads			
Single with children	0.61	0.45	0.44
Married with children	0.21	0.14	0.03
Single, no children	0.20	0.09	0.04
Married, no children	0.10	0.02	0.02

Notes: Weighted means of food stamp participation rates using families in the 1976–1978 Panel Study of Income Dynamics. These years were chosen because by 1976 all counties had implemented food stamp programs yet it was before the elimination of the purchase requirement in 1979.

[24] The table calculates FSP participation rates for families by the head's race, marital status, and educational attainment, and presence of children using the PSID pooled over the years 1976–78, which is the period in our data after the FSP was implemented in all counties.

Table 4
Impacts of food stamp introduction on labor supply and family income, placebo samples.

	All high-income (>50 K) households		High-income married w/ children		White high-income married w/ children	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Any food stamps = 1						
County FSP implemented	0.005	0.005	0.003	0.006	0.005	0.007
	(0.003)*	(0.003)*	(0.004)	(0.004)	(0.005)	(0.007)
Number of observations	16,658	16,658	9692	9692	1372	1372
B. Head any work = 1						
County FSP implemented	−0.008	−0.015	−0.001	0.000	−0.010	−0.012
	(0.006)	(0.003)*	(0.004)	(0.005)	(0.012)	(0.020)
C. Head annual hours						
County FSP implemented	−16	−35	14	7	−48	−73
	(23)	(32)	(27)	(36)	(61)	(82)
D. Head annual earnings						
County FSP implemented	532	−1063	981	52	3683	421
	(1066)	(1376)	(1366)	(2013)	(4078)	(6113)
E. Spouse any work last year						
County FSP implemented			0.006	−0.011	0.010	0.009
			(0.023)	(0.029)	(0.064)	(0.083)
F. Spouse annual hours						
County FSP implemented			−6	20	−54	−41
			(35)	(45)	(76)	(110)
G. Spouse annual earnings						
County FSP implemented			−159	−106	−1103	−1042
			(539)	(728)	(1380)	(1906)
H. Log(family income)						
County FSP implemented	0.011	0.006	0.003	0.003	0.000	−0.033
	(0.010)	(0.014)	(0.014)	(0.017)	(0.038)	(0.053)
1960 cty vars × linear time	X	X	X	X	X	X
Year and county fixed effects	X	X	X	X	X	X
Per capita cty transfers	X	X	X	X	X	X
State × linear time	X		X		X	
State × year FE		X		X		X

Notes: Each parameter is from a separate regression of the outcome variable on a dummy variable equal to 1 if the county-year observation had a Food Stamp Program in place by January of that year. The sample includes nonelderly PSID families using interview years 1968–1978. Observations from Alaska are dropped because of missing data on Food Stamp Program start dates. All outcome variables correspond to annual measures taken as of the interview (in spring of the interview year). Demographic controls include dummies for education, number of children, number of adults, race, urban location and state unemployment rate. 1960 county variables include log of population, percent of land in farming, percent of population black, urban, age <5, age >65 and with income less than $3000, each interacted with a linear time trend. Per capita county transfer income comes from the BEA REIS and includes measures for public assistance (AFDC, General Assistance), medical care (Medicare, Medicaid, military), and retirement and disability benefits. Estimates are weighted using the PSID weight and clustered on county. Standard errors are in parenthesis and * indicates that the estimates are statistically significant at the 10% level.

households). In Table 3, we present results limited to that subsample. The first two columns present estimates for all female-headed households while columns 3 and 4 limit the sample further to nonwhite female heads of household. This sample reports participating in the program at much higher rates: between 21 and 22% report receiving benefits after the program is introduced in their county. Employment and earnings responses for this sample are completely in line, although not always statistically significantly, with the theoretical predictions outlined in Fig. 2. Among all female heads, the point estimate on annual employment status shows a statistically insignificant 4 percentage point reduction with FSP introduction. Further, although employment rates do not significantly change, there is evidence that female heads reoptimize along the intensive margin as predicted by economic theory by reducing their annual hours of work. In the specification controlling for linear time trends, annual hours are reduced by an average of 183 h (on a mean of 1068). The results are slightly larger when state-by-year fixed effects are included in the model.[25]

Although the effects are not precisely estimated, head's annual earnings and log of family income both appear to decline as well. To benchmark the potential impact on earnings, we calculate the implied impact by assuming constant wages and multiplying by the treatment effect on hours worked. At the mean hourly wage among female heads, this implies about a $2400 reduction in total earnings which lies within the confidence interval for the estimated earnings effect. The implied reduction at the 25th percentile and median also fall within the confidence interval (see Appendix Table 5).

The estimated coefficients in Table 2 represents the intention-to-treat effect, which presumably includes many "zero" impacts for households that are ineligible or do not take up the program. In order to estimate the treatment-on-treated, then, we can inflate the estimated coefficient by the program participation rate, which we get from Table 3.[26] The treatment-on-the-treated effect along with its 95% confidence interval is presented in Table 2. Using this approach, the treatment-on-treated estimate is a 12 percentage point reduction in employment (with a 95% confidence interval spanning

[25] As another robustness check, we interacted the 3 county characteristics that best predict the adoption of the program with year fixed effects (see Appendix Table 5). In addition, we have estimated a Poisson model for all dependent variables measured as counts, and the results are quite similar (available upon request).

[26] We do not use the coefficient estimated in panel A because of missing data and inconsistent definitions of FSP receipt in the early years, and for consistency with the triple-difference results reported further below.

from −0.320 to 0.084). The treatment-on-treated impact on annual hours worked implies a 505-hour reduction (with a 95% confidence interval of −923 to −88).

Finally, we examine the impacts for the nonwhite female-headed families in columns (3)–(4). The estimated first stage of food stamp takeup rate is largest among these families with an estimated 28 to 38% reporting receipt of food stamps after the program is introduced. These families appear to adjust on the extensive margin, with estimates varying from insignificant 4.8 to a statistically significant 9.7 percentage point reduction in labor force participation. Annual hours worked and earnings (which, as noted above, include zeros for non-workers) decline commensurately and are statistically significantly different from zero. The impact on overall family cash income is imprecise and not consistently signed across the specifications.

Existing evidence suggests that the labor supply of a family's secondary worker is more responsive to income transfer programs, so we might expect the spouse's labor supply to decline by more than the head's in a married-couple family (Eissa and Hoynes, 2006; Hoynes, 1996). To test this, we estimated the impact of the introduction of the Food Stamp Program in a sample limited to only married-couple families. Only 2% of families in this sample report receiving benefits after the FSP is introduced, and we found no evidence in this sample that heads or spouses alter their labor supply efforts in response to the introduction of the program.[27]

As a check to insure that the labor market findings are not spurious, in Table 4 we present placebo regressions looking for signs of an "effect" of the FSP on groups that are unlikely to have received treatment. Columns (1) and (2) report results for all high-income (defined as annual income greater than $50,000 in inflation-adjusted 2005 dollars) families, the next pair of columns limits the sample to high-income married couples with children, and the final pair limits the sample even further to white married high-income families with children. As shown in Panel A, these families are very unlikely to report receiving any food stamp benefits. The coefficients on head's employment, hours worked and earnings are small and not statistically different from zero. The sign on annual earnings is generally wrong-signed—indicating an increase in earnings after FSP is introduced in their county—with sizeable standard errors. Effects on spouse reporting any employment here are small and generally wrong-signed, while spousal hours worked and earnings bounce around and are imprecise. These findings indicate that there is no "treatment effect" on families that do not receive treatment and lend further credibility to our research design.

Table 5
Impact of food stamp introduction on labor supply and family income: triple-difference estimates for nonelderly sample.

	(1)	(2)	(3)
Head any work last year			
County FSP implemented × Pg	−0.267	−0.273	−0.242
	(0.081)***	(0.083)***	(0.067)***
County FSP implemented	0.020	0.029	0.018
	(0.007)***	(0.009)***	(0.006)***
Head annual hours			
County FSP implemented × Pg	−281	−302	−269
	(183)	(188)	(168)
County FSP implemented	26	52	25
	(24)	(27)*	(23)
Head annual earnings			
County FSP implemented × Pg	−6728	−6929	−4418
	(5935)	(6168)	(5440)
County FSP implemented	744	–22	600
	(905)	(1099)	(891)
Log(family income)			
County FSP implemented × Pg	−0.139	−0.137	−0.122
	(0.123)	(0.124)	(0.122)
County FSP implemented	0.014	0.005	0.012
	(0.017)	(0.021)	(0.017)
Number of observations	48,148	48,148	48,148
1960 Cty Vars × linear time	X	X	X
Per capita cty transfers	X	X	X
Group fixed effects, group × linear time	X	X	X
Year-fixed effects (main and x Pg)	X	X	X
County fixed effects	X	X	X
State × linear time	X		X
State × year FE		X	
Pg × other covariates (except area fixed effects)			X

Notes: Each parameter is from a separate regression of the outcome variable on the food stamp implementation dummy multiplied by a group food stamp participation rate. The food stamp implementation dummy equals one if the county-year observation had a Food Stamp Program in place by January of that year. The group food stamp participation rate is calculated for each education–race–marital status–presence of children cell using the 1976–78 PSID. The sample includes all years 1968–78 and excludes observations from Alaska. For details on this sample selection, see text. All outcome variables correspond to annual measures taken as of the interview (in spring of the interview year). Demographic controls include dummies for education, number of children, number of adults, race, urban location and state unemployment rate. 1960 county variables include log of population, percent of land in farming, percent of population black, urban, age <5, age >65 and with income less than $3000, each interacted with a linear time trend. Per capita county transfer income comes from the BEA REIS and includes measures for public assistance (AFDC, General Assistance), medical care (Medicare, Medicaid, military), and retirement and disability benefits. Estimates are weighted using the PSID weight and clustered on county. Standard errors are in parenthesis and *** and * indicate that the estimates are statistically significant at the 1% and 10% levels, respectively.

6.2. Triple-difference approach

Unlike virtually all other U.S. public assistance programs, there is no categorical eligibility for the Food Stamp Program. That is, eligibility depends on income and asset tests but it is not targeted to particular demographic groups, such as single parents with children. This argues for using broad samples, such as all nonelderly families or low educated nonelderly families as we did above. However, it is clear from the analysis in Table 1 that we may not have the power to detect effects in a broad sample.

At the same time, the participation rates in Table 3 show that while food stamp participation is highest among single parent families with children, the participation is widespread and varying across many demographic groups. For example, among married families with children where the head has less than 12 years of education, 11% of whites and 21% of blacks participate in food stamps (compared to 41 and 61%, respectively, among single parent families with children with less than a high school education). In the end, in choosing our preferred sample for this analysis, we face a tradeoff between sample size (using the larger sample of nonelderly families but with overall lower participation rates) and targeting (using the smaller sample female heads of household with higher participation rates).

Here, we refine these earlier results by using the entire nonelderly sample, but adopting a triple-difference specification that accounts for different probabilities of being affected by food stamps. Using the participation rates in Table 3, we consider 24 groups g that are defined by race (white, nonwhite), marital status (married, not married), presence of children (yes, no), and educational attainment (<12, 12, >12). Following Bleakley (2007) and Hoynes and Schanzenbach (2009), we use these participation rates (P_g) to scale the food stamp treatment variable, and estimate the following model where the unit of observation is the family-year:

$$y_{ict} = \alpha + \varphi FSP_{ct} + \delta FSP_{ct}P_g + \sigma CB60_c * t + \gamma REIS_{ct} + \theta_g + \theta_g * t + \eta_c + \lambda_t + \mu_{st} + \varepsilon_{ict} \quad (2)$$

[27] See Hoynes (2010) for the detailed results.

In addition to the variables included in Eq. (1), we also include fixed effects for group and group linear time trends. We include the main effect for FSP as well as the interaction with P_g. The coefficient on the main effect can be interpreted as the impact on a group with zero risk of being on food stamps, and we therefore expect the coefficient on the main effect to be zero.

The results are presented in Table 5. In column (1) we include state linear time trends, in column (2) we include state-by-year fixed effects, and in column (3) we augment Eq. (2) with interaction terms between P_g and all control variables. This is a triple-difference specification: across counties, over time, and across groups.[28] Because we inflate by participation rate, the coefficient of interest, δ, can be interpreted as the impact of the treatment on the treated.

The results in Table 5 are consistent with the theoretical predictions that the introduction of an income support program induces households to consume more leisure. Further, the results are remarkably consistent across specifications. Participating heads reduce employment by 24 to 27 percentage points when the FSP is introduced in the county. Head's annual hours also appear to decline with p-values hovering just over 10%. The point estimates on earnings and log family income suggest a commensurate decline, but the estimates never approach statistical significance. In each case, as expected, the coefficients on the main effect for FSP are uniformly small and either wrong-signed or not different from zero. This is a placebo-test of sorts and the small coefficients give us additional confidence about the validity of our design.[29]

6.3. Event study analysis

Our results so far have been quite robust to different ways of controlling for confounding background characteristics. This suggests that our results are identified off of discrete changes in outcomes at the time of FSP introduction. If instead identification were being driven by underlying trends or variation in other characteristics, we would expect more sensitivity of our coefficients to changes in the manner we account for these possible confounders.

To directly examine the timing of the shifts in labor supply relative to the introduction of FSP, we return to the female-headed household sample and conduct an event study analysis. We use the female-headed household sample because they have the highest FSP participation rates and largest responses to FSP introduction in the results presented above. The event study approach requires a discrete treatment, so using the full sample with adjustment for group specific participation rates does not lend itself to this approach.

The advantage of an event study analysis is it traces out the trend in labor supply year-by-year for the periods leading up to and following FSP adoption. This allows for rigorous testing of a "pre-trend," the presence of which would raise concerns about our identification strategy. It also allows for an analysis of the dynamics of the policy effect—for example whether they grow with time since adoption. We are able to estimate an event study analysis here because we have a discrete policy variable that is implemented at different times across the counties in our sample.

Specifically, we estimate the following equation:

$$y_{ict} = \alpha + \sum_{j=-3}^{3} \pi_j 1(\tau_{ct} = j) + \sigma CB60_c * t + \gamma REIS_{ct} + \eta_c + \lambda_t + \mu_{st} + \varepsilon_{ict} \quad (3)$$

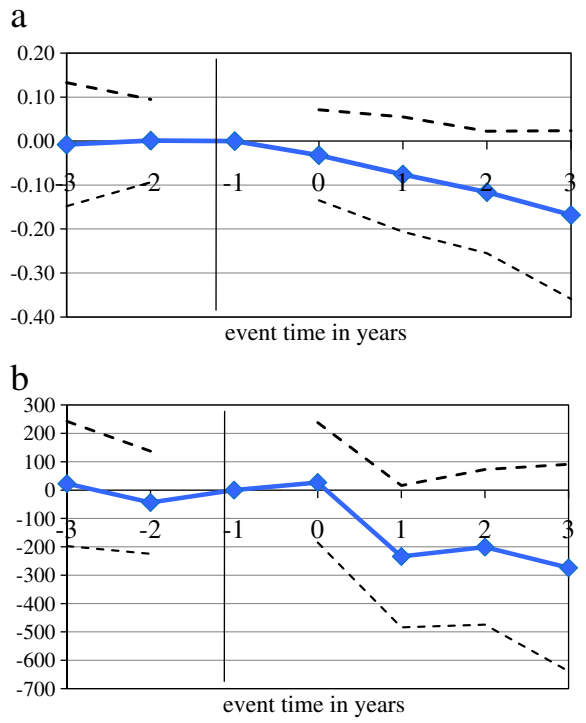

Fig. 3. a: Event study estimates of impact of FSP on head any work last year female-headed households with children only. b: Event study estimates of impact of FSP on head annual hours female-headed households with children only. Notes: the graphs plot estimates and 95% confidence intervals from an event study analysis described by Eq. (3). Coefficients are defined as years relative to the year the Food Stamp Program is implemented in the county. Year 0 is the first year that the county was treated for the entire year. The sample is a balanced county sample, where county is included only if there are observations for all 7 event periods.

where τ_{ct} denotes the year relative to the introduction of the FSP, defined so that $\tau_{ct} = 0$ if for county c the FSP began at any point in calendar year t. For $\tau_{ct} < 0$, work effort decisions were made prior to the introduction of the FSP. All coefficients are measured relative to the omitted coefficient ($\tau_{ct} = -1$). As we did in the tables, we include controls for potential FSP policy endogeneity (that is, the 1960 county characteristics interacted with linear time trends and REIS county controls) and include state linear time trends (μ_{st}). In order to eliminate potential compositional effects, we restrict the sample to a balanced panel of counties having observations for all 7 event periods (i.e., 3 years prior to FSP implementation and 3 years afterwards). Because our data begin in 1968, this means we exclude counties where FSP started in 1970 or earlier. [30]

We produce event study graphs for employment (Fig. 3a) and annual hours worked (Fig. 3b) for the female-headed households. Fig. 3a plots the event-year coefficients from the estimation of Eq. (3) on whether the head was employed last year, with 95% confidence interval bands in dotted lines, and a vertical line at t = −1 indicating that we expect responses to begin in period 0. The figure shows an absence of pre-trends, and shows a sharp turn downward after FSP is introduced, though the coefficients are imprecisely estimated and do not reach statistical significance. Fig. 3b repeats the exercise with annual hours worked by the head and show a qualitatively similar pattern to the employment event study: there is no pre-trend in hours worked prior to the introduction of FSP and after the program is introduced (here one year later), there is a sharp drop in hours worked that is persistent. Results for earnings and log family income look generally similar, and are presented in Appendix Fig. 1.

These results show useful evidence that our county adoption of FSP is identifying the causal impact of the Food Stamp Program.

[28] To implement a full triple-difference specification we would also need to include county-by-group fixed effects, but we lack the statistical power to implement this. Nonetheless, we can push toward this by adding interactions between group and 1) other covariates (in Table 5) and 2) state-linear time (available on request). Including these covariates changes the results little.

[29] Results for married-couple families, available on request, show no statistically significant responses to FSP. Consistent with the literature, we do find that relative to their baseline values the responses are larger for spouses than heads.

[30] Results change little if we exclude the county controls for FSP endogeneity, see Hoynes and Schanzenbach (2010).

Table 6
Impact of food stamp introduction on labor supply and family income 1960, 1970, 1980 census STF analysis.

	Labor force participation rate			
	Females 16 and over	Males 16 and over	Females with children<6	Family income<$10,000 (1979$)
	(1)	(2)	(3)	(4)
A. All races				
County FSP implemented	−0.002	−0.003	0.004	0.009
	(0.001)*	(0.001)**	(0.008)	(0.002)***
Treatment on treated	*−0.027*	*−0.041*	*0.054*	*0.122*
Dependent variable mean	0.396	0.762	0.337	0.238
Number of observations	7898	7898	7898	7898
B. Nonwhites				
County FSP implemented	0.002	−0.004	n/a	0.027
	(0.004)	(0.003)		(0.006)***
Dependent variable mean	0.457	0.703		0.455
Number of observations	7443	7321		7093
1960 cty vars × decade fixed effects	X	X	X	X
Per capita county transfer payments	X	X	X	X
Decade fixed effects	X	X	X	X
County fixed effects	X	X	X	X

Notes: Each parameter is from a separate regression of the outcome variable on a dummy variable equal to 1 if the county-year observation had a Food Stamp Program in place in that year. Data are from 1960 to 1980 Census county-level STF files. Counties in Alaska are dropped because of missing data on Food Stamp Program start dates. Demographic controls include dummies for education, number of children, number of adults, race, urban location and state unemployment rate. 1960 county variables include log of population, percent of land in farming, percent of population black, urban, age<5, age>65 and with income less than $3000, each interacted with a linear time trend. Per capita county transfer income comes from the BEA REIS and includes measures for public assistance (AFDC, General Assistance), medical care (Medicare, Medicaid, military), and retirement and disability benefits. Estimates are weighted using the PSID weight and clustered on county. Standard errors are in parenthesis and ***, **, and * indicate that the estimates are statistically significant at the 1%, 5% and 10% levels, respectively.

These figures show no trend in the outcome variables leading up to the program introduction and sharp changes in the outcome following county program introduction (although, admittedly the estimates are quite imprecise). In our prior work analyzing the impact of the FSP introduction on infant health (Almond et al., 2011) we used county FSP caseload data to show that the food stamp programs seemed to ramp up quite quickly in the 1–2 years following program adoption. This is consistent with the outcome results in Fig. 3.

6.4. Specification check using decennial census

As a final robustness check, we estimate the impact of FSP program introduction on labor supply using public-use decennial Census data from the county-level aggregate files—known as the STF files. Here, we are limited to the variables and samples that have been pre-tabulated by the Census. We are able to estimate models for female labor force participation rates, male labor force participation rates, labor force participation rates for women with a child under age six, and the fraction of families with income below 10,000 (in real 1979 dollars).[31] We pool county outcomes from the 1960, 1970 and 1980 Censuses and estimate models similar to those presented above, with decade fixed effects replacing the year fixed effects. The regressions are weighted using county population and the standard errors are clustered on county. The results are presented in Table 6—with panel A reporting estimates for all persons and panel B reporting estimates for nonwhites (for variables that are available).[32]

It is important to note that these treatment groups are broader than the targeted samples used in the PSID and there is no way to weight the treatment by group participation rate because the data are county (or multi county) averages. With that said, the results show a relatively statistically precise, small negative estimated work disincentive effect. For example, the estimate for males (females) shows that implementing a food stamp program leads to a statistically significant 0.003 (0.002) decrease in the labor force participation rate compared to the mean value of 0.76 (0.40). Using an average participation rate of 7.4, these imply TOT impacts on women of 2.7 percentage points (in italics in the table) or about 7%. For men, the TOT impacts are 4.1 percentage points, or 5% of the mean. In addition, the probability that overall family cash income (not including food stamps) is less than $10,000 per year (in 1979 dollars) increases by a statistically significant 1 to 3 percentage points.

7. Magnitude of the impacts

Overall, the evidence from the PSID and decennial Census is consistent with our theoretical predictions. We generally find that the introduction of the Food Stamp Program leads to lower rates of employment and hours worked. The evidence is less clear for earnings and family income as our results never approach statistical significance for these outcomes. The validity of the research design and estimates is supported by several additional results, although not all reach statistical significance. First, the impacts are found in subsamples with higher participation rates (e.g. female heads of household) and are essentially zero for placebo groups such as highly educated married couples. Second, the event study results show that the policy introduction is unrelated to trends in the outcome variables, and labor supply sharply declines with the introduction of the program. Beyond the sign and statistical significance, however, what do our magnitudes tell us about the work disincentive effects of the Food Stamp Program? How do our results compare to the literature?

Our preferred estimates use the pooled sample and the triple-difference specification (Table 5), as this takes advantage of the universal nature of the FSP program. However, the existing food stamp literature does not provide any estimates for comparison. As a check and to gauge the magnitude of the expected labor supply effects of the Food Stamp Program, we simulated the impact of the program on annual hours worked in our PSID sample using estimated labor supply elasticities from the Negative Income Tax experiments (Robins and West, 1983). More specifically, we assign the compensated wage

[31] The labor force participation rates are defined for persons age 16 and over.

[32] The smallest geographic area identified in the census public use micro data is the *county group*, these are contiguous groups of counties with a combined population of 100,000 or more We estimated models using the 1970 and 1980 public use micro data IPUMS data where we aggregated the FSP treatment across all counties in the country group. Further, we had to combine county groups to accommodate the changing county group boundaries between the 1970 and 1980 Census. In the end, this aggregation was substantial and the results had very low power. The results are available on request.

and income effects based on family type (female-headed households are assigned the effects for "single females" and all male headed singles/families are assigned the effects for "husbands"). We assign maximum food stamp benefits and the food stamp breakeven income level by family size using the 1975 food stamp parameters from Table 2.2 of MacDonald (1977). The change in net wages is −0.3 reflecting the food stamp benefit reduction rate. We assume no nonlinear response to the implementation of food stamps, and therefore simulate the change in hours for those families with head's earnings below the food stamps breakeven point.[33] The simulations for the full nonelderly sample predicted a 20–24-hour reduction in annual hours, which scaled up by the group FS participation rate implies about a 300-hour reduction in hours for participating households. This is quite similar to our (insignificant) estimated reduction in Table 5. Overall, we take this as a useful exercise which corroborates our estimates of moderate work incentive effects in the Food Stamp Program in the full sample.

In the female-headed household subsamples where we find our largest and most robust estimates, there is more scope to compare our results to the literature. We focus on the results for annual hours, since that is the measure used in the literature. Table 21 results show a decrease in hours of 183–238 for all female heads, implying a TOT impact of 500–600 h/year (or more than a 50% reduction). This is considerably larger than the Fraker and Moffitt (1988) results, whose structural model results indicate that food stamp participants reduce their work effort by less than 100 h/year.

It is also instructive to compare our results to the related literature examining the impacts of the AFDC program on labor supply. As summarized above in Section 2, AFDC is estimated to reduce labor supply (annual hours) among program participants by 10 to 50% (e.g., Moffitt, 1983 reports a 546-hour reduction). Our estimates for female-headed households are on the high end of this range. At the same time we would have expected the FSP to yield smaller impacts than AFDC given the program's lower benefit reduction rate and lower guarantee.

8. Conclusion

In this paper we present evidence on the work incentive effects of food stamps, the largest cash or near-cash transfer program in the safety net. This paper provides an important contribution to the literature on work incentives of social welfare programs, and is the first paper on employment effects of the Food Stamp Program that uses a quasi-experimental research design. The impacts of food stamps on work behavior have been difficult for researchers to isolate because there is little cross-state or overtime program variation to exploit. Here we use county variation in the adoption of the program between 1963 and 1975 to identify the impact of food stamps. Using the PSID, our results are consistent with the theoretical predictions, but are limited by power given the combined effect of the relatively low food stamp participation rate and the PSID sample size. We find no significant impacts on the overall sample, but when we limit to single-parent households with a female head—a group much more likely to participate in the program—we find a significant intent-to-treat estimate of a reduction of 183 annual hours (treatment-on-the-treated reduction of 505 annual hours). Our triple-difference estimates imply a 24 to 27 percentage point reduction (treatment-on-the-treated) in the employment rate. We find no significant impacts of the FSP on earnings or family income, though the estimates are imprecise. Overall, our results suggest a larger work disincentive effect of FSP for female heads than had previously been concluded in the literature.

Even though there have been changes in the population of the United States, in the labor force attachment of women, and some of the parameters of FSP have changed a little bit, we argue that these results are still relevant for today's policy debates because they provide insight into labor force responses to income support programs that are not explicitly tied to work. To date, there is little credible evidence on the impact of FSP on work behavior. Understanding the incentives in FSP is important since such a large fraction of Americans rely on FSP, and it is one of the few remaining safety net programs that does not have a substantial work requirement component.

[33] Full results of the simulation are available on request.

References

Acemoglu, D., Autor, D., Lyle, D., 2004. Women war and wages: the impact of female labor supply on the wage structure at mid-century. Journal of Political Economy 112 (3), 497–551 (June).

Almond, D., Hoynes, H.W., Schanzenbach, D.W., 2011. The impact of the Food Stamp Program on infant outcomes. Review of Economics and Statistics 93 (2), 387–403.

Berry, J.M., 1984. Feeding Hungry People: Rulemaking in the Food Stamp Program. Rutgers University Press, New Brunswick, NJ.

Bleakley, H., 2007. Disease and development: evidence from hookworm eradication in the American South. Quarterly Journal of Economics 122 (1), 73–117.

Burtless, G., Hausman, J., 1978. The effect of taxation on labor supply: evaluating the Gary negative income tax experiment. Journal of Political Economy 86 (6), 1103–1130.

Cascio, E., Gordon, N., Lewis, E., Reber, S., 2010. Paying for progress: conditional grants and the desegregation of southern schools. The Quarterly Journal of Economics 125 (1), 445–482.

Citizens' Board of Inquiry, 1968. Hunger, U.S.A. Beacon Press, Boston.

Currie, J., 2003. U.S. Food and Nutrition Programs. In: Moffitt, R. (Ed.), Means-tested Transfer Programs in the United States. University of Chicago Press, Chicago.

Eissa, N., Hoynes, H.W., 2006. The hours of work response of married couples: taxes and the earned income tax credit. In: Agell, J., Sørensen, P.B. (Eds.), Tax Policy and Labor Market Performance. MIT Press (2006).

Finklestein, A., McKnight, R., 2008. What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. Journal of Public Economics 92, 1644–1669.

Fraker, T., Moffitt, R., 1988. The effect of food stamps on labor supply: a bivariate selection model. Journal of Public Economics 35, 25–56.

Hagstrom, P., 1996. The food stamp participation and labor supply of married couples: an empirical analysis of joint decisions. Journal of Human Resources 31 (2), 383–403.

Hausman, J., 1980. The effects of wages, taxes and fixed costs on women's labor force participation. Journal of Public Economics 14 (2), 161–194.

Hausman, J., 1981. Labor supply. In: Aaron, H., Pechman, J. (Eds.), How Taxes Affect Economic Behavior. Brookings Institution, Washington, D.C.

Hoynes, H.W., 1996. Welfare transfers in two-parent families: labor supply and welfare participation under the AFDC-UP Program. Econometrica 64 (2), 295–332.

Hoynes, H.W., 1997. Work, welfare, and family structure: what have we learned? In: Auerbach, A.J. (Ed.), Fiscal Policy: Lessons from Economic Research.

Hoynes, H.W., Schanzenbach, D.W., 2009. Consumption responses to in-kind transfers: evidence from the introduction of the Food Stamp Program. American Economic Journal—Applied Economics 1 (4), 109–139.

Hoynes, H.W., Schanzenbach, D.W., 2010. Work incentives and the Food Stamp Program. NBER Working Paper #16198.

Johnson, L.B., 1964. Remarks upon Signing the Food Stamp Act. Available at http://www.presidency.ucsb.edu/ws/index.php?pid=26472.

Keane, M., Moffitt, R., 1998. A structural model of multiple welfare program participation and labor supply. International Economic Review 39 (3), 553–589.

Kotz, N., 1969. Let them Eat Promises: The Politics of Hunger in America. Doubleday, New York.

Ludwig, J., Miller, D., 2007. Does head start improve children's outcomes? Evidence from a regression discontinuity design. Quarterly Journal of Economics 122 (1), 159–208.

Moffitt, R., 1983. An economic model of welfare stigma. American Economic Review 73 (5), 1023–1035.

Moffitt, R., 1992. Incentive effects of the U.S. welfare system: a review. Journal of Economic Literature 30 (1), 1–61.

Moffitt, R., 2002. Welfare programs and labor supply. In: Auerbach, A.J., Feldstein, M. (Eds.), *Handbook of Public Economics*. Amsterdam: North Holland.

Robins, P., West, R., 1983. Labor supply response. In: International, S.R.I. (Ed.), Final Report of the Seattle/Denver Income Maintenance Experiment. United States Government Printing Office, Washington D.C.

U.S. Department of Agriculture (various years). "Food Stamp Program, Year-End Participation and Bonus Coupons Issues," Technical report, Food and Nutrition Service.