Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach

Patrick Kline
UC Berkeley/NBER

Melissa Tartari*
University of Chicago

May, 2014

Abstract

We study the impact of Connecticut's Jobs First (JF) welfare reform experiment on the labor supply decisions of a sample of welfare applicants and recipients. Although the experiment identifies the distribution of choices made in the absence and presence of reform, each woman's counterfactual choice is unknown. We show that economic theory restricts the counterfactual choices compatible with each woman's actual choice. We use these restrictions to develop bounds on the frequency of intensive and extensive margin responses to reform. Our results indicate that the JF experiment led some women to work and others to reduce their earnings.

Keywords: labor supply, bounds, intensive/extensive margin response, revealed preference

JEL Codes: J22, H20, C14

*This paper previously circulated under the title "What Distributional Impacts Mean: Welfare Reform Experiments and Competing Margins of Adjustment." We thank Andres Santos and seminar participants and discussants at the NBER Public Economics and Labor Studies Program Meetings, the Harris School, the New Economic School, Bates White, Mannheim University, the Minneapolis Federal Reserve, MIT, Queens University, UC Berkeley, University College London, UCLA, UCSD, Stanford, the Federal Reserve Bank of Chicago, and Yale for useful comments. Stuart Craig and Attila Lindner provided outstanding research assistance. This research was supported by NSF Grant #0962352.

The U.S., like other advanced economies, has an extensive system of transfer programs designed to provide social insurance and equity. By affecting work incentives, these programs can induce individuals to enter or exit the labor force (extensive margin responses) or to alter how much they earn conditional on working (intensive margin responses).[1] The relative size of these responses is an important input to the optimal design of tax and transfer schemes (Diamond, 1980; Saez, 2002; Laroque, 2005).

Much of the empirical literature concludes that adjustment to policy reforms occurs primarily on the extensive margin.[2] Two sorts of evidence are often cited in support of this position. First, several studies exploiting policy variation fail to find qualitative evidence of intensive margin responses (Ashenfelter, 1983; Eissa and Liebman, 1996; Eissa and Hoynes, 2006; Meyer and Rosenbaum, 2001; Meyer, 2002). Second, in both survey and administrative data, earnings tend not to exhibit much bunching at the budget "kinks" induced by tax and transfer policies (Heckman, 1983; Saez, 2010). Since the excess mass at kink-points is a nonparametric indicator of intensive margin responsiveness (Saez, 2010), many have concluded that such responses are often sharply constrained (e.g., as in Chetty et al., 2011).

We revisit this conclusion by examining the impact of Connecticut's Jobs First welfare reform experiment on the labor supply and program participation decisions of a sample of welfare applicants and recipients. The JF program created strong incentives to work on welfare by instituting strict work requirements and disregarding earnings up to an eligibility threshold (or "notch"). Bitler, Gelbach, and Hoynes (BGH, 2006) show that the JF reform induced a nuanced pattern of quantile treatment effects (QTEs) on earnings qualitatively consistent with substantial intensive margin responsiveness. Specifically, they find that JF boosted the middle quantiles of earnings while lowering the top quantiles, yielding a mean earnings effect near zero. The negative impacts on upper quantiles provide suggestive evidence of an "opt-in" response to welfare (Ashenfelter, 1983), whereby working women are induced to lower their earnings in order to qualify for benefits.

However, to formally conclude that adjustment occurred along a given margin, one must infer what choices an agent would have made if policy rules had been different. In the JF experiment, this means inferring features of the *joint* distribution of potential earnings across two policy regimes. As noted by Heckman, Smith, and Clements (1997), such a joint distribution is only point identified by QTEs under a "rank invariance" assumption that treatment effects don't induce rank reversals in outcomes. There are many reasons to be dubious of this assumption.[3] For instance, rank reversals

[1] Blundell and Macurdy (1999), Moffitt (2002), and Grogger and Karoly (2005) provide reviews.

[2] Heckman (1993), for instance, concludes that "elasticities are closer to 0 than 1 for hours-of-work equations (or weeks-of-work equations) *estimated for those who are working*. A major lesson of the past 20 years is that the strongest empirical effects of wages and nonlabor income on labor supply are to be found at the extensive margin." (emphasis in original). Likewise, many modern models of aggregate labor supply are now predicated on the notion that labor supply is "indivisible" (Hansen, 1985; Rogerson, 1988; Ljungqvist and Sargent, 2011). See Chetty et al. (2011) for an assessment of how macro estimates of these models compare to estimates from micro data.

[3] In a related analysis, BGH (2005) find evidence of rank reversals in the Canadian Self-Sufficiency Project experiment.

are likely if reform induces some skilled women to work (or to stop working).

We replace the rank invariance assumption with a set of theoretically motivated restrictions on earnings and program participation behavior. These restrictions follow from three basic observations. First, utility maximizing women will not choose a dominated option. Second, they will not respond to reform by choosing an alternative made less attractive. Third, they will not respond by choosing an alternative unaffected by reform unless the experiment reduces the payoff to the option they would have chosen in the absence of reform. We use these restrictions to develop bounds on the frequency of intensive and extensive margin responses to the JF reform.

We begin our analysis with an adaptation of the standard frictionless model of labor supply and a discussion of its key qualitative predictions for the impact of the JF program on earnings and program participation choices. Empirically, most of these predictions appear to be satisfied in the data. However, one notable prediction of the frictionless model does not hold. The opt-in effects of welfare reform should be accomplished, in part, via a point mass of women locating at the JF eligibility notch. Using administrative data, we show that there is no evidence of bunching at this notch, despite its obvious salience. We suggest this is likely due to a combination of under-reporting behavior and constraints on the ability of workers to choose their level of earnings. Indeed, many women with ineligible earning levels are observed in our data to receive benefits, suggesting that under-reporting of earnings is likely to be common among program participants, a finding in line with much of the previous literature (Greenberg, Moffitt, and Friedman, 1981; Greenberg and Halsey, 1983; Hotz, Mullin, and Scholz, 2003).

We thus turn to developing the qualitative implications of an augmented model in which workers may lack fine control over their earnings and can, with some cost, under-report their earnings to the welfare agency. The addition of under-reporting opportunities introduces margins of adjustment that have not typically been incorporated into prior models but are potentially important for many transfer programs. Notably, the JF reform may induce some women who would not have worked to earn relatively large amounts which are then under-reported to the welfare agency – an additional violation of the rank invariance condition.

Unlike traditional parametric models of labor supply (e.g. Burtless and Hausman, 1978; Hoynes, 1996; Keane and Moffitt, 1998), our constrained model has no refutable predictions for the cross-sectional distribution of earnings and program participation under a given policy regime.[4] We show however that the model places strong restrictions on the manner in which a woman's earnings and welfare participation choices may respond to reform. These restrictions stem from prior knowledge of the impact of the experiment on work incentives. The JF reform made working on welfare more attractive due to an enhanced earnings disregard but made claiming cash benefits without working less attractive due to strict work requirements. It also left the desirability of many alternatives (e.g. working off welfare) unaltered. By revealed preference, the reform cannot induce women to

[4]See Macurdy, Green, and Paarsch (1990) for an early critique of parametrically structured econometric models of labor supply with nonlinear budget sets.

choose an alternative made less attractive. We use this, and other, simple exclusion restrictions to develop analytical bounds on the probability of responding along each of eight allowable margins, some of which involve intensive margin adjustments and some of which involve extensive margin adjustments.

Applying our identification results, we find evidence of substantial intensive and extensive margin responses to reform. Jobs First incentivized many women who would not have worked to do so and many who would have worked off welfare at low earnings to participate in the JF program. Corroborating the interpretation of BGH, we find a significant opt-in response among women who would have worked at relatively high earnings levels, demonstrating that intensive margin responses are, in fact, present in our sample. We also find that the JF work requirements induced some women to work but under-report their earnings in order to maintain eligibility for benefits. We conclude with a discussion of the robustness of our results to a richer modeling of the policy environment and dynamic considerations.

Our results demonstrate that simple revealed preference arguments allow researchers studying policy reforms to bound the size of competing response margins under very weak assumptions. This insight was anticipated by Heckman, Smith, and Clements (1997) who, in the context of an application to the U.S. Job Training Partnership Act, consider the identifying power of Roy (1951)-type models of optimization for the joint distribution of potential outcomes. Our results are applicable to settings that do not obey strong Roy-style dependence between choices and outcomes. In our setting, partial identification of impact distributions allows us to demonstrate that, while undoubtedly constrained, women do respond to price changes along the intensive margin.[5] Our approach is generalizable to other reforms which alter the value of alternatives in known directions (e.g. EITC expansions, health care reform).

Finally, our results also contribute to a recent literature on partial identification of labor supply models. The bounding approach developed here is similar in spirit to the work of Manski (2012) who uses revealed preference arguments to set-identify features of policy counterfactuals. However, our analysis explores the identifying power of experimental variation in isolating response margins, a subject he does not consider. We also work with a richer model of labor supply and program participation behavior. Blundell, Bozio, and Laroque (2011a,b) also implement a bounds based analysis of labor supply behavior but are concerned with a statistical decomposition of fluctuations in aggregate hours worked rather than formal identification of policy counterfactuals. Their findings, which are compatible with ours, indicate that adjustments along both the intensive and extensive margins are important contributors to fluctuations in aggregate hours worked. Chetty (2012) considers bounds on labor supply elasticities in a class of semi-parametric models with optimization frictions. He also finds evidence of non-trivial intensive margin responsiveness, but relies on stronger parametric assumptions and does not model program participation.

[5] See Heckman and Smith (1995), Heckman, Smith, and Clements (1997), Heckman (2001) and Deaton (2009) for discussions of the importance of understanding the distribution of program impacts.

The remainder of the paper is structured as follows. Section 1 describes the Jobs First Experiment. Section 2 lays out a model without constraints or under-reporting behavior and uses this baseline model to illustrate the possible responses to the JF reform. Section 3 describes the data used. Section 4 provides experimental impacts on a variety of outcomes and assesses their consistency with the baseline model's predictions. Section 5 describes our full model, one that incorporates earnings constraints and under-reporting behavior. Section 6 studies identification of response margins under the model's restrictions. Section 7 provides our main empirical results and Section 8 discusses the robustness of our results to a variety of extensions. Section 9 concludes.

1 The Jobs First Program

Background

With the passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) in 1996, all fifty states were required to replace their Aid to Families with Dependent Children (AFDC) programs with Temporary Assistance to Needy Families (TANF) programs. This change involved a series of major reforms including the imposition of time limits, work requirements, and enhanced financial incentives to work. The state of Connecticut responded to PRWORA by implementing the Jobs First (JF) program which involved each of the major features of TANF.

To study the effectiveness of the reform, the state contracted with the Manpower Development Research Corporation (MDRC) to conduct a randomized evaluation, comparing the Jobs First TANF program to the earlier state AFDC program. Between January 1996 and February 1997, MDRC collected a baseline sample of about 4,800 single parent AFDC recipients and applicants and randomly assigned them to either the new JF program or the old AFDC program with equal probability. To conduct the evaluation, administrative data on earnings and welfare participation were collected for several years prior to and following the date of random assignment and merged with a baseline survey conducted by MDRC. This experiment has been heavily studied, with several analyses finding that the program encouraged work and had important effects on the distribution of both earned and unearned income (Bloom et al., 2002; BGH, 2006, 2010).

Appendix Table 1 provides a summary of the key JF and AFDC program features. In Connecticut, as elsewhere in the nation, earned income reduced the AFDC benefits paid out to a recipient down from a maximum, or "base grant", amount. However, some earned income was disregarded. Specifically, AFDC recipients were eligible for a fixed earnings disregard of $120 for the 12 months following the first month of employment while on assistance and $90 afterwards. They were also eligible for a proportional disregard of any additional earnings. Connecticut's implementation of AFDC followed "fill-the-gap budgeting" rules which set the proportional disregard at 51% for the 4 months following the first month of employment while on assistance and 27% afterwards (see

Bloom et al., 2002).[6] Accordingly, we refer to the 51% rate as the *unreduced* proportional disregard and to the $120 set aside as the *unreduced* fixed disregard.

A distinguishing feature of the JF program is its 100% earnings disregard which provides a dramatic reduction in the implicit tax on earnings faced by women on welfare relative to AFDC. The JF earnings disregard was meant to incentivize work but also created an eligibility "notch" in the transfer scheme, with a windfall loss of the entire grant amount occurring if a woman earned a dollar more than the monthly federal poverty line.[7]

Another important feature of JF, like all TANF programs, is the imposition of time limits. As documented in MDRC's final report (Bloom et al., 2002) and BGH (2006), this feature induced time varying effects of the program as some JF cases eventually became ineligible for program benefits. For this reason, we restrict our analysis to the first 7 quarters of the JF experiment, a horizon over which no case was in danger of reaching the limit and during which behavioral responses appear roughly stationary.[8] Finally, the JF program involved more stringent work requirements and imposed sanctions on cases that failed to seek work. This created further incentives to work while on welfare, as non-working women receiving benefits were effectively "hassled" into either getting off welfare or working.

A Monthly Budget Set

Figure 1 provides a stylized depiction of the monthly budget set faced by a long term welfare recipient with two children under the AFDC and JF policy rules respectively. This woman is assumed to have access only to the reduced fixed and proportional disregards. The vertical axis of the graph gives total income (earned income plus transfers) while the horizontal axis gives earned income E. $\overline{G}$ is the base grant amount which is common to JF and AFDC. The JF budget set exhibits a large discontinuity, a "notch", at the federal poverty line (FPL): at earnings below the FPL the woman receives a transfer equal to $\overline{G}$ while at earnings beyond FPL she is ineligible for assistance. The presence of the notch implies that total income when monthly earnings equal FPL exceeds that for earnings in the range $(FPL, FPL+\overline{G})$. The JF budget set is to be contrasted with the AFDC budget set which exhibits no discontinuities: the transfer phases out smoothly reflecting

[6] Fill the gap budgeting lowers the implicit AFDC tax rate on earnings in Connecticut by a factor of .73. For example in the first four months of employment the usual tax rate would be 2/3rds (as part of the 30 + 1/3rd policy) but in Connecticut it is $.73 \times \frac{2}{3} \times 100 = 49\%$. After that the usual tax rate would be 100% but in Connecticut it becomes 73%. One minus the tax rate is the disregard rate.

[7] The program also induced a second notch for the sample of JF *applicants* who faced a strict earnings test in order to *establish* eligibility. AFDC did not have an earnings test for applicants, but benefits for that program phased out at an amount above the JF earnings test. Hence, it became harder under JF for high earning applicants to establish eligibility. Because our analysis is static, we follow BGH in ignoring this feature of the policy. See Card and Hyslop (2005) for an empirical analysis of dynamic responses to eligibility incentives. We have experimented with restricting our analysis to recipients only and found similar results however this group contains substantially fewer high earners which limits our ability to detect opt-in behavior.

[8] If agents are forward looking, this restriction may not fully remove the influence of the time limits on behavior (Grogger and Michalopoulos, 2003). We return to this concern in Section 8.

an implicit tax rate of 73% on earnings above a \$90 disregard.

We formalize these policy rules by means of the transfer function $G_i^t(E)$ which gives the monthly grant amount associated with welfare participation at earnings level E under policy regime $t \in \{a, j\}$ (AFDC or JF respectively). The i subscript acknowledges that the grant amount may vary across women with the same earnings due to variation in the size of the assistance unit (AU) or based upon their recent assistance history.[9] The regime specific transfer functions are:

$$\begin{aligned} G_i^a(E) &= \max\left\{\overline{G}_i - 1[E > \delta_i](E - \delta_i)\tau_i, 0\right\} \\ G_i^j(E) &= 1[E \leq FPL_i]\,\overline{G}_i, \end{aligned} \tag{1}$$

where $\delta_i \in \{90, 120\}$ and $1-\tau_i \in \{.27, .51\}$ are the fixed and proportional AFDC earnings disregards, both of which may vary across women based upon their assistance history. The base grant amount $\overline{G}_i$ and the federal poverty line FPL_i also vary due to differences in AU size. Although in Figure 1 the AFDC transfer is fully exhausted at an earnings level that is strictly below the federal poverty level (that is, $G_i^a(FPL_i) = 0$), this is not always the case. In Connecticut, a woman with access to the *unreduced* proportional and fixed disregards exhausts her AFDC transfer at an earnings level slightly above the federal poverty level. Specifically, $0 < G_i^a(FPL_i) < \$75$ for any such woman.

Figure 1 illustrates five alternatives that a long term recipient may choose under AFDC: on assistance and zero earnings (point A), off assistance and zero earnings (A'), on assistance and positive earnings (point B), off assistance and earnings low enough to imply AFDC eligibility (point B'), and off assistance and earnings high enough to preclude AFDC eligibility (points C, D, and E). The next section shows that a conventional model of labor supply and program participation constrains how such a woman's choices under AFDC may be paired with choices under JF.

2 Optimization and Behavioral Responses

A Model

To study the effects of the JF experiment, we begin by considering a conventional static utility maximization framework where women have complete control over their earnings. We assume women have heterogeneous utility functions $U_i(E, C)$, defined over earnings and a consumption equivalent C. As in Saez (2010), utility is increasing in consumption C but decreasing in earnings E which require effort to generate. In this framework, one can think of heterogeneity in the disutility of earnings as capturing both variation in labor market skills and preferences for leisure.

The consumption equivalent C incorporates a variety of psychic and monetary costs and takes

[9] The assistance unit consists of the woman receiving welfare plus eligible dependent children. Children are eligible if they are under age 18 or under age 19 and in school.

the form:

$$C = E + \left(G_i^t(E) - \phi_i - \eta_i^t 1[E=0]\right) D - \mu_i 1[E>0]. \tag{2}$$

The variable $D \in \{0,1\}$ is an indicator for the woman participating in welfare. The parameter $\phi_i \geq 0$ gives the dollar value of welfare stigma (Moffitt, 1983), which may vary across women and explains why some eligible women don't participate in welfare. We also allow for a fixed cost of work μ_i which, for example, might capture the monthly cost of commuting to work. Fixed costs discourage work at low earnings levels and create the possibility that non-working women respond to marginal changes in work incentives by earning large amounts (Cogan, 1981). To capture the effects of work requirements, $\eta_i^t \geq 0$ gives the dollar value of the "hassle" a woman faces when not working and receiving benefits under regime t. In some cases these hassle costs may be pecuniary as women may be sanctioned if they fail to seek work. Because JF includes stronger work requirements and sanctions, we assume that $\eta_i^j \geq \eta_i^a \ \forall i$.

Finally, the woman's decision problem is to:

$$\max_{E \geq 0,\ D \in \{0,1\}} U_i(E, C) \text{ subject to } (1) \text{ and } (2).$$

We note in passing that the primitives of this simple model are given by the joint distribution of $\left(\phi_i, \eta_i^a, \eta_i^j, \mu_i, U_i(.)\right)$. Empirical modeling usually proceeds by placing parametric restrictions on this joint distribution so that its parameters are identified by the observed data. We depart from this standard practice by leaving this joint distribution unrestricted. Our interest centers not on the primitives themselves, but on the set of margins along which reform might induce women to adjust their behavior. This is a topic on which the theory speaks clearly, as we shall see next.

Response Margins

The above model places restrictions on how a woman's choice may vary with the policy regime. These restrictions follow from simple revealed preference arguments. The JF reform made working on welfare more attractive, it potentially made not working while on welfare less pleasant due to increased hassle, and it had no impact on the appeal of other choices (e.g. working while off welfare). Thus, since preferences are invariant to the policy regime, the model implies that women will not be induced to choose options made less attractive by reform, nor abandon options made more attractive by it. These restrictions are illustrated in Figures 2a-2c, which show the choices women may make under the JF regime given their choices under the AFDC regime.[10]

Consider first a woman who, under AFDC, chooses not to work or participate in welfare (point A' in Figure 2a). She must face high welfare stigma and/or hassle as she is willing to forgo the full grant amount $\overline{G}_i$ under AFDC. Not working and participating in JF will be at least as unattractive to such a woman as not working and participating in AFDC since the base grant

[10] A formal treatment of these arguments is provided in Section 6 and the Appendix.

amount is unaffected by the reform and JF may entail greater hassle. Thus, by revealed preference, such a woman will not choose point A under JF. She will also not choose to earn above the FPL under JF, as she could have done so under AFDC and chose not to. Likewise, she will not choose to earn below the FPL and be off assistance, as this option was available to her under AFDC. She may however choose to earn below the FPL on assistance under JF, as this option entails higher consumption than under AFDC. Hence, by revealed preference, the only ways such a woman may respond to reform are to work on welfare at earnings levels below the poverty line or to remain at point A' (no adjustment).

By contrast, a woman at point A in Figure 2a who, under AFDC, would participate in welfare without working, may be incentivized by the JF rules to adjust in several ways. First, she may be induced to work while on welfare both by the reduction in implicit tax rates on earnings and the increased hassle associated with JF. If fixed costs are large enough, the additional hassle associated with not working under JF may induce her to leave welfare and earn more than the federal poverty line. Alternatively, she may respond to the hassle by opting out of welfare and continuing to not work. Finally, she may not respond at all and continue to stay on welfare and not work. The sole restriction on her response is that she cannot choose to work off welfare at earnings levels below the FPL (e.g. point B'). This is because her choice under AFDC implies that her stigma is less than the base grant amount $\overline{G}_i$, which implies working on JF is preferable to working off welfare at all earnings levels below the FPL.

Corresponding arguments can be constructed for women making each of the other potential choices under AFDC, namely, points B'-E. A woman who works off welfare at point B' in Figure 2b under AFDC must dislike welfare participation but may be induced by reform to work on welfare under JF because doing so entails greater consumption than under AFDC. By revealed preference, she will not choose to participate in welfare without working since this option was available under AFDC. Nor will she be induced to work at high earnings levels off welfare since this was also available to her. Thus, her only possible avenue of adjustment is to work on welfare, which may entail an income effect leading to a reduction in her earnings.

A woman working on welfare at point B in Figure 2b under AFDC will face a reduction in her implicit tax rate under JF. Like any uncompensated increase in the wage, this change could lead to increases or decreases in the amount of work undertaken, but in any case will lead her to continue working on welfare. If her substitution effect dominates her income effect, the woman will be expected to work more but not more than the federal poverty line as this level of earnings was in her choice set under AFDC. Likewise, a woman working on welfare under AFDC at point C in Figure 2b may choose to participate in JF which would offer an increase in income for the same amount of work. This may result in a reduction in earnings due to income effects. But if the woman has high welfare stigma, she may choose to remain at point C when offered JF.

A woman choosing point D under AFDC in Figure 2c may choose to reduce her earnings and participate in the JF program, which would offer additional income for less work. However, with sufficiently high welfare stigma, she may nonetheless choose to remain off welfare and continue to earn at point D. A high earning woman who locates at point E in Figure 2c may be incentivized to opt-in to welfare under JF which would involve lowering her earnings to the federal poverty line. At this point she will unambiguously want to work more, as participation involves a reduction in income for her. She will not do so because of the large eligibility notch she faces. Thus, we have the sharp nonparametric prediction that opt-in should lead to a mass point of workers locating exactly at the poverty line.[11]

Testable Predictions

The above arguments catalogue how womens' choices may depend upon the policy regime they face. In practice, it is impossible to observe the same woman under two regimes at a given point in time.[12] Although we cannot directly observe the joint distribution of choices under both policy regimes, the Jobs First experiment allows us to infer the regime specific marginal distributions of outcomes. The key qualitative predictions of the model for these distributions are that: 1) the fraction of women who work should be greater under JF because the program cannot induce anyone to chose not to work and may induce some to work, 2) the fraction of women on welfare and not working should be lower under JF because the hassle feature of the program makes this choice less attractive, and 3) a mass of women should locate exactly at the FPL under JF because of the eligibility notch. We turn now to an examination of whether these predictions are satisfied in the data.

3 Data

Data Sources and Measures of AU Size

Our data come from the MDRC Jobs First Public Use Files. They contain a baseline survey of demographic and family composition variables merged with longitudinal administrative information on welfare participation, rounded welfare payments, family composition, and rounded earnings covered by the state unemployment insurance (UI) system for at least seven quarters prior to random assignment to at least 16 quarters post random assignment.

There are a number of important limitations to the Public Use Files that place constraints on our analysis. While welfare payments are measured monthly, UI earnings data are only available quarterly. To put them on a consistent time scale, we aggregate welfare participation to the

[11] Women who would locate at point D under AFDC may also choose to locate exactly at the poverty line under JF.

[12] Since preferences and constraints may change month to month, panel data will not help much in this regard without strong assumptions about how these factors evolve over time. We return to this issue in Section 8.

quarterly level.[13]

Another difficulty is that the administrative measure of AU size is missing for most cases. For the Jobs First sample we are able to infer an AU size in most months from the grant amount while the women are on welfare. However if AU size changes while off welfare we are not able to detect this change.[14] Moreover, in some cases the grant amount does not match any of the base grant amounts. This is presumably because the woman reported some unearned income or because of failure to comply with employment-related mandates.[15] In both of these situations, we use the grant amount in other months to impute AU size in a month for which we cannot observe it. For the AFDC sample, the grant amount depends on many unobserved factors, preventing us from inferring the AU size from the administrative data.

Thus, when computing treatment effects by AU size, we rely on a variable collected in the baseline survey named "kidcount." This variable records the number of children in the household at the time of random assignment and is top-coded at three children. Appendix Table 2 gives a cross-tabulation, in the JF sample, of kidcount with our more reliable AU size measure inferred from grant amounts. The tabulation suggests the kidcount variable is a reasonably accurate measure of AU size over the first 7 quarters post-random assignment conditional on the number of children at baseline being less than three. As might be expected, the kidcount variable tends to underestimate the true AU size as women may have additional children over the 7 quarters following the baseline survey. To deal with this problem we try inflating the kidcount based AU size by one in order to avoid understating the location of the poverty line for most households.[16]

Baseline Characteristics of the Analysis Sample

Table 1 shows descriptive statistics for our analysis sample. We have 4,642 cases with complete pre-random assignment characteristics and nonmissing values of the kidcount variable.[17] There are some

[13] The public use files do not report the month of randomization. However, we were able to infer it by contrasting monthly assistance payments with an MDRC constructed variable providing quarterly assistance payments. For each case, we found that a unique month of randomization leads the aggregation of the monthly payments to match the quarterly measure to within rounding error.

[14] Changes in AU size are typically due to a birth or to the fact that a child reaches age 18 (or 19 if enrolled in school) hence becomes categorically ineligible for welfare. Under AFDC, the AU size also changes when the adult is removed from the unit due to sanctions for failure to comply with employment-related mandates. Empirically this source of time variation in AU size seems quantitatively minor. Bloom et al. (2002) report that 5 percent of AFDC group members had their benefits reduced owing to a sanction within four years after random assignment.

[15] Under JF a recipient's cash grant was reduced by 20 percent for three months in response to the first instance of noncompliance and by 35 percent for three months in response to the second instance. A third instance resulted in cancellation of the entire grant for three months.

[16] That is, we use the following mapping from kidcount to AU size: $0 \to 3$, $1 \to 3$, $2 \to 4$, $3 \to 5$, which maps each kidcount value to the modal inferred AU size in Appendix Table 2 plus one. This mapping is conservative in ensuring that earnings levels below the FPL are indeed below it. We have found that our results are robust to alternate codings including inflating the AU size by two and not inflating it at all.

[17] This yields essentially the same baseline sample as in BGH (2006). Relative to their analysis, we impose the additional restriction that the kidcount variable be non-missing. We also drop one AFDC case from our analysis with unrealistically high quarterly earnings that sometimes led to erratic results.

mildly significant differences between the AFDC and JF groups in their baseline characteristics, however these differences are not jointly significant. We follow BGH (2006) in using propensity score reweighting to adjust for these baseline differences in order to increase the efficiency of our estimation procedures. These techniques are described in the Appendix. After adjustment, the means of the AFDC and JF groups are very similar.

We also examine three subgroups studied by BGH (2010). These groups are characterized by their earnings seven quarters prior to random assignment, with a "zero earnings" group having no earnings in pre-assignment quarter 7, and a "low" and a "high" earning group having quarterly earnings below or above the median of the non-zero observations respectively.[18] Descriptive statistics for these groups are also provided in Table 1. Because pre-assignment earnings proxy for tastes and earnings ability, the JF reform likely presented these groups with different incentives, which makes them useful for exploring treatment effect heterogeneity.

4 Experimental Impacts, Bunching, and Under-reporting

Experimental Impacts on Earnings and Participation

We turn now to examining the testable predictions of the model introduced in Section 2. Figure 3a provides reweighted empirical distribution functions (EDFs) of earnings in the AFDC and JF samples using quarterly earnings data for the seven quarters following random assignment. We rescale earnings relative to three times the monthly FPLs faced by the sample women: 3FPL is the maximum amount that a woman can earn in a quarter while maintaining welfare eligibility throughout the quarter. We base the FPL on the survey measure of AU size. To avoid understating the location of the poverty line these women face, we inflate the implied AU size by one.

Several of the theoretical predictions can be assessed from this figure alone. A reweighted Kolmogorov-Smirnov test strongly rejects the null hypothesis that the two EDFs are identical.[19] Clearly more quarters exhibit positive earnings in the JF sample than in the AFDC sample, indicating that JF successfully incentivized many women to work. The earnings EDF rises more quickly in the JF sample than under AFDC, signaling excess mass at low earnings levels. Also, the EDFs cross below the notch leading the fraction earning less than $3FPL$ to be slightly greater for the JF sample than among the AFDC controls. A large increase in the fraction earning less than $3FPL$ would be suggestive evidence of an opt-in response, however the impact here is small and statistically insignificant. Notably, there is no evidence of a spike in the distribution of earnings at $3FPL$, an issue we explore more carefully below using a different measure of AU size.

These distributional effects conceal substantial heterogeneity across subgroups. Figures 3b-3d provide corresponding EDFs in three subsamples defined by their earnings in the seventh quarter

[18] The seventh quarter prior to random assignment is a useful stratifying variable because welfare applicants and recipients generally experience a severe dip in earnings in the quarters immediately prior to random assignment.

[19] See the Appendix for details on the computation of this test.

prior to random assignment. These groups are of interest because pre-random assignment earnings are a strong predictor of post-random assignment earnings, hence they proxy for the relevant range of the budget set an agent would face under AFDC. Accordingly, units with high pre-random assignment earnings should be most likely to exhibit an opt-in effect while units with zero earning should be more likely to be pushed into the labor force by JF. The figures confirm that the expected pattern of heterogeneity is in fact present, with the high earnings group experiencing no impact on the fraction of cases working but a large (though marginally significant) impact on the fraction of cases with earnings less than or equal to three times the monthly poverty line. The zero earnings group, by contrast, exhibits a large impact on the fraction of cases working, but essentially no impact on the fraction of cases with earnings less than or equal to three times the monthly poverty line.

Table 2 quantifies the impacts of JF on the earnings distribution and provides standard errors generally confirming the visual impression of the prior figures. JF yielded a large decrease in the fraction of quarters with earnings below the monthly poverty line, suggesting many women were incentivized to work greater amounts than before by the reduction in implicit tax rates. We also see a tendency for the fraction earning less than three times the poverty line to increase, especially in the higher earning groups, suggesting the possible presence of an opt-in effect. Consistent with the model's other predictions, welfare participation grew sharply under JF in each subgroup and the fraction of quarters spent on welfare with no earnings fell. We also see that average earnings while on welfare grew, which the theory would explain as the result of opt-in to welfare by high earners and substitution effects by working welfare recipients.

Bunching and Under-reporting Behavior

Our evidence so far largely corroborates the predictions of the simple model of Section 2, with the distributional impacts varying across subgroups in the expected manner. However, the model predicts that bunching should be present at the JF eligibility notch, which seems not to be the case in the EDFs of the previous section. We can examine this prediction more carefully by looking for bunching in the JF sample, where we can accurately measure AU size.[20]

Figure 4a provides a histogram of earned income rescaled relative to the federal poverty line. Not only do we fail to detect a spike in the mass of observations located at the notch, the earnings density actually appears to be declining through this point. Moreover, this decline is relatively smooth through the notch which should bound, to its right, a dominated earnings region. Compared to women not on welfare in the quarter (Figure 4c), there is arguably an excess "mound" in the density of earnings below the notch for women on welfare throughout the quarter (Figure 4b).

One explanation for these findings is that women cannot fully control their earnings (Altonji

[20] Our measure of AU size is inferred from the grant amount upon the assumption that size has not changed since the last time the woman was on welfare and had an unreduced grant amount.

and Paxson, 1988; Chetty et al, 2011; Dickens and Lundberg, 1993). Such constraints were ignored in the idealized model of Section 2, but are potentially important in understanding labor supply responses .

We also suspect that under-reporting behavior is important for the patterns in Figure 4b, which shows that, conditional on assistance, earnings stretch well beyond the FPL. Interpreting this as evidence of under-reporting behavior is consistent with the results of Hotz, Mullin, and Scholz (2003), who analyzed data from a welfare reform experiment in California. Comparing administrative earnings records from the California Unemployment Insurance system with earnings reported to welfare, they find a highly nonlinear relationship, with cases having quarterly UI earnings above $2,500 being disproportionately more likely to under-report and under-reporting greater amounts on average. This threshold is close to the eligibility threshold of the California program.

There is an alternative explanation for both the lack of bunching and "excessive" earnings by women on assistance: imperfect enforcement of the JF eligibility threshold. Unfortunately, we lack data on earnings reported to the welfare agency, which prohibits us from directly examining whether lapses in enforcement took place. Appendix Figure 1 provides some indirect evidence that enforcement was extensive. We plot the empirical relationship between quarterly UI earnings and welfare payments for women in the JF sample. Because the poverty line varies by AU size, we rescale earnings of each case relative to the relevant poverty line.[21] Appendix Figure 1 shows that the fraction of women on assistance throughout the quarter falls quickly as *actual* earnings exceed the eligibility notch, suggesting that women with *reported* earnings above the notch were likely denied benefits. This drop is equally pronounced in several different measures of the duration of welfare participation which suggests lags between earnings and eligibility decisions are not particularly important. We use this evidence to justify the assumption, maintained in the model introduced next, that women with earnings above the relevant eligibility threshold must under-reporting their earnings in order to receive cash assistance.

5 An Augmented Model

Thus far we have found that most of the qualitative predictions of the frictionless static model regarding earnings and program participation appear to hold in the data but that predictions regarding a mass point at the eligibility notch fail. We also found that many seemingly ineligible households receive welfare benefits. We consider now a generalized version of the standard labor supply model of Section 2 that can accommodate these observations by allowing constraints on worker choices and under-reporting opportunities. Based on this model, we derive in the next section an approach to identification of the magnitude of the intensive margin opt-in effect, the

[21] The close correspondence at low earnings levels between the median monthly grant (which includes cases not on welfare) and the statutory unreduced grant amount indicates that grants are rarely reduced, which is comforting given our use of the grant amount to infer the size of the AU.

extensive margin decision to work, and various sorts of under-reporting responses.

We begin by assuming that, in addition to the welfare participation decision (D) and the choice of earned income (E), women may also choose a level of earnings (E^r) to report to the welfare agency. We assume that women can under-report but not over-report earnings so that $E^r \leq E$.[22] Grant amounts are determined based upon reported earnings, so that the transfer function becomes:

$$\begin{aligned} G_i^a (E^r) &= \max\left\{\overline{G}_i - 1\left[E^r > \delta_i\right](E^r - \delta_i)\,\tau_i, 0\right\} \\ G_i^j (E^r) &= 1\left[E^r \leq FPL_i\right]\overline{G}_i. \end{aligned} \tag{3}$$

The costs of under-reporting earnings are given by $\kappa_i > 0$ which may vary arbitrarily across women. One can think of κ_i as the expected psychic and pecuniary costs of concealing a job from the welfare agency.[23] Like stigma and hassle, under-reporting costs are assumed to enter as dollar equivalents in consumption. Thus, we augment our earlier specification of the consumption equivalent to be:

$$C = E + \left(G_i^t (E^r) - \phi_i - \eta_i^t 1\left[E^r = 0\right] - \kappa_i 1\left[E^r < E\right]\right) D - \mu_i 1\left[E > 0\right]. \tag{4}$$

Note that women face hassle when they report zero earnings regardless of whether or not their actual earnings are zero. Hence, a woman concealing her earnings will choose to report positive earnings equal to or below the fixed disregard level δ_i under AFDC or below FPL_i under JF.

We make one additional assumption on psychic costs relative to Section 2 which will greatly simplify our analysis in the next section: we assume $\phi_i > G_i^a (FPL_i)$. This assumption guarantees that women will not choose to report earnings above the federal poverty line while on AFDC.[24]

To model constraints, we suppose that each woman draws a pair of earnings offers $\left(O_i^1, O_i^2\right)$ from an unrestricted bivariate distribution $F_i(.)$ with support on the (strictly) positive orthant. She can choose between these offers or reject them both, in which case she earns nothing. The offers drawn are invariant to the policy regime t to which the woman is assigned. Thus, the woman's

[22] Allowing over-reporting behavior would essentially nullify the JF work requirements. In practice, concocting a fictitious job would have been difficult as employment had to be verified by case workers.

[23] See Saez (2010) for a related analysis involving a fixed "moral" cost of mis-reporting income to tax authorities.

[24] This restriction is useful for two reasons. First, it allows us to identify women on assistance with earnings above the poverty line as under-reporters regardless of the policy regime (AFDC or JF). Second, it allows us to rule out the possibility that women are induced by JF to stop working. This would happen if the only earnings offers received were above the poverty line, in which case working on welfare would not be an option under JF. To check the plausibility of this restriction, we examined the fraction of women in the AFDC sample who participated in welfare while receiving a rounded benefit of \$50. We found this fraction to be approximately 1% which corroborates the notion that most women in our sample do not find it worthwhile to participate in welfare in exchange for very low benefit levels. Additionally, as pointed out above, the bound on stigma is not very stringent: at worst it requires a monthly stigma of \$75 but most often a strictly positive stigma suffices since many of the women in our sample are long term participants likely to be ineligible for the *unreduced* proportional disregard.

objective is to:

$$\max_{E\in\left\{0,O_i^1,O_i^2\right\},\ D\in\{0,1\},\ E^r\in[0,E]} U_i\left(E,C\right) \text{ subject to } (3) \text{ and } (4)$$

Note that the presence of two offers provides the possibility of an intensive margin earnings response to the policy change. This response may or may not be constrained as the offers $\left(O_i^1,O_i^2\right)$ could coincide with the woman's unconstrained choices of E_i under the two policy regimes.[25] The presence of constraints provides a simple explanation for the absence of a spike in the earnings distribution at the JF eligibility notch. Likewise, our allowance for under-reporting behavior provides an explanation for welfare participation among households with UI earnings above the eligibility notch.

The introduction of under-reporting behavior introduces new margins of adjustment not present in the model of Section 2. Figure 5 illustrates the decision problem in earnings and consumption equivalent space for two women with substantial fixed costs of work and low costs of under-reporting. The effective budget sets are discontinuous at zero earnings due to the fixed costs of work μ and hassle costs. As depicted, the hassle costs η^j of not working under JF are larger than those under AFDC represented by η^a, but both are smaller than μ. In comparison with the fixed costs of work and hassle, the costs of under-reporting (κ) are depicted as being relatively small. The under-reporting line is equivalent under AFDC and JF because in either case the woman can report earnings arbitrarily close to zero, obviating any implicit taxes on her true earnings.

A woman with the configuration of psychic costs and preferences found in Figure 5a would work on welfare under AFDC but under-report her earnings (point A). However, under JF, she would truthfully report her earnings (point B), as the JF disregard reduces the return to under-reporting. Hence, reform may induce a reduction in under-reporting.

By contrast, Figure 5b shows a scenario where the hassle effects of JF are larger, the costs of under-reporting are smaller, and preferences over earnings are such that the disutility of work is lower. This woman would receive benefits without working (point A) under AFDC but, under JF, will choose to earn above the poverty line and under-report her earnings (point B) in order to maintain eligibility. This occurs because the JF work requirements remove point A from her budget set – such a woman has effectively been hassled off welfare into under-reporting. Thus, the JF reform may have mixed effects on reporting behavior which can lead to an increase or a decrease in the total rate of under-reporting. In the next section we seek to formally identify the magnitude of these and other responses.

[25] An alternate approach, which would achieve the same result, is to treat the offers as independent draws from an unknown continuous univariate distribution, and to let the number of offers vary across women in an unrestricted fashion. This too would nest the unconstrained case, as the number of offers may approach infinity for each woman.

6 Identification and Estimation of Response Margins

Our augmented model is sufficiently general that it places no restrictions on the cross-sectional distribution of earnings and program participation choices under a given policy regime: there is a mix of preferences and earnings opportunities that can support any earnings and program participation choice. Hence, the right mix of preferences and offers across women can support any cross-sectional distribution of choices. However, as we have already discussed, the model does restrict how different groups of women may respond to policy variation. Specifically, it rules out certain combinations of earnings and program participation choices under the two policy regimes.

In this section, we enumerate the restrictions that the augmented model places on how a woman's earnings range and welfare participation may vary across policy regimes. Our focus on coarse earnings categories follows from the observation that the JF reform made a broad range of earnings choices (those below the FPL) more attractive conditional on welfare participation, potentially reduced the attractiveness of not working at all while on welfare, and had no effect on the return to working at earnings levels above the FPL.[26] Because we have not parametrically structured utility, we cannot quantify exactly how much more attractive each choice within these broad ranges has become. However, we can deduce whether women will make different broad earnings and program participation choices in response to the JF reform. We use these restrictions to test the model and form bounds on the fraction of women who respond to reform along each of the allowable margins of adjustment.

Response margins

We begin by introducing some notation. Our coarsened earnings variable $\widetilde{E}_i$ is defined by the relation

$$\widetilde{E}_i \equiv \begin{cases} 0 & \text{if } E_i = 0 \\ 1 & \text{if } E_i \leq FPL_i \\ 2 & \text{if } E_i > FPL_i \end{cases}.$$

That is, $\widetilde{E}_i$ indicates whether a woman works, and if so, whether she earns enough to be ineligible for benefits under JF. This choice of earnings categories is crucial as the model rules out many responses to reform involving changes across (but not within) these categories. Moreover, our assumptions so far imply that, under either policy regime, a woman with $\widetilde{E}_i = 2$ who participates in welfare must be under-reporting her earnings to the welfare agency.

[26] A coarsening of the choice set is common practice in the structural labor supply literature (e.g. Hoynes, 1996; Keane and Moffitt, 1998; Blundell et al, 2013). Typically, the coarsening adopted in these works rests on data-driven categories such as part-time and full-time work. Our approach is fundamentally different. We do not assume that agents lack the ability to choose earnings levels within earnings categories – constraints in our framework are summarized by the earnings offer distribution $F_i(.)$ which may be continuous.

Pairing these earning categories with the decision to participate in welfare and the under-reporting decision yields seven earning / participation / reporting choices allowed by the model, which we henceforth refer to as *states*. The set of possible states is: $\mathcal{S} \equiv \{0n, 1n, 2n, 0r, 1r, 1u, 2u\}$. The number associated with each state refers to the woman's earnings category while the letter describes her combined welfare participation and reporting decisions. Specifically, the letter n denotes welfare non-participation, r denotes participating in welfare while truthfully reporting earnings, and u denotes participating in welfare while under-reporting earnings. The state $0u$ is ruled out, as it is not meaningful to "under-report" zero earnings, and the state $2r$ is not allowed by either the JF eligibility rules or, given the lower bound imposed on stigma, the AFDC eligibility rules.

Let S_i^a denote woman i's potential state under AFDC and S_i^j her potential state under JF. Let $\pi_{s^a,s^j} \equiv P\left(S_i^j = s^j | S_i^a = s^a\right)$ be the conditional probability of occupying state $s^j \in \mathcal{S}$ under JF given the choice of state $s^a \in \mathcal{S}$ under AFDC. These probabilities constitute our parameters of interest as they summarize the frequency of adjustment along the various margins through which agents can respond to the JF reform.[27] We show in the Appendix that the augmented model of Section 5 implies the 7×7 matrix Π of response probabilities takes the form:

State under AFDC	Earnings / Reporting State under JF						
	$0n$	$1n$	$2n$	$0r$	$1r$	$1u$	$2u$
$0n$	$1-\pi_{0n,1r}$	0	0	0	$\pi_{0n,1r}$	0	0
$1n$	0	$1-\pi_{1n,1r}$	0	0	$\pi_{1n,1r}$	0	0
$2n$	0	0	$1-\pi_{2n,1r}$	0	$\pi_{2n,1r}$	0	0
$0r$	$\pi_{0r,0n}$	0	$\pi_{0r,2n}$	$1-\pi_{0r,0n}-\pi_{0r,2n}$ $-\pi_{0r,1r}-\pi_{0r,2u}$	$\pi_{0r,1r}$	0	$\pi_{0r,2u}$
$1r$	0	0	0	0	1	0	0
$1u$	0	0	0	0	1	0	0
$2u$	0	0	0	0	$\pi_{2u,1r}$	0	$1-\pi_{2u,1r}$

The zero entries in the matrix reflect the many responses that the model says cannot occur. No woman will choose state $1u$ under JF because under-reporting is costly and earnings below the poverty line are fully disregarded under JF. Thus, all the entries in the column corresponding to state $1u$ under JF are restricted to zero. The remaining zeros stem from revealed preference arguments. First, the JF reform makes the value of choices in the set $\mathcal{C}^1 \equiv \{0n, 1n, 2n, 1r, 1u, 2u\}$ no lower and the value of the choices in the set $\mathcal{C}^2 \equiv \{0n, 1n, 2n, 0r, 1u, 2u\}$ no higher. Therefore, by revealed preference, no woman will pair any of the choices in $\mathcal{C}^1$ under AFDC with a (different) choice in $\mathcal{C}^2$ under JF. This explains most of the zero restrictions. Finally, note (as explained

[27]Note that these probabilities are functionals of the joint distribution of the model primitives $\left(\phi_i, \eta_i^a, \eta_i^j, \mu_i, \kappa_i, U_i(.), F_i(.)\right)$.

in Section 2) that the choice of $0r$ under AFDC implies stigma is below the base grant amount, which in turn implies that choice $1n$ is dominated by the choice $1r$ under JF. It follows then that $\pi_{0r,1n} = 0$, which is the last zero restriction to be accounted for in the matrix.

The non-zero elements of the matrix reveal that the non-deterministic responses are driven by eight response margins – that is, eight ways in which a woman might respond to the JF experiment by changing her behavior.[28] Note that each of the allowable responses has already been discussed in Sections 2 or 5. Reading the above matrix from the top left element to the bottom right element, these margins are: 1) a fraction $\pi_{0n,1r}$ of women who would not participate in welfare or work under AFDC will take up welfare and work under JF, 2) a fraction $\pi_{1n,1r}$ of women who would work but earn less than the poverty line and not participate in welfare under AFDC will work while on welfare under JF, 3) a fraction $\pi_{2n,1r}$ of women who would earn more than the poverty line and be off assistance under AFDC will "opt-in" to welfare and reduce their earnings below the poverty line under JF, 4) a fraction $\pi_{0r,0n}$ of women who participate in welfare without working under AFDC will leave welfare and continue not to work, 5) a fraction $\pi_{0r,2n}$ of women who participate in welfare without working under AFDC will leave welfare and earn above the poverty line, 6) a fraction $\pi_{0r,1r}$ of women who participate in welfare without working under AFDC will take up work and remain on welfare under JF, 7) a fraction $\pi_{0r,2u}$ of women who participate in welfare without working under AFDC will earn above the poverty line but remain on welfare by under-reporting earnings, and 8) a fraction $\pi_{2u,1r}$ of women who earn above the poverty line but under-report earnings in order to qualify for benefits under AFDC will reduce their earnings below the poverty line under JF and truthfully report earnings.

Note that the response probabilities $(\pi_{0r,2n}, \pi_{0r,2u})$ involve moving from earnings category 0 to category 2, while the probabilities $(\pi_{2n,1r}, \pi_{2u,1r})$ involve moving from earnings category 2 to category 1. Thus, the model allows for rank reversals in earnings and therefore violates the standard rank-invariance condition. Of course, a variety of other adjustments may occur within each of our three earnings ranges. Thus the restriction that $\pi_{1r,1r} = 1$ should not be taken to imply that no response is present among women who work on welfare, as many such women may adjust their earnings. Without further restrictions, however, we cannot infer the magnitude of any such adjustments.

Observable States

Our data do not allow us to measure reporting decisions other than by contrasting a woman's administrative earnings with the eligible maximum. Hence, states $1u$ and $1r$ are not empirically distinguishable. Accordingly, we define a function $g: \mathcal{S} \to \widetilde{\mathcal{S}}$ that reduces the "latent" states $\mathcal{S}$ to

[28]There is also one deterministic response: a woman choosing state $1u$ under AFDC must choose state $1r$ under JF with probability one.

"observable" states $\widetilde{\mathcal{S}}$ that can be measured in our data. Formally,

$$g(s) \equiv \begin{cases} s & \text{if } s \in \{0n, 1n, 2n\} \\ 0p & \text{if } s = 0r \\ 1p & \text{if } s \in \{1u, 1r\} \\ 2p & \text{if } s = 2u \end{cases}.$$

As before, the number of each state refers to the woman's earnings category and the letter n refers to welfare non-participation. The letter p denotes welfare participation, which is directly observable. Note that state $2p$ can only be occupied via under-reporting.

Let $\widetilde{S}_i^t$ denote the observed potential state of a woman whose latent potential state under policy regime t is S_i^t, that is, $\widetilde{S}_i^t \equiv g\left(S_i^t\right)$ for $t \in \{a, j\}$. Also, define the probability of occupying state $\widetilde{s}$ under policy regime t as

$$p_{\widetilde{s}}^t \equiv P\left(\widetilde{S}_i^t = \widetilde{s}\right) = \sum_{s:\, \widetilde{s}=g(s)} P\left(S_i^t = s\right). \tag{5}$$

Let the 6×6 matrix $\widetilde{\Pi}$ be composed of response probabilities of the form $P\left(\widetilde{S}_i^j = \widetilde{s}^j | \widetilde{S}_i^a = \widetilde{s}^a\right)$. From (5), $\widetilde{\Pi}$ may be written:

State under AFDC	Earnings / Participation State under JF					
	$0n$	$1n$	$2n$	$0p$	$1p$	$2p$
$0n$	$1-\pi_{0n,1r}$	0	0	0	$\pi_{0n,1r}$	0
$1n$	0	$1-\pi_{1n,1r}$	0	0	$\pi_{1n,1r}$	0
$2n$	0	0	$1-\pi_{2n,1r}$	0	$\pi_{2n,1r}$	0
$0p$	$\pi_{0r,0n}$	0	$\pi_{0r,2n}$	$1-\pi_{0r,0n}-\pi_{0r,2n}$ $-\pi_{0r,1r}-\pi_{0r,2u}$	$\pi_{0r,1r}$	$\pi_{0r,2u}$
$1p$	0	0	0	0	1	0
$2p$	0	0	0	0	$\pi_{2u,1r}$	$1-\pi_{2u,1r}$

.

Identification of Response Probabilities

We are now ready to discuss identification of the eight response probabilities appearing in the matrix $\widetilde{\Pi}$ based on the distribution of observed states in the AFDC and JF populations. Let $T_i \in \{a, j\}$ denote the policy regime into which a woman is randomized by the JF evaluation. Random assignment ensures that her potential (latent) states are independent of treatment. Formally,

$$T_i \perp \left(S_i^a, S_i^j\right), \tag{6}$$

where the symbol $\perp$ denotes independence. By (6) and (5), the marginal probabilities $p_{\widetilde{s}}^t$ are identified by the relation $p_{\widetilde{s}}^t = P\left(\widetilde{S}_i^t = \widetilde{s} | T_i = t\right)$. This is the well-known result that experimental variation identifies the marginal distributions of potential outcomes (here called states).

Define the vectors of observable state probabilities $\mathbf{p}^j \equiv \left[p^j_{0n}, p^j_{1n}, p^j_{2n}, p^j_{0p}, p^j_{1p}, p^j_{2p}\right]'$ and $\mathbf{p}^a \equiv \left[p^a_{0n}, p^a_{1n}, p^a_{2n}, p^a_{0p}, p^a_{1p}, p^a_{2p}\right]'$. By the law of total probability $\mathbf{p}^j = \widetilde{\mathbf{\Pi}}\mathbf{p}^{\mathrm{a}}$. This system consists of six equations, one of which is redundant given that state probabilities sum to one in each policy regime. The five non-redundant equations can be given an intuitive representation as:

$$
\begin{aligned}
p^j_{0n} - p^a_{0n} &= -p^a_{0n}\pi_{0n,1r} + p^a_{0p}\pi_{0r,0n} \\
p^j_{1n} - p^a_{1n} &= -p^a_{1n}\pi_{1n,1r} \\
p^j_{2n} - p^a_{2n} &= -p^a_{2n}\pi_{2n,1r} + p^a_{0p}\pi_{0r,2n} \\
p^j_{0p} - p^a_{0p} &= -p^a_{0p}\left(\pi_{0r,1r} + \pi_{0r,2u} + \pi_{0r,2n} + \pi_{0r,0n}\right) \\
p^j_{2p} - p^a_{2p} &= p^a_{0p}\pi_{0r,2u} - p^a_{2p}\pi_{2u,1r}
\end{aligned}
\tag{7}
$$

The left hand side of (7) catalogues the experimental impacts of the JF reform on the observable state probabilities. The right hand side rationalizes these impacts in terms of flows into and out of each state as allowed by the labor supply model. The identifying power of the theory derives from the fact that only a handful of response probabilities appear in each equation. Despite these restrictions, the system in (7) is clearly under-determined, with eight unknown response probabilities and only five equations. Still, it is immediate to see that the second equation of (7) uniquely identifies the response probability $\pi_{1n,1r}$. The remaining 4 equations constrain (without uniquely determining) the remaining seven response probabilities.

Some of the restrictions embedded in (7) are testable. First, the model implies that the experiment cannot generate an increase in the frequency of states $1n$ or $0p$. Second, the increase in the proportion of women choosing state $1p$ must be at least as large as the decrease in the fraction of women choosing state $1n$. Formally, the testable restrictions are:

$$
p^j_{0p} - p^a_{0p} \leq 0, p^j_{1n} - p^a_{1n} \leq 0, p^j_{1p} - p^a_{1p} \geq p^a_{1n} - p^j_{1n}. \tag{8}
$$

Violation of any of these conditions would imply that our framework failed to allow a response actually present in the data.

Subject to the restrictions in (8) holding, we can use the system in (7) to bound the seven remaining response probabilities. The upper and lower bounds on each of the response probabilities can be represented as the solution to a pair of linear programming problems of the form

$$
\max_{\pi} \lambda'\pi \text{ subject to (7) and } \pi \in [0,1]^7, \tag{9}
$$

where $\pi \equiv [\pi_{0n,1r}, \pi_{0r,0n}, \pi_{2n,1r}, \pi_{0r,2n}, \pi_{0r,1r}, \pi_{0r,2u}, \pi_{2u,1r}]'$. For example, solving the above problem for $\lambda = [0,0,0,0,0,0,1]'$ yields the upper bound on $\pi_{2u,1r}$, while choosing $\lambda = [0,0,0,0,0,0,-1]'$ yields the lower bound.

We can also use this representation to derive bounds on linear combinations of the response probabilities. We consider four "composite" margins of adjustment: $\pi_{0r,n} \equiv \pi_{0r,0n} + \pi_{0r,2n}$, $\pi_{p,n} \equiv \frac{p^a_{0p}}{p^a_{2p}+p^a_{1p}+p^a_{0p}} \left(\pi_{0r,2n} + \pi_{0r,0n}\right)$, $\pi_{n,p} \equiv \frac{p^a_{2n}\pi_{2n,1r}+p^a_{1n}\pi_{1n,1r}+p^a_{0n}\pi_{0n,1r}}{p^a_{2n}+p^a_{1n}+p^a_{0n}}$, and $\pi_{0,1+} \equiv \frac{p^a_{0p}(\pi_{0r,1r}+\pi_{0r,2n}+\pi_{0r,2u})+p^a_{0n}\pi_{0n,1r}}{p^a_{0p}+p^a_{0n}}$. The parameter $\pi_{0r,n}$ gives the fraction of women who would claim benefits without working under AFDC that are induced to get off welfare under JF. Upper and lower bounds for this response probability can be had by solving (9) with $\lambda = [0,1,0,1,0,0,0]$ and $[0,-1,0,-1,0,0,0]$ respectively. We also examine the fraction $\pi_{p,n}$ of all women who would participate in welfare under AFDC that are induced to leave welfare under JF, the fraction $\pi_{n,p}$ of women who are induced to take up welfare under JF, and the fraction $\pi_{0,1+}$ who are induced by JF to work. Because no women who would work under AFDC will choose not to work under JF, this last fraction is point identified by the proportional reduction in the fraction of women not working under JF relative to AFDC.

It is useful to construct analytic expressions for the bounds as a function of the regime-specific marginal distributions entering the constraints in (9). We accomplished this by solving the relevant linear programming problems by hand (a straightforward though cumbersome process). The resulting expressions are listed in the Appendix. An example is given by the bounds on the opt-in probability $\pi_{2n,1r}$ which take the form:

$$\max\left\{0, \frac{p^a_{2n}-p^j_{2n}}{p^a_{2n}}\right\} \leq \pi_{2n,1r} \leq \min\left\{\begin{array}{c} 1, \\ \frac{p^a_{2n}-p^j_{2n}+p^a_{0p}-p^j_{0p}}{p^a_{2n}}, \\ \frac{p^a_{2n}-p^j_{2n}+p^a_{0p}-p^j_{0p}+p^a_{0n}-p^j_{0n}}{p^a_{2n}}, \\ \frac{p^a_{2n}-p^j_{2n}+p^a_{0p}-p^j_{0p}+p^a_{2p}-p^j_{2p}}{p^a_{2n}}, \\ \frac{p^a_{2n}-p^j_{2n}+p^a_{0p}-p^j_{0p}+p^a_{0n}-p^j_{0n}+p^a_{2p}-p^j_{2p}}{p^a_{2n}} \end{array}\right\}.$$

Note that there are two possible solutions for the lower bound, one of which is zero. This turns out to be a generic feature of the lower bounds for each of the seven response probabilities. Which solution will be relevant is unknown a priori since the population vectors $\left(\mathbf{p}^a, \mathbf{p}^j\right)$ are unknown. The upper bound on $\pi_{2n,1r}$ admits five possible solutions. Other response probabilities can have fewer or more solutions.[29]

Estimation and Inference

Consistent estimators of the upper and lower bounds of interest can be had by using sample analogs of the marginal probabilities and computing the relevant $\min\{.\}$ and $\max\{.\}$ expressions. Inference is complicated by the fact that the limit distribution of the upper and lower bounds depends upon uncertainty in which of the constraints in (9) bind – i.e. in which of the bound solutions is relevant.

[29] The bounds for each parameter are functions of $\left(\mathbf{p}^a, \mathbf{p}^j\right)$, which leads to interesting patterns of dependence among them. For instance, among each pair of response probabilities ($\pi_{2n,1r}$,$\pi_{0r,2n}$), ($\pi_{0n,1r}$,$\pi_{0r,0n}$), and ($\pi_{2u,1r}$,$\pi_{0r,2u}$), only one probability may have an informative lower bound.

Naive bootstrap inference on the empirical $\min\{.\}$ and $\max\{.\}$ of the sample analogues of the bound solutions will fail to provide coverage of the parameters in question with fixed probability (Andrews and Han, 2009).

We report confidence intervals for the response probabilities based upon two inference procedures. The first simply ignores the uncertainty in which constraints bind – that is, it assumes the bound solution that appears relevant given the sample analogues is the only possible solution. In such a case, results from Imbens and Manski (2002) imply a 95% confidence interval for the parameter in question can be constructed by extending the upper and lower bounds by $1.65\widehat{\sigma}$ where $\widehat{\sigma}$ is a standard bootstrap estimate of the standard error of the sample moment used to define the relevant bound.[30] The second approach is a conservative bootstrap procedure described in the Appendix which covers the parameter with asymptotic probability greater than or equal to 95% regardless of which constraints bind. The lower limit of this confidence interval coincides with that of the naive procedure because sampling uncertainty only affects one of the bound solutions in the $\max\{.\}$ operator. However, the upper limit of the confidence interval from our conservative procedure generally exceeds that from the naive procedure, often by a substantial amount.

7 Results

Table 3 reports the estimated probabilities of occupying the six observable earnings and welfare participation states under each policy regime. Notably, the sign restrictions in (8) are satisfied by the point estimates. There is a small but statistically significant increase in the fraction of quarters on welfare with earnings above the quarterly poverty line indicating that, on net, JF induced more women to under-report earnings than it induced to truthfully report them.

Table 4 provides estimates of the response probabilities that rationalize the impacts in Table 3. The point identified response probability $\pi_{1n,1r}$ is computed by plugging in its sample analogue $\frac{\hat{p}^a_{1n}-\hat{p}^j_{1n}}{\hat{p}^a_{1n}}$. JF has a strong effect on entry into the program by the working poor. The bootstrap confidence intervals suggest between 31% and 46% of the women who would have worked off welfare under AFDC at earnings levels below the poverty line were induced to participate in JF at eligible earning levels.

There is a substantial opt-in response among women who would have worked off welfare at earning levels above the poverty line. The estimated bounds imply that $\pi_{2n,1r} \geq .28$. That is, at least 28% of those women with ineligible earnings under AFDC decided to work at eligible levels

[30] For example, if the relevant lower bound for $\pi_{2n,1r}$ is $\frac{p^a_{2n}-p^j_{2n}}{p^a_{2n}}$ and the relevant upper bound is $\frac{p^a_{2n}-p^j_{2n}+p^a_{0p}-p^j_{0p}}{p^a_{2n}}$, then the 95% bootstrap confidence interval for $\pi_{2n,1r}$ is: $\left[\max\left\{0, \frac{\widehat{p}^a_{2n}-\widehat{p}^j_{2n}}{\widehat{p}^a_{2n}} - 1.65\widehat{\sigma}_l\right\}, \min\left\{1, \frac{\widehat{p}^a_{2n}-\widehat{p}^j_{2n}+\widehat{p}^a_{0p}-\widehat{p}^j_{0p}}{\widehat{p}^a_{2n}} + 1.65\widehat{\sigma}_u\right\}\right]$ with $\widehat{\sigma}_l$ the bootstrap standard error of $\frac{\widehat{p}^a_{2n}-\widehat{p}^j_{2n}}{\widehat{p}^a_{2n}}$ and $\widehat{\sigma}_u$ the bootstrap standard error of $\frac{\widehat{p}^a_{2n}-\widehat{p}^j_{2n}}{\widehat{p}^a_{2n}} + \frac{\widehat{p}^a_{0p}-\widehat{p}^j_{0p}}{\widehat{p}^a_{2n}}$ and where hats on probabilities denote reweighted sample analogues.

under JF and participate in welfare. Accounting for sampling uncertainty in the bounds extends this lower limit to 19%, which is still quite substantial. The upper bounds for this parameter are not informative leading us to conclude that the opt-in probability lies in the interval $[.19, 1]$ with 95% probability.

We also find suggestive evidence of a second opt-in effect from non-participation. The sample bounds imply $\pi_{0n,1r} \in [.06, .62]$. However, uncertainty in the bounds prevents us from rejecting the null that this response probability is actually zero. We also find a small but significant under-reporting response attributable to the hassle effects of JF. A conservative 95% confidence interval for $\pi_{0r,2u}$ is $[.02, .13]$. Thus, JF induced at least one subpopulation to under-report earnings, and in the process violate the standard rank invariance condition implicit in BGH's analysis.

The remaining response probabilities $(\pi_{0r,0n}, \pi_{0r,2n}, \pi_{0r,1r}, \pi_{2u,1r})$ each have zero lower bounds. However, we can reject the null that they are jointly zero. From (7) such a joint restriction implies $p^j_{0p} - p^a_{0p} = -\left(p^j_{2p} - p^a_{2p}\right)$, which is easily rejected by our data. Thus, at least some of these margins of adjustment are present. Among the probabilities in question, the candidate that seems most likely to be positive is $\pi_{0r,1r}$ which is the extensive margin response through which welfare reform has traditionally been assumed to operate. However, we cannot be sure that the abundance of women working at low earning levels under JF are in fact coming from state $0r$ rather than state $2u$.

The last four rows of Table 4 report the estimated bounds, and corresponding confidence intervals, for the composite margins described in the previous section. First is the probability $\pi_{0,1+}$ that a woman responds along the extensive margin from nonwork to work. A conservative 95% confidence interval for this probability is $[0.13, 0.21]$. Thus, JF induced a substantial fraction of women who would not have worked under AFDC to obtain employment under JF.

The confidence interval on the fraction $\pi_{n,p}$ of women induced to take up welfare by JF is relatively tight. Although JF unambiguously increased the fraction of women on welfare, our model suggests some women may also have been induced to leave welfare, breaking point identification of this margin. According to our conservative inference procedure, at least 19% (and at most 51%) of women off welfare under AFDC were induced to claim benefits under JF. Conversely, the fraction $\pi_{p,n}$ of women induced by JF to leave welfare is estimated to be at least zero and at most 17%.

Finally, we cannot reject the null hypothesis that JF failed to induce any of the women who would have not worked while claiming AFDC benefits to leave welfare under JF, as the lower bound for the response probability $\pi_{0r,n}$ is zero. We are however able to conclude that at most 24% of such women left welfare, which may limit concerns that the JF reforms pushed a large fraction of women potentially unable to work off assistance.

8 Discussion

Here we discuss a number of potential extensions to our approach and issues which may affect the interpretation of our results. First, we explain our decision not to use the panel structure of the data to infer counterfactuals. Second, we argue that incorporating Food Stamps and the Earned Income Tax Credit into the budget set has no effect on our identification results. Finally, we discuss why time limits, an inherently dynamic feature of the JF intervention, could yield a violation of some of the exclusion restrictions used for identification. Following Grogger and Michaelopolous (2003), we conduct a simple test for anticipatory behavior which leads us to conclude that anticipation effects are likely to be small in the JF experiment.

The Panel Structure and Counterfactuals

Our empirical analysis has pooled together person-quarter observations, effectively treating each woman as a different decisionmaker each quarter. This reflects our reticence to take a stand on how preferences and constraints evolve across quarters. Consider a women within the JF sample who occupies state $0p$ in the quarter prior to random assignment. Does the fraction of such women who end up in state $1p$ after reform identify $\pi_{0p,1p}$? In general, the answer is no: womens' states would evolve even in the absence of reform, preventing us from equating temporal flows with adjustments between counterfactual states.

The problem is illustrated in Appendix Table 3 which provides state probabilities in quarters 1 through 7 among the subsample of women assigned to AFDC who chose state $0p$ in the quarter prior to random assignment. Even in the first quarter after random assignment, many of these women have switched states, suggesting substantial drift in preferences and constraints. This time variation prevents us from inferring the counterfactual behavior of a woman assigned to JF who transitions from state $0p$ to another state after reform. While it is possible to structure the evolution of unobserved factors in a way that provides additional information about counterfactuals in each period, such an approach runs counter to the spirit of our exercise which seeks to establish under minimal assumptions whether intensive margin responses are present in the data.

Policy Interactions

Cash assistance is part of a broader web of tax and transfer programs that can interact in subtle ways. Fortunately, these interactions do not change our basic insight that JF incentivized women to work on welfare below the poverty line and to avoid drawing cash assistance without working.

Recall that Figure 1 depicts the budget set of a long term recipient under the simplifying, albeit unrealistic, assumption that the only available transfer program is welfare under either AFDC or JF. In Appendix Figure 2 we depict this woman's budget set accounting for the other tax and transfer programs that are empirically relevant for would-be welfare recipients, namely the

Food Stamps (FS) program and the federal tax system, including payroll taxes and the Earned Income Tax Credit (EITC).[31] The FS program interacts with cash assistance both because welfare recipients are categorically eligible for FS and because welfare benefits are treated as income in the determination of the FS transfer. By contrast, the EITC and other taxes do not directly interact with cash and in-kind assistance because income from FS and welfare is not counted in the determination of taxes and tax credits.

Notably, the JF intervention changed the earnings disregard for both welfare and the FS program. Conditional on joint take up, earnings up to the FPL were disregarded in full for both the determination of the welfare grant and the FS grant under JF. This feature of the JF intervention is clearly visible in Appendix Figure 2: the segment of the budget set corresponding to joint participation in welfare and FS under JF entails a combined grant that does not change with earnings up to the FPL, at which point it falls to zero. Thus, JF's impact on the FS program amplifies the notch at the FPL.

Accounting for the FS program and the tax code does not affect how women can respond to the JF intervention when the outcomes of interest are participation in welfare and coarsened earnings.[32] As in the scenario without multiple transfer programs, it is never optimal to under-report earnings below the FPL under JF which implies the exclusion restrictions pertaining to the pairing of any state under AFDC with state 1u under JF. Additionally, the JF intervention leaves the value of states off welfare assistance unchanged, potentially worsens the value of welfare assistance when paired with non-employment, and increases the value of welfare assistance when paired with employment. These three facts suffice to apply the revealed preference argument underlying most of the remaining exclusion restrictions embedded in matrix Π. Specifically, no woman will pair any of the states in $\mathcal{C}^1$ under AFDC with a (different) state in the set $\mathcal{C}^2$ under JF because the former choices continue to be made no worse by reform while the latter choices are made no better. Finally, the exclusion restriction pertaining to the pairing of $0r$ under AFDC with $1n$ under JF also continues to hold. Indeed, irrespective of whether under AFDC a woman chooses state $0r$ in tandem with FS or by itself, such a choice reveals that, given any positive earnings below the FPL, she prefers to be on welfare assistance under JF. In sum, all the exclusion restrictions embedded in matrix Π are robust to the introduction of the FS program and the tax code.

Forward Looking Behavior

Our results rely upon a myopic model of decision making. In practice, women are likely to make choices taking into account both current and future payoffs. From our perspective, these dynamic

[31] The code used to construct Appendix Figure 2 is available upon request. Figure assumes woman considers using "Advance" EITC which allows claiming monthly tax credit.

[32] The richer policy environment, and the fact that the administrative data contains information on FS take up, suggests explicitly augmenting the model to account for multiple program participation and deriving restrictions on an expanded matrix of responses involving multiple participation states. We leave such an extension to future research.

motives are only of concern if they rationalize responses prohibited by our model. For this to be the case, alternative specific continuation values would need to differ across AFDC and JF in ways that undermine our static conclusions regarding which choices were made more or less attractive by reform.

The most obvious culprit for such effects are the JF time limits, which create incentives for a risk averse woman to save months of welfare eligibility for later periods when her earnings may be lower (e.g. due to job loss). Thus, under some conditions, JF may actually make working on welfare (state $1r$) less attractive, as this choice requires sacrificing the option value of using welfare an additional month in the future, which could be very costly if one is risk averse. If state $1r$ is made less attractive in response to reform then additional responses are possible. For example, a woman might choose state $1r$ under AFDC but state $1n$ under JF, a response prohibited by the current myopic model.

Following Grogger and Michaelopolous (2003), we conduct a simple test for whether the JF time limits yield anticipatory effects. Our test compares the impact of reform on the welfare use of women who at baseline had a youngest child age 16-17 (for whom the time limits were irrelevant) to impacts on the welfare use of women who had younger children. As shown in Appendix Table 4, we cannot reject the null hypothesis that the average impact of JF on monthly welfare take-up is the same for both groups of women. In fact, our point estimates suggest that the response of women with younger children to reform was actually slightly greater than the response of women with children ages 16-17, which is the opposite of what banking behavior would suggest. While this finding does not prove that the women in our sample were myopic, it does suggest that anticipatory responses to the time limits were probably small. An extension of the methods developed here to dynamic optimizing models is an important area for future research.

9 Conclusion

Our analysis of the Jobs First experiment suggests that women responded to the policy incentives of welfare reform along several margins, some of which are intensive and some of which are extensive. This finding is in accord with BGH's original interpretation of the JF experiment and with recent evidence from Blundell, Bozio, and Laroque (2011a,b) who find that secular trends in aggregate hours worked appear to be driven by both intensive and extensive margin adjustments. Our conclusions are also qualitatively consistent with recent studies relying on dynamic parametrically structured labor supply models (e.g., Blundell et al., 2012; Blundell et al., 2013).

An important question is the extent to which our finding of intensive margin responsiveness might generalize to other transfer programs that lack sharp budget notches but still involve phase-out regions that should discourage work. It seems plausible that the JF notch would yield larger disincentive effects than, say, the budget kink induced by the EITC phase-out region. However, BGH (2008) show that experimental responses to a Canadian reform inducing such a gradual benefit phaseout generated a pattern of earnings QTEs similar to that found in the JF experiment. More conclusive evidence on this question may be had via an application of the methods developed here to other policy reforms.

Though we studied a randomized experiment, our approach is easily generalized to quasi-experimental settings. Estimates of the relevant counterfactual choice probabilities can be formed using one's research design of choice (e.g., a difference in differences design), subject to the usual caveat that different designs may identify counterfactuals for different treated subpopulations.[33] With the two sets of marginal choice probabilities, bounds on response probabilities can then be had by a direct application of the methods developed in this paper.

As with most methods designed for the study of treatment effects, we cannot, without additional assumptions, predict the responses likely to arise from new interventions outside the range of observed policy variation. In cases where data are available on many different sorts of policy interventions, one can fit a curve summarizing how the bounds on response probabilities vary with policy parameters and attempt a statistical extrapolation. Otherwise, restrictions on model primitives will be necessary for prediction. A natural approach would be to parameterize features of utility and/or the process governing the labor supply constraints (e.g. as in Chetty et al., 2011), in which case bounds can be developed on finite dimensional structural parameters rather than transition probabilities. A challenge to such approaches is establishing consensus on functional forms, as inappropriate parametric restrictions tend to overstate, rather than simply approximate, what is known in partially identified settings (Ponomareva and Tamer, 2010; Kline and Santos, 2013). We leave the development of such semi-parametric methods to future work.

References

1. Bloom, Dan, Susan Scrivener, Charles Michalopoulos, Pamela Morris, Richard Hendra, Diana Adams-Ciardullo, Johanna Walter, and Wanda Vargas. 2002. *Jobs First: Final Report on Connecticut's Welfare Reform Initiative.* New York: Manpower Demonstration Research Corporation.

2. Andrews, Donald and S. Han. 2009. "Invalidity of the Bootstrap and m Out of n Bootstrap for Interval Endpoints Defined by Moment Inequalities." *Econometrics Journal* 12(s1): S172-S199.

3. Ashenfelter, Orley. 1983. "Determining Participation in Income-Tested Social Programs." *Journal of the American Statistical Association* 78(383):517-525.

4. Bitler, Marianne, Jonah Gelbach, and Hilary Hoynes. 2005. "Distributional Impacts of the Self-Sufficiency Project." *NBER Working Paper #11626.*

5. Bitler, Marianne, Jonah Gelbach, and Hilary Hoynes. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review* 96(4): 988-1012.

[33] For example, if one uses an instrumental variables design, counterfactuals are, under suitable restrictions, identified only for the subpopulation of "compliers" (Imbens and Rubin, 1997).

6. Bitler, Marianne, Jonah Gelbach, and Hilary Hoynes. 2010. "Can Variation in Subgroups' Average Treatment Effects Explain Treatment Effect Heterogeneity? Evidence from a Social Experiment." working paper.

7. Blundell, Richard and Thomas Macurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics*, Vol. 3A, ed. Orley Ashenfelter and David Card, pp.1559-1695. Amsterdam: Elsevier.

8. Blundell, Richard, Antoine Bozio, and Guy Laroque. 2011a. "Extensive and Intensive Margins of Labour Supply: Working Hours in the US, UK and France." IZA Discussion Paper #6051.

9. Blundell, Richard, Antoine Bozio, and Guy Laroque. 2011b. "Labour Supply and the Extensive Margin." *American Economic Review, Papers and Proceedings* 101(3): 482-486.

10. Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2012. "Consumption Inequality and Family Labor Supply." NBER Working Paper #18445.

11. Blundell, Richard, Monica Costas Dias, Costas Meghir, and Jonathan Shaw. 2013. "Female Labor Supply, Human Capital, and Welfare Reform." NBER Working Paper #19007.

12. Burtless, Gary and Jerry Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *The Journal of Political Economy* 86(6):1103-1130.

13. Card, David and Dean Hyslop. 2005. "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers." *Econometrica*, 73(6): 1723-1770.

14. Chetty, Raj. 2012. "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica*, 80(3):969-1018.

15. Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2011. "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review* 101(3): 471-475.

16. Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence From Danish Tax Records." *Quarterly Journal of Economics* 126(2): 749-804.

17. Chetty, Raj, John Friedman, and Emmanuel Saez. 2012. "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." NBER Working Paper #18232.

18. Cogan, John. 1981. "Fixed Costs and Labor Supply." *Econometrica* 49 (4): 945–63.

19. Deaton, Angus. 2009. "Instruments of Development: Randomization in the tropics, and the search for the elusive keys to economic development." NBER Working Paper #14690.

20. Diamond, Peter. 1980. "Income Taxation with Fixed Hours of Work." *Journal of Public Economics* 13(1): 101-110.

21. Eissa, Nada, and Jeffrey B. Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics*, 111(2): 605-37.

22. Eissa, Nada, and Hilary Hoynes. 2006. "The Hours of Work Response of Married Couples: Taxes and the Earned Income Tax Credit." In *Tax Policy and Labor Market Performance*, ed. Jonas Agell and Peter Birch Sorensen, pp.187-227. Cambridge: MIT Press.

23. Greenberg, David, Robert Moffitt, and John Friedmann. 1981. "Underreporting and Experimental Effects on Work Effort: Evidence from the Gary Income Maintenance Experiment." *Review of Economics and Statistics* 63(4): 581-589.

24. Greenberg, David and Harlan Halsey. 1983. "Systematic Misreporting and Effects of Income Maintenance Experiments on Work Effort: Evidence from the Seattle-Denver Experiment." *Journal of Labor Economics* 1(4): 380-407.

25. Grogger, Jeffrey and Charles Michalopoulos. 2003. "Welfare Dynamics under Time Limits." *Journal of Political Economy* 111(3): 530-554.

26. Grogger, Jeffrey and Lynn Karoly. 2005. *Welfare Reform: Effects of a Decade of Change*. Cambridge, MA: Harvard University Press.

27. Hansen, Gary D. 1985. "Indivisible Labor and the Business Cycle." *Journal of Monetary Economics* 16:309-327.

28. Heckman, James and Jeffrey Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9(2): 85-110.

29. Heckman, James, Jeffrey Smith, and Nancy Clements. 1997. "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts." *Review of Economic Studies* 64(4): 487-535.

30. Heckman, James J. 1983. "Comment." In *Behavioral Simulation Methods in Tax Policy Analysis*, ed. Martin Feldstein, 70–82. Chicago: University of Chicago Press.

31. Heckman, James. 1993. "What Has Been Learned About Labor Supply in the Last 20 Years?" *American Economic Review Papers and Proceedings* 83(2): 116-121.

32. Heckman, James. 2001. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109(4): 673-748.

33. Hotz, Joseph , Charles Mullin, and Karl Scholz. 2003. "The Effects of Welfare Reform on Employment and Income: Evidence from California." working paper.

34. Hoynes, Hilary. 1996. "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP." *Econometrica* 64(2): 295-332.

35. Imbens, Guido and Donald Rubin. 1997. "Estimating Outcome Distributions for Compliers in Instrumental Variables Models." *Review of Economic Studies* 64(4): 555-574.

36. Keane, Michael, and Robert Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review*, 39(3): 553-89.

37. Kline, Patrick and Andres Santos. 2013. "Sensitivity to Missing Data Assumptions: Theory and an Evaluation of the U.S. Wage Structure." *Quantitative Economics*, forthcoming.

38. Laroque, Guy. 2005. "Income Maintenance and Labor Force Participation." *Econometrica* 73(2): 341-376.

39. Ljungqvist, Lars and Thomas Sargent. 2009. "A Labor Supply Elasticity Accord?" *American Economic Review Papers and Proceedings* 101(3): 487-491.

40. Macurdy, Thomas, David Green, and Harry Paarsch. 1990. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources* 25(3): 415-490.

41. Manski, Charles F. 2012. "Identification of Preferences and Evaluation of Income Tax Policy." (No. w17755). National Bureau of Economic Research.

42. Meyer, Bruce and Dan Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 116(3): 1063-1114.

43. Meyer, Bruce. 2002. "Labor Supply at the Intensive and Extensive Margins: the EITC, Welfare, and Hours Worked." *American Economic Review* 92(2): 373-379.

44. Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *The American Economic Review* 73(5): 1023-1035.

45. Moffitt, Robert. 1990. "The Econometrics of Kinked Budget Constraints." *Journal of Economic Perspectives* 4(2): 119-139.

46. Moffitt, Robert. 2002. "Welfare Programs and Labor Supply." In *Handbook of Public Economics.* Vol. 4, ed. Alan J. Auerbach and Martin S. Feldstein, 2393-2430. New York: North-Holland.

47. Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, 3:145-146.

48. Ponomareva, Maria and Elie Tamer. 2010. "Misspecification in moment inequality models: back to moment equalities?" *The Econometrics Journal* 14(2): 186-203.

49. Rogerson, Richard. 1988. "Indivisible Labor, Lotteries, and Equilibrium." *Journal of Monetary Economics* 21:3-16.

50. Saez, Emmanuel. 2002. "Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses." *Quarterly Journal of Economics* 117(3): 1039-1073.

51. Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2(3): 180-212.

Table 1: Mean Sample Characteristics

	Overall Sample				Zero Earnings Q7 pre-RA				Low Earnings Q7 pre-RA				High Earnings Q7 pre-RA			
	Jobs First	AFDC	Difference	Difference (adjusted)	Jobs First	AFDC	Difference	Difference (adjusted)	Jobs First	AFDC	Difference	Difference (adjusted)	Jobs First	AFDC	Difference	Difference (adjusted)
Demographic Characteristics																
White	0.374	0.360	0.014	0.001	0.340	0.331	0.009	-0.001	0.461	0.435	0.026	-0.002	0.446	0.407	0.039	0.004
Black	0.380	0.384	-0.004	0.000	0.370	0.360	0.010	0.001	0.382	0.443	-0.061	0.000	0.426	0.426	0.000	-0.005
Hispanic	0.214	0.224	-0.010	-0.001	0.258	0.275	-0.017	0.000	0.131	0.102	0.030	0.000	0.090	0.134	-0.044	-0.002
Never married	0.654	0.661	-0.007	0.000	0.658	0.654	0.003	0.000	0.701	0.736	-0.034	0.009	0.589	0.609	-0.021	-0.002
Div/wid/sep/living apart	0.332	0.327	0.005	0.000	0.327	0.334	-0.007	0.000	0.287	0.246	0.041	-0.008	0.402	0.382	0.020	0.002
HS dropout	0.350	0.334	0.017	0.000	0.390	0.394	-0.004	0.000	0.323	0.237	0.086	-0.002	0.192	0.179	0.012	0.011
HS diploma/GED	0.583	0.604	-0.021	0.000	0.550	0.555	-0.005	-0.001	0.623	0.707	-0.084	0.003	0.699	0.706	-0.007	-0.013
More than HS diploma	0.066	0.062	0.004	0.000	0.060	0.051	0.009	0.000	0.054	0.056	-0.002	0.000	0.109	0.115	-0.006	0.002
More than 2 Children	0.235	0.214	0.021	0.000	0.260	0.250	0.010	0.000	0.187	0.128	0.059	-0.006	0.165	0.150	0.015	0.005
Mother younger than 25	0.287	0.298	-0.011	-0.003	0.287	0.268	0.019	-0.001	0.376	0.456	-0.080	0.006	0.200	0.262	-0.062	-0.001
Mother age 25-34	0.412	0.414	-0.003	0.005	0.410	0.419	-0.009	0.000	0.367	0.346	0.021	-0.001	0.464	0.467	-0.003	-0.002
Mother older than 34	0.301	0.287	0.014	-0.002	0.303	0.313	-0.010	0.001	0.257	0.198	0.059	-0.006	0.336	0.270	0.066	0.003
Average quarterly pretreatment values																
Earnings	673	750	-76*	4	174	185	-11	2	763	856	-94*	-4	2943	3066	-123	17
	[1306]	[1379]	(40)	(6)	[465]	[479]	(17)	(4)	[679]	[706]	(51)	(31)	[1911]	[1957]	(145)	(124)
Cash welfare	903	845	58**	1	1050	1022	28	0	804	701	102**	6	307	238	70**	-21
	[805]	[784]	(23)	(2)	[811]	[799]	(28)	(3)	[721]	[662]	(51)	(23)	[525]	[418]	(35)	(83)
Food stamps	356	344	12	0	399	398	1	1	335	300	35*	1	171	157	14	-13
	[320]	[304]	(9)	(1)	[326]	[310]	(11)	(1)	[301]	[268]	(21)	(16)	[235]	[221]	(16)	(48)
Fraction of pretreatment quarters with																
Any earnings	0.319	0.347	-0.029***	0.000	0.137	0.143	-0.007	0.000	0.661	0.694	-0.033*	0.000	0.839	0.861	-0.022	0.002
	[0.362]	[0.370]	(0.011)	(0.001)	[0.211]	[0.215]	(0.008)	(0.001)	[0.273]	[0.258]	(0.019)	(0.009)	[0.218]	[0.180]	(0.015)	(0.017)
Any cash welfare	0.581	0.551	0.030*	-0.001	0.650	0.636	0.014	0.000	0.578	0.535	0.043	0.004	0.259	0.204	0.055**	-0.012
	[0.451]	[0.450]	(0.013)	(0.001)	[0.439]	[0.439]	(0.015)	(0.001)	[0.445]	[0.442]	(0.032)	(0.014)	[0.369]	[0.307]	(0.025)	(0.045)
Any food stamps	0.613	0.605	0.008	0.000	0.670	0.674	-0.004	0.001	0.611	0.591	0.020	0.003	0.349	0.322	0.027	-0.012
	[0.437]	[0.431]	(0.012)	(0.001)	[0.427]	[0.421]	(0.015)	(0.001)	[0.431]	[0.424]	(0.031)	(0.013)	[0.394]	[0.364]	(0.028)	(0.042)
# of cases	2,318	2,324			1,630	1,574			343	384			345	366		

Notes: Sample units with kidcount missing are excluded. Adjusted differences are computed via propensity score reweighting. Numbers in brackets are standard deviations and numbers in parentheses are standard errors calculated via 1,000 block bootstrap replications (resampling at case level). ***, **, and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels, respectively (significance indicators provided only for difference estimates).

Table 2: Mean Outcomes Post-Random Assignment

	Overall			Zero Earnings Q7 pre-RA			Low Earnings Q7 pre-RA			High Earnings Q7 pre-RA		
	Jobs First	AFDC	Adjusted Difference	Jobs First	AFDC	Adjusted Difference	Jobs First	AFDC	Adjusted Difference	Jobs First	AFDC	Adjusted Difference
Average Earnings	1,191	1,086	**105**	930	751	**179**	1,362	1,291	**70**	2,124	2,362	**-238**
	(29)	(30)	**(36)**	(32)	(30)	**(42)**	(66)	(101)	**(118)**	(114)	(151)	**(179)**
Fraction of quarters with positive earnings	0.520	0.440	**0.080**	0.445	0.349	**0.096**	0.691	0.590	**0.100**	0.680	0.690	**-0.011**
	(0.008)	(0.007)	**(0.010)**	(0.009)	(0.009)	**(0.012)**	(0.018)	(0.022)	**(0.028)**	(0.022)	(0.028)	**(0.035)**
Fraction of quarters with earnings below monthly FPL (AU size implied by kidcount+1)	0.665	0.710	**-0.046**	0.731	0.789	**-0.058**	0.570	0.636	**-0.066**	0.477	0.438	**0.039**
	(0.004)	(0.005)	**(0.006)**	(0.004)	(0.004)	**(0.006)**	(0.012)	(0.017)	**(0.020)**	(0.019)	(0.025)	**(0.030)**
Fraction of quarters with earnings below 3FPL (AU size implied by kidcount+1)	0.906	0.897	**0.009**	0.938	0.940	**-0.002**	0.906	0.881	**0.025**	0.777	0.722	**0.055**
	(0.007)	(0.007)	**(0.009)**	(0.008)	(0.008)	**(0.010)**	(0.019)	(0.021)	**(0.027)**	(0.023)	(0.030)	**(0.036)**
Fraction of quarters on welfare	0.748	0.674	**0.074**	0.771	0.718	**0.053**	0.764	0.674	**0.091**	0.637	0.475	**0.162**
	(0.007)	(0.007)	**(0.010)**	(0.008)	(0.008)	**(0.011)**	(0.019)	(0.022)	**(0.029)**	(0.024)	(0.033)	**(0.040)**
Average earnings in quarters with any month on welfare	929	526	**403**	762	404	**359**	1,123	694	**428**	1,524	1,075	**449**
	(24)	(19)	**(28)**	(25)	(18)	**(30)**	(53)	(48)	**(69)**	(96)	(119)	**(147)**
Fraction of quarters with no earnings and at least one month on welfare	0.363	0.437	**-0.074**	0.426	0.508	**-0.082**	0.231	0.334	**-0.103**	0.221	0.220	**0.001**
	(0.007)	(0.007)	**(0.010)**	(0.009)	(0.009)	**(0.012)**	(0.015)	(0.022)	**(0.027)**	(0.019)	(0.028)	**(0.033)**
# of cases	2,318	2,324		1,630	1,574		343	384		345	366	

Notes: Sample covers quarters 1-7 post-random assignment. Sample units with kidcount missing are excluded. Adjusted differences are computed via propensity score reweighting. Numbers in parentheses are standard errors calculated via 1,000 block bootstrap replications (resampling at case level).

Table 3: Probability of Earnings / Participation States

	Overall			Overall - Adjusted		
	Jobs First	AFDC	Difference	Jobs First	AFDC	Difference
Pr(State=0n)	0.127	0.136	-0.009	0.128	0.135	**-0.007**
				(0.006)	(0.006)	**(0.008)**
Pr(State=1n)	0.076	0.130	-0.055	0.078	0.126	**-0.048**
				(0.004)	(0.005)	**(0.006)**
Pr(State=2n)	0.068	0.099	-0.031	0.069	0.096	**-0.027**
				(0.004)	(0.005)	**(0.006)**
Pr(State=0p)	0.366	0.440	-0.074	0.359	0.449	**-0.090**
				(0.008)	(0.008)	**(0.012)**
Pr(State=1p)	0.342	0.185	0.157	0.343	0.184	**0.159**
				(0.008)	(0.006)	**(0.009)**
Pr(State=2p)	0.022	0.009	0.013	0.023	0.009	**0.014**
				(0.002)	(0.001)	**(0.002)**
# of quarterly observations	16,226	16,268		16,226	16,268	

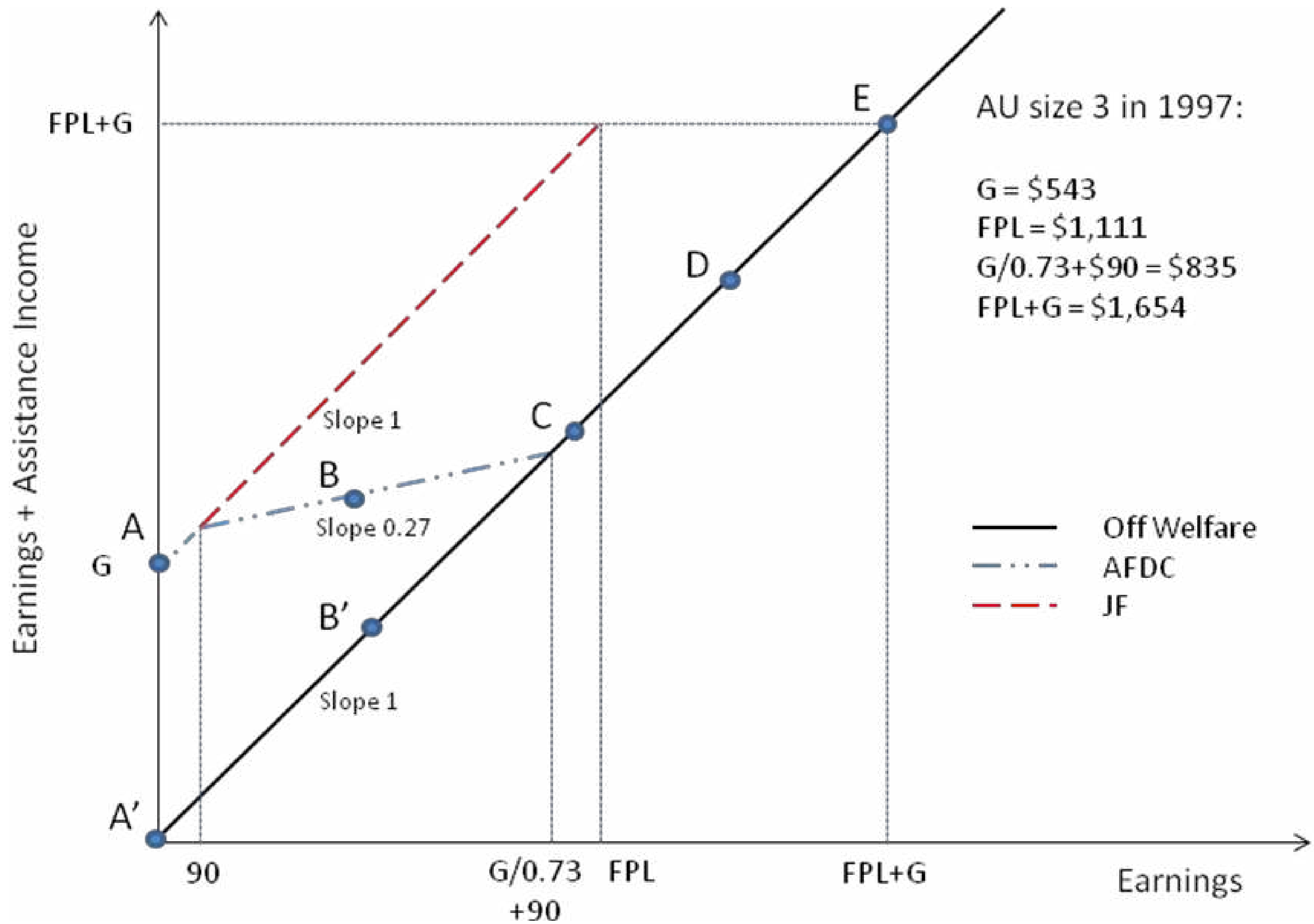
Notes: Sample covers quarters 1-7 post-random assignment during which individual is either always on or always off welfare. Sample units with kidcount missing are excluded. Number of state refers to earnings level, with 0 indicating no earnings, 1 indicating earnings below 3 times the monthly FPL, and 2 indicating earnings above 3FPL. The letter n indicates welfare nonparticipation throughout the quarter while the letter p indicates welfare participation throughout the quarter. Poverty line computed under assumption AU size is one greater than amount implied by baseline kidcount variable. Adjusted probabilities are adjusted via the propensity score reweighting algorithm described in the Appendix. Standard errors computed using 1,000 block bootstrap replications (resampling at case level).

Table 4: Point and Set-identified Response Margins

	Estimate	95% CI (deterministic bound)	95% CI (conservative)
Point-identified Margins			
Off welfare, low earnings → On welfare, low earnings ($\pi_{1n,1r}$)	0.381	[0.306, 0.455]	[0.306, 0.455]
Set-identified Margins			
On welfare, not working → Off welfare, not working ($\pi_{0r,0n}$)	{0.000, 0.169}	[0.000, 0.210]	[0.000, 0.245]
On welfare, not working → On welfare, low earnings ($\pi_{0r,1r}$)	{0.000, 0.169}	[0.000, 0.210]	[0.000, 0.250]
On welfare, not working → Off welfare, high earnings ($\pi_{0r,2n}$)	{0.000, 0.154}	[0.000, 0.170]	[0.000, 0.302]
On welfare, not working → On welfare, high earnings ($\pi_{0r,2u}$)	{0.031, 0.051}	[0.022, 0.059]	[0.022, 0.131]
Off welfare, not working → On welfare, low earnings ($\pi_{0n,1r}$)	{0.059, 0.618}	[0.000, 0.755]	[0.000, 0.875]
Off welfare, high earnings → On welfare, low earnings ($\pi_{2n,1r}$)	{0.281, 1.000}	[0.193, 1.000]	[0.193, 1.000]
On welfare, high earnings → On welfare, low earnings ($\pi_{2u,1r}$)	{0.000, 1.000}	[0.000, 1.000]	[0.000, 1.000]
Composite Margins			
Not working → Working ($\pi_{0,1+}$)	0.168	[0.129, 0.206]	[0.129, 0.206]
Off welfare → On welfare ($\pi_{n,p}$)	{0.232, 0.444}	[0.188, 0.486]	[0.188, 0.509]
On welfare → Off welfare ($\pi_{p,n}$)	{0.000, 0.118}	[0.000, 0.147]	[0.000, 0.173]
On welfare, not working → Off welfare ($\pi_{0r,n}$)	{0.000, 0.169}	[0.000, 0.210]	[0.000, 0.244]

Notes: Estimates refer to the probability of occupying the state to the right of the arrow under JF rules among women who would occupy the state to the left of the arrow under AFDC rules. Estimates inferred from probabilities in Table 3, see text for formulas. Low earnings refers to quarterly earnings less than or equal to three times the monthly federal poverty line (E=1), high earnings refer to quarterly earnings above three times the monthly federal poverty line (E=2). Numbers in braces are estimated upper and lower bounds, numbers in brackets are 95% confidence intervals. Column labeled "deterministic bound" ignores uncertainty in which moment inequalities bind. Column labeled "conservative" uses inference procedure described in Appendix which covers true parameter with at least 95% probability regardless of which constraints bind. See text for details.

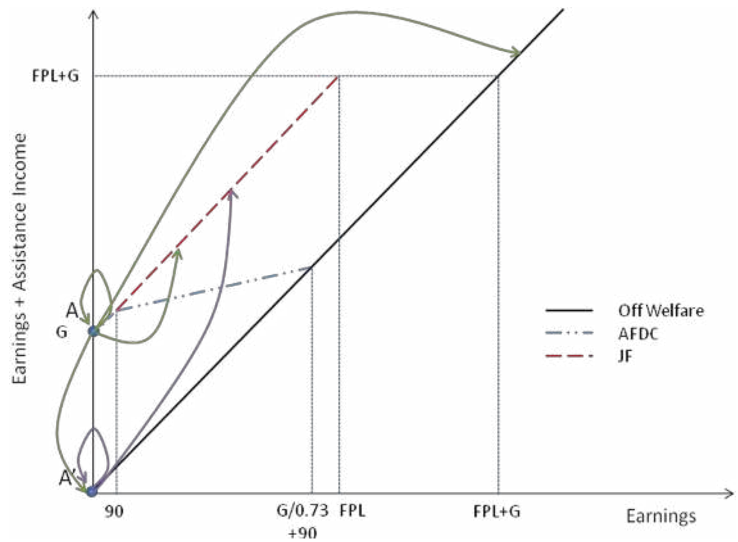
Figure 1: Hypothetical Budget Sets under AFDC and JF



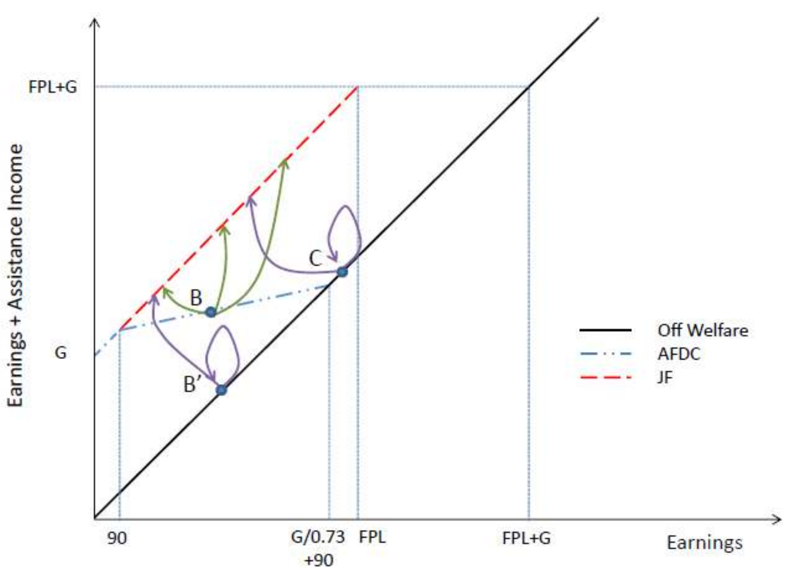
Notes: Figure depicts hypothetical monthly budget faced by assistance unit of size 3 under AFDC and Jobs First policy rules as of 1997. Illustration assumes household only has access to fixed $90 disregard under AFDC. FPL refers to federal poverty line and G is base grant amount.

Figure 2: Allowable Responses to JF by Hypothetical Choice under AFDC

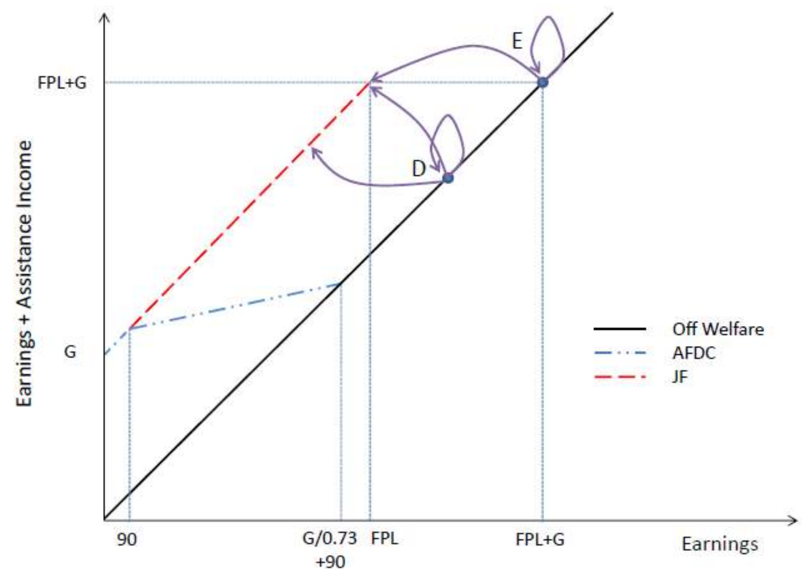
a) Zero Earnings



b) Positive Earnings, below FPL



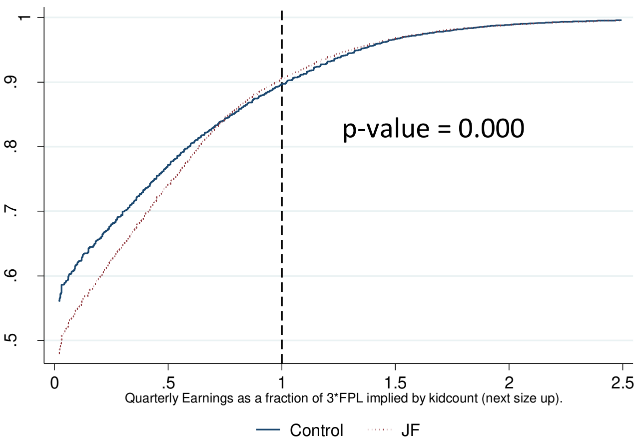
c) Positive Earnings, above FPL



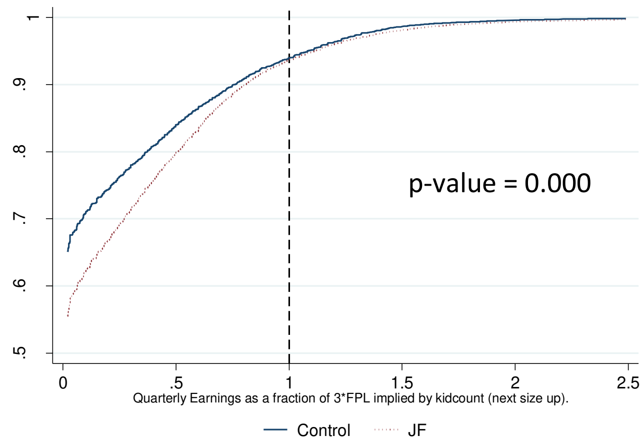
Notes: Figures give the choices that may emerge under JF among groups of women defined by their hypothetical choices under AFDC.

Figure 3: CDFs of Quarterly Earnings Relative to 3 x Federal Poverty Line

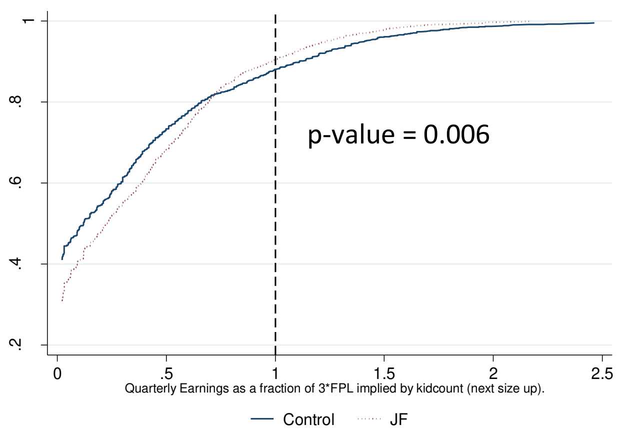
a) Overall



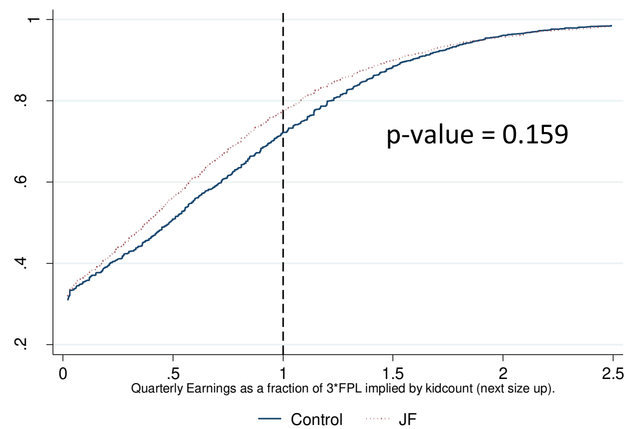
b) Zero Earnings pre-RA



c) Low Earnings pre-RA



d) High Earnings pre-RA



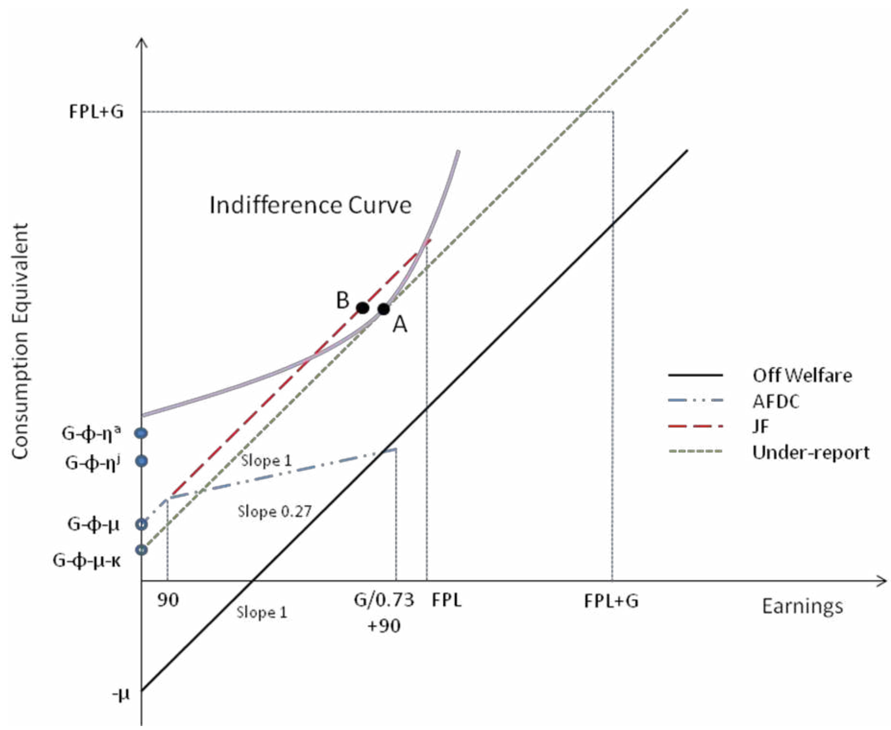
Notes: Figures give reweighted CDFs of quarterly UI earnings (in quarters 1-7 post-RA) in JF and AFDC samples relative to 3 x the monthly federal poverty line associated with year and AU size. Zero earnings pre-RA refers to women with zero earnings in the 7th quarter prior to random assignment, while low/high earnings refers to women with positive earnings below/above the median conditional on positive in the 7th quarter prior to random assignment. AU size determined by baseline survey question kidcount. To deal with increases in family size since random assignment, we use next AU size up relative to size directly implied by kidcount (see text for details). The p-value refers to a Kolmogorov-Smirnov test of equality of the two distributions (based on 1,000 replications, see Appendix for details).

Figure 4: Distribution of Quarterly Earnings Centered at 3 x Monthly Federal Poverty Line

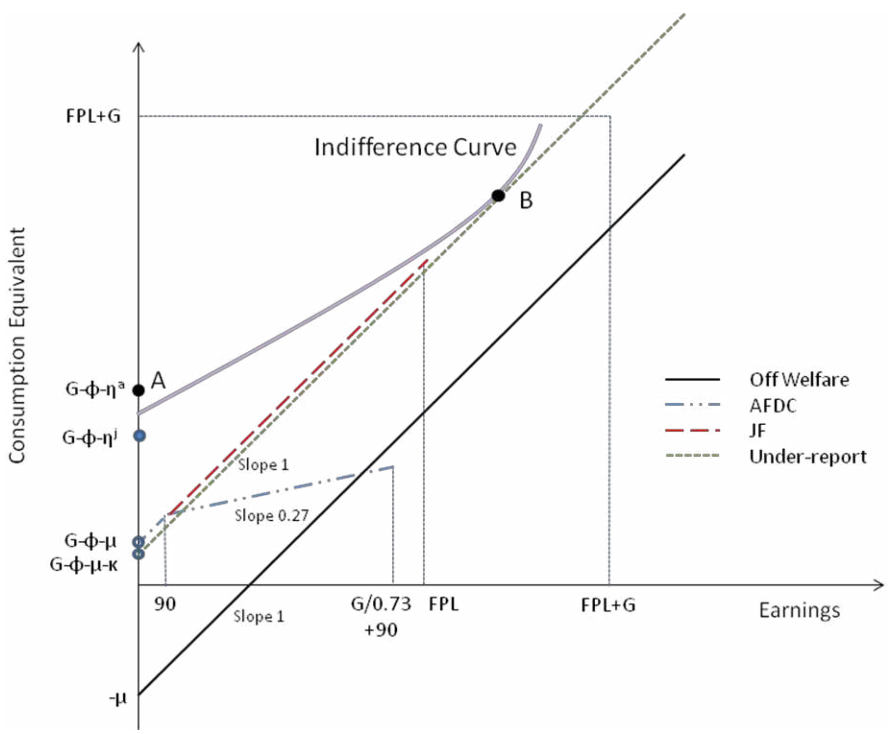
Notes: Restricted to Jobs First sample quarters 1-7 post-RA. Assistance unit size has been inferred from monthly aid payment. AU sizes above 8 have been excluded. The bins in the histograms are $100 wide with bin 0 containing three times the monthly federal poverty line corresponding to the size and the calendar year of the quarterly observation. Vertical line indicates JF eligibility threshold at three times the monthly federal poverty line.

Figure 5: Earnings and Participation Choices with Under-reporting

a) From Under-reporting to Truthful Reporting



b) From Truthful Reporting to Under-reporting



Notes: Panel (a) depicts a scenario where reform induces a woman who would participate in welfare, work, and under-report her earnings (point A) under AFDC to work and truthfully report her earnings (point B) under JF. Figure (b) depicts a scenario where reform induces a woman who would participate in welfare without working (point A) under AFDC to work and under-report her earnings (point B) in order to avoid welfare hassle under JF.

Online Appendix to Kline and Tartari (2013) - Not For Publication

Road-Map

In Section 1 of this supplementary appendix we prove that the response matrix Π takes the form described in Section 6 of the paper. We start by introducing definitions and restating the Assumptions made in the paper. We then prove a few intermediate lemmas. Specifically, we show that no woman will truthfully report earnings above the federal poverty level while on assistance (Lemma 1) and that no woman will under-report earnings that are below the federal poverty under JF (Lemma 2). The main revealed preference argument is given in Lemma 3, which states that no woman will pair a state chosen under AFDC whose utility value is at least as high under AFDC as under JF with a state under JF whose utility value is at most as high under JF as under AFDC. We conclude with a formal proof that matrix Π incorporates all the restrictions implied by the model. Specifically, we show that the zero entries of the matrix correspond to responses that, given the augmented model, cannot occur (Proposition 1). This also takes care of the entries of the matrix that equal one. Then we show that the *free* nonzero entries of the matrix correspond to responses that may or may not occur (Proposition 2). That is, there are no zeros or ones missing from the matrix. In Section 2 we describe the propensity score reweighting method that we use to adjust for chance imbalances in baseline characteristics. In Section 3 we explain how we construct the test for equality of distributions whose *p-values* are reported in Figure 3 of the paper. In Section 4 we list the analytical expressions for the bounds on the response probabilities, explain how we have derived the bounds, and describe the construction of the 95% confidence intervals reported in Table 4. Finally, we list the Appendix Figures and Tables.

1 The Response Matrix

Notation, Definitions, and Assumptions

Notation: Throughout, we use a to refer to AFDC and j to refer to JF. The policy regime is denoted by $t \in \{a, j\}$.

Definition 1 *Earnings range 0 refers to zero earnings. Earnings range 1 refers to the interval* $(0, FPL_i]$ *where* FPL_i *is woman i's federal poverty line. Earnings range 2 refers to the interval* (FPL_i, ∞).

Definition 2 *The regime dependent transfer functions are* $G_i^a(E^r) \equiv \max\left\{\overline{G}_i - 1[E^r > \delta_i](E^r - \delta_i)\tau_i, 0\right\}$ *and* $G_i^j(E^r) \equiv 1[E^r \leq FPL_i]\overline{G}_i$. *The parameter* $\delta_i \in \{90, 120\}$ *gives woman i's fixed disregard and the parameter* $\tau_i \in \{.49, .73\}$ *governs her proportional disregard.* $\overline{G}_i, FPL_i > 0$ *vary across women due to differences in AU size.*

Definition 3 *Define woman i's regime dependent consumption equivalent as* $C_i^t(E, D, E^r) \equiv E - \mu_i 1\{E > 0\} + D\left(G_i^t(E^r) - \phi_i - \eta_i^t 1\{E^r = 0\} - \kappa_i 1\{E^r < E\}\right)$.

Definition 4 *Woman i's "state" is defined by the following function:*

$$s_i(E, D, E^r) \equiv \begin{cases} 0n & \text{if } E = 0,\ D = 0 \\ 1n & \text{if } E \text{ in range 1},\ D = 0 \\ 2n & \text{if } E \text{ in range 2},\ D = 0 \\ 0r & \text{if } E = 0,\ D = 1 \\ 1r & \text{if } E \text{ in range 1},\ D = 1,\ E^r = E \\ 1u & \text{if } E \text{ in range 1},\ D = 1,\ E^r < E \\ 2u & \text{if } E \text{ in range 2},\ D = 1,\ E^r < E \\ 2r & \text{if } E \text{ in range 2},\ D = 1,\ E^r = E \end{cases}$$

Definition 5 *An allocation is an earnings and consumption equivalent pair (E, C). For simplicity, we refer to C as consumption.*

Definition 6 *We say that an allocation (E, C) is compatible with state s under regime t for woman i if there exists a pair $(D, E^r) \in \{0, 1\} \times [0, E]$ such that $s = s_i(E, D, E^r)$ and $C = C_i^t(E, D, E^r)$.*

Definition 7 *Unless specified otherwise, we denote earnings offers by O_i^k or O_i^l. It is implicit that $O_i^k, O_i^l \in \{O_i^1, O_i^2\}$, where (O_i^1, O_i^2) are woman i's two earning offers drawn from the bivariate distribution $F_i(.)$ with support on the strictly positive orthant. The statement $\forall O_i^k$ is shorthand for $\forall O_i^k \in \{O_i^1, O_i^2\}$.*

Definition 8 *We say that an allocation (E, C) is available and compatible with state s under regime t for woman i if it is compatible with state s under regime t and E corresponds to an earning draw O_i^k or $E = 0$.*

Definition 9 *We say that a state s is unpopulated under regime t if no available allocation compatible with s under regime t is chosen by any woman.*

Definition 10 *We say that a state s is no better (worse) under JF than under AFDC if, for any woman i, and any (E, D, E^r) such that $s = s_i(E, D, E^r)$, $U_i\left(E, C_i^j(E, D, E^r)\right) \leq (\geq) U_i(E, C_i^a(E, D, E^r))$. We say that a state s is equally attractive under JF and AFDC if, for any woman i, and any (E, D, E^r) such that $s = s_i(E, D, E^r)$, $U_i\left(E, C_i^j(E, D, E^r)\right) = U_i(E, C_i^a(E, D, E^r))$.*

Definition 11 *Define $\mathcal{S} \equiv \{0n, 1n, 2n, 0r, 1r, 1u, 2u\}$, $\mathcal{C}_+ \equiv \{1r\}$, $\mathcal{C}_- \equiv \{0r\}$ and $\mathcal{C}_0 \equiv \{0n, 1n, 2n, 1u, 2u\}$.*

Definition 12 *Consider those women who under AFDC choose a triplet (E, D, E^r) such that $s^a = s_i(E, D, E^r)$. We denote by π_{s^a, s^j} the fraction of them who under JF choose a triplet $(E', D', E^{r\prime})$ such that $s^j = s_i(E', D', E^{r\prime})$.*

Assumption 1 *Woman i's utility function $U_i(.,.)$ is decreasing in its first argument (earnings) and increasing in its second argument (consumption).*

Assumption 2 *For each woman i, $\left(\mu_i, \eta_i^a, \eta_i^j\right)$ are non-negative, $\eta_i^j \geq \eta_i^a$, $\kappa_i > 0$, and $\phi_i > \underline{\phi}_i \equiv G_i^a(FPL_i)$.*

Assumption 3 *Under regime t, woman i makes choices by solving the optimization problem:*

$$\max_{E \in \{0, O_i^1, O_i^2\}, D \in \{0,1\}, E^r \in [0, E]} U_i\left(E, C_i^t(E, D, E^r)\right).$$

Assumption 4 *Women break indifference in favor of the same allocation irrespective of the regime.*

Intermediate Lemmas

Lemma 1 *Given Assumptions 1, 2, and 3, state* $2r$ *is unpopulated.*

Proof. State $2r$ is unpopulated under regime j because $\phi_i > 0$ for all women (Assumptions 2) and the JF grant amount is zero whenever a woman reports earnings above FPL_i (Assumptions 1 and 3). We next show that state $2r$ is also unpopulated under regime a. Define woman i's break-even earnings level under a as $\overline{E}_i \equiv \frac{\overline{G}_i}{\tau_i} + \delta_i$, this is the level at which benefits are exhausted. If $\overline{E}_i \leq FPL_i$, she will not choose to truthfully report earnings above FPL_i (range 2) because $\phi_i > 0$ (Assumptions 2) and the AFDC grant amount is zero whenever she reports earnings above $\overline{E}_i$ (Assumptions 1 and 3). We now prove by contradiction that even when $\overline{E}_i > FPL_i$ a woman will not choose to truthfully report earnings above FPL_i (range 2). Suppose that woman i chooses an allocation that entails earnings $O_i^k \in (FPL_i, \overline{E}_i]$ and reports these earnings truthfully. By Assumption 3, her choice reveals that $U_i\left(O_i^k, O_i^k - \mu_i + G_i^a\left(O_i^k\right) - \phi_i\right) \geq U_i\left(O_i^k, O_i^k - \mu_i\right)$ which, because consumption is a good (Assumption 1), bounds her stigma from above, namely, $G_i^a\left(O_i^k\right) \geq \phi_i$. We thus have $G_i^a\left(O_k^i\right) \geq \phi_i > \underline{\phi}_i = G_i^a\left(FPL_i\right)$ because ϕ_i is bounded from below by $\underline{\phi}_i$ (Assumption 2). This yields a contradiction because $G_i^a\left(O_k^i\right) < G_i^a\left(FPL_i\right)$ for any $O_i^k > FPL_i$. Finally, suppose that under a woman i chooses an allocation that entails earnings $O_i^k > \overline{E}_i$ and reports these earnings truthfully. We again have a contradiction because $\phi_i > 0$ for all women (Assumption 2) and the AFDC grant amount is zero whenever she reports earnings above $\overline{E}_i$ (Assumptions 1 and 3). ■

Lemma 2 *Given Assumptions 1-4: a) the optimal reporting rule, while on assistance, entails either truthful reporting or reporting an amount in the range* $[0, FPL_i]$ *under JF or in the range* $[0, \delta_i]$ *under AFDC, b) when earnings are positive, reporting zero earnings is only optimal if stigma under the relevant regime is zero, and c) state* $1u$ *is unpopulated under JF.*

Proof. Let E and E^r denote the actual and reported earnings corresponding to an optimal allocation so that $E = O_i^k$ for some earning draw or $E = 0$. Consider first regime j and focus on three alternative optimal allocations: an allocation with E equal zero, an allocation with E in range 1, and an allocation with E in range 2. We now show that, by their optimality, each of these allocations entails either $E^r = E$ or $E^r \in [0, FPL_i]$. Women cannot over-report earnings (Assumption 3). Thus, truthful reporting is trivially optimal for a non-working woman. When E is in range 1, consumption while on welfare depends on reported earnings as follows:

$$C_i^j\left(E, 1, E^r\right) = \begin{cases} E - \mu_i + \overline{G}_i - \phi_i - \kappa_i - \eta_i^j & \text{if } E^r = 0,\ E \text{ in range 1} \\ E - \mu_i + \overline{G}_i - \phi_i - \kappa_i & \text{if } 0 < E^r < E,\ E \text{ in range 1} \\ E - \mu_i + \overline{G}_i - \phi_i & \text{if } E^r = E,\ E \text{ in range 1} \end{cases}.$$

Thus, truthful reporting maximizes consumption since $\kappa_i > 0, \eta_i^j \geq 0$ (Assumption 2). Hence, by Assumptions 1 and 3, truthful reporting must be optimal. When E is in range 2, consumption while on welfare depends on reported earnings as follows:

$$C_i^j\left(E, 1, E^r\right) = \begin{cases} E - \mu_i + \overline{G}_i - \phi_i - \kappa_i - \eta_i^j & \text{if } E^r = 0,\ E \text{ in range 2} \\ E - \mu_i + \overline{G}_i - \phi_i - \kappa_i & \text{if } E^r \text{ in range 1},\ E \text{ in range 2} \end{cases}.$$

When $\eta_i^j > 0$, reports in range 1 maximize consumption (and therefore utility). Since the grant amount is unaffected by earnings in this range, she may report any earnings in the range $(0, FPL_i]$. If $\eta_i^j = 0$, the woman is indifferent between reporting zero and amounts in the range $[0, FPL_i]$. This establishes parts a) and b) of the Lemma under regime j.

Consider next regime a and focus on four alternative optimal allocations: an allocation with $E = 0$, an allocation with $0 < E \leq \delta_i$ (by construction in range 1), an allocation with $E > \delta_i$ in range 1, and an allocation with E in range 2. We now show that, by optimality, each of these allocations entails either $E^r = E$ or $E^r \in (0, \delta_i]$. First, $E^r = E$ when $E = 0$ (Assumption 3). Thus, truthful reporting is optimal for a non-working woman. When $0 < E \leq \delta_i$ consumption while on welfare depends on reported earnings as follows:

$$C_i^a(E, 1, E^r) = \begin{cases} E - \mu_i + \overline{G}_i - \phi_i - \kappa_i - \eta_i^a & \text{if } E^r = 0,\ E \in (0, \delta_i] \\ E - \mu_i + \overline{G}_i - \phi_i - \kappa_i & \text{if } 0 < E^r < E,\ E \in (0, \delta_i] \\ E - \mu_i + \overline{G}_i - \phi_i & \text{if } E^r = E,\ E \in (0, \delta_i] \end{cases}.$$

Thus, truthful reporting maximizes consumption (and hence utility) because $\kappa_i > 0, \eta_i^a \geq 0$. When $E > \delta_i$ and is in range 1, consumption while on welfare depends on reported earnings as follows:

$$C_i^a(E, 1, E^r) = \begin{cases} E - \mu_i + \overline{G}_i - \phi_i - \kappa_i - \eta_i^a & \text{if } E^r = 0,\ E \in (\delta_i, FPL_i] \\ E - \mu_i + \overline{G}_i - \phi_i - \kappa_i & \text{if } 0 < E^r \leq \delta_i,\ E \in (\delta_i, FPL_i] \\ E - \mu_i + G_i^a(E^r) - \phi_i - \kappa_i & \text{if } \delta_i < E^r < E,\ E \in (\delta_i, FPL_i] \\ E - \mu_i + G_i^a(E^r) - \phi_i & \text{if } E^r = E,\ E \in (\delta_i, FPL_i] \end{cases}.$$

Thus, only truthful reports or under-reports in $[0, \delta_i]$ are optimal since $G_i^a(.)$ is a decreasing function and $\kappa_i > 0, \eta_i^a \geq 0$ (Assumption 2). When $\eta_i^a > 0$ such a woman is indifferent among the reports in the interval $(0, \delta_i]$. When $\eta_i^a = 0$ she is indifferent among the reports in the interval $[0, \delta_i]$. When E is in range 2, women must be under-reporting (Lemma 1), hence consumption while on welfare depends on reported earnings as follows:

$$C_i^a(E, 1, E^r) = \begin{cases} E - \mu_i + \overline{G}_i - \phi_i - \kappa_i - \eta_i^a & \text{if } E^r = 0,\ E \text{ in range 2} \\ E - \mu_i + \overline{G}_i - \phi_i - \kappa_i & \text{if } 0 < E^r \leq \delta_i,\ E \text{ in range 2} \\ E - \mu_i + G_i^a(E^r) - \phi_i - \kappa_i & \text{if } \delta_i < E^r < E,\ E \text{ in range 2} \end{cases}.$$

Thus, reports below δ_i maximize consumption since $G_i^a(.)$ is a decreasing function and $\kappa_i \geq 0, \eta_i^a \geq 0$. If $\eta_i^a > 0$, reports in the interval $(0, \delta_i]$ are optimal, while if $\eta_i^a = 0$ reports in the interval $[0, \delta_i]$ are optimal. This establishes parts a) and b) of the Lemma under a.

It is straightforward to verify that, for any woman i, the consumption associated with optimally under-reporting is $E - \mu_i + \overline{G}_i - \phi_i - \kappa_i$ under either regime. Hence, a woman will only under-report if $E - \mu_i + \overline{G}_i - \phi_i - \kappa_i \geq C_i^t(E, 1, E)$ which occurs only when $\kappa_i \leq \overline{G}_i - G_i^t(E)$. Because $\overline{G}_i - G_i^j(E) = 0$ for any E in range 1, women will never choose state $1u$ under j. ■

Lemma 3 *Consider any pair of states (s^a, s^j) obeying: I) $s^j \neq s^a$; II) state s^a is no worse under JF than under AFDC; III) state s^j is no better under JF than under AFDC. Then, if Assumptions 3 and 4 hold, the response probability π_{s^a,s^j} equals zero.*

Proof. The proof is by contradiction. Suppose that $\pi_{s^a,s^j} > 0$ for some pair of states (s^a, s^j) satisfying properties I)-III). Then, there exists a woman i who chooses a triple (E, D, E^r) under a obeying $s_i(E, D, E^r) = s^a$ and a triple $(E', D', E^{r\prime})$ under j obeying $s_i(E', D', E^{r\prime}) = s^j$. By Property II) $U_i\left(E, C_i^j(E, D, E^r)\right) \geq U_i(E, C_i^a(E, D, E^r))$. The choice of state s_a under a reveals that $U_i(E, C_i^a(E, D, E^r)) \geq U_i(E', C_i^a(E', D', E^{r\prime}))$ (Assumption 3). By Property III) $U_i(E', C_i^a(E', D', E^{r\prime})) \geq U_i\left(E', C_i^j(E', D', E^{r\prime})\right)$. Combining these inequalities we have:

$$U_i\left(E, C_i^j(E, D, E^r)\right) \geq U_i(E, C_i^a(E, D, E^r)) \geq U_i(E', C_i^a(E', D', E^{r\prime})) \geq U_i\left(E', C_i^j(E', D', E^{r\prime})\right).$$

If any of the inequalities is strict, optimality of choice $(E', D', E^{r\prime})$ under j is contradicted. If no inequality is strict, woman i is indifferent between the two allocations $\left(E, C_i^j(E, D, E^r)\right)$ and $(E', C_i^a(E', D', E^{r\prime}))$ which contradicts her chosing the first allocation under a and the second under j (Assumption 4 and Property I). ■

Lemma 4 *Given Assumptions 1-4, the states in $\mathcal{C}_+$ are no worse under JF than under AFDC, the states in $\mathcal{C}_-$ are no better under JF than under AFDC, and the states in $\mathcal{C}_0$ are equally attractive under JF and AFDC.*

Proof. From Assumption 1, it is sufficient to verify that $C_i^j(E, D, E^r) \geq C_i^a(E, D, E^r)$ for all (E, D, E^r) such that $s_i(E, D, E^r) \in \mathcal{C}_+$, that $C_i^j(E, D, E^r) \leq C_i^a(E, D, E^r)$ for all (E, D, E^r) such that $s_i(E, D, E^r) \in \mathcal{C}_-$, and that $C_i^j(E, D, E^r) = C_i^a(E, D, E^r)$ for all (E, D, E^r) such that $s_i(E, D, E^r) \in \mathcal{C}_0$.

Start with a triple (E, D, E^r) obeying $s_i(E, D, E^r) \in \mathcal{C}_+$. Since $\mathcal{C}_+$ consists of state $1r$, this means that E is in range 1, $D = 1$, and $E^r = E$. Because $G_i^a(E) < \overline{G}_i$ for all E in range 1,

$$C_i^j(E, 1, E) = E - \mu_i + \overline{G}_i - \phi_i \geq E - \mu_i + G_i^a(E) - \phi_i = C_i^j(E, 1, E),$$

which verifies the desired inequality. Consider next a triple (E, D, E^r) obeying $s_i(E, D, E^r) \in \mathcal{C}_-$. Since $\mathcal{C}_-$ consists of state $0r$, this means that $E = E^r = 0$ and $D = 1$. Because $\eta_i^j \geq \eta_i^a$ (Assumption 2),

$$C_i^j(0, 1, 0) = \overline{G}_i - \phi_i - \eta_i^j \leq \overline{G}_i - \phi_i - \eta_i^a = C_i^a(0, 1, 0),$$

which verifies the desired inequality. Finally, consider a triple (E, D, E^r) obeying $s_i(E, D, E^r) \in \mathcal{C}_0$. If $s_i(E, D, E^r) \in \{0n, 1n, 2n\}$, consumption is unaffected by the regime (Assumption 3): it is either zero, when $s_i(E, D, E^r) = 0n$, or $E - \mu_i$, when $s_i(E, D, E^r) \in \{1n, 2n\}$. If $s_i(E, D, E^r) \in \{1u, 2u\}$, consumption is unaffected by the regime because optimal under-reporting yields a transfer of $\overline{G}_i$ under either regime (Lemma 2). Specifically, consumption is $E - \mu_i + \overline{G}_i - \phi_i - \kappa_i$ under either regime. ■

Main Propositions

For convenience we reproduce here the matrix Π:

State under AFDC	Earnings / Reporting State under JF						
	$0n$	$1n$	$2n$	$0r$	$1r$	$1u$	$2u$
$0n$	$1-\pi_{0n,1r}$	0	0	0	$\pi_{0n,1r}$	0	0
$1n$	0	$1-\pi_{1n,1r}$	0	0	$\pi_{1n,1r}$	0	0
$2n$	0	0	$1-\pi_{2n,1r}$	0	$\pi_{2n,1r}$	0	0
$0r$	$\pi_{0r,0n}$	0	$\pi_{0r,2n}$	$1-\pi_{0r,0n}-\pi_{0r,2n}$ $-\pi_{0r,1r}-\pi_{0r,2u}$	$\pi_{0r,1r}$	0	$\pi_{0r,2u}$
$1r$	0	0	0	0	1	0	0
$1u$	0	0	0	0	1	0	0
$2u$	0	0	0	0	$\pi_{2u,1r}$	0	$1-\pi_{2u,1r}$

.

Proposition 1 *Given Assumptions 1-4, the responses corresponding to the zero entries of matrix* Π *cannot occur and the responses corresponding to unitary entries of the matrix must occur.*

Proof. We begin with the zeros. State $1u$ is unpopulated under j (Lemma 2). Therefore $\pi_{s^a,1u} = 0$ for any $s^a \in \mathcal{S}$. Next, by Lemmas 3 and 4, the response probability π_{s^a,s^j} equals zero for all (s^a, s^j) in the collection:

$$\left\{(s^a, s^j) : s^a \in \mathcal{C}_0 \cup \mathcal{C}_+, s^j \in \mathcal{C}_0 \cup \mathcal{C}_-, s^a \neq s^j\right\}. \tag{A.1}$$

It suffices to show that properties I)-III) of Lemma 3 are met. Property I) holds trivially and properties II) and III) hold by Lemma 4. Therefore, the responses in (A.1) have probability zero.

We now show that $\pi_{0r,1n} = 0$. The proof is by contradiction. Suppose there is a woman i who selects an allocation compatible with state $0r$ under a and selects an allocation compatible with state $1n$ under j, entailing earnings O_k^i. By Assumption 3, her choice under a reveals that $U_i(0, \overline{G}_i - \phi_i - \eta_i^a) \geq U_i(0, 0)$ which implies $\overline{G}_i - \phi_i \geq \eta_i^a$. Her choice under j reveals that $U_i(O_i^k, O_i^k - \mu_i) \geq U_i(O_i^k, O_i^k - \mu_i + \overline{G}_i - \phi_i)$ which implies

$\overline{G}_i - \phi_i \leq 0$. Thus, $0 \leq \eta_i^a \leq \overline{G}_i - \phi_i \leq 0$ (Assumption 1). If $\eta_i^a > 0$ or $\eta_i^a = 0$ and $\overline{G}_i \neq \phi_i$, a contradiction ensues. If $\eta_i^a = 0$ and $\overline{G}_i = \phi_i$, the woman must be indifferent between the allocation compatible with state $0r$ and that compatible with $0n$ under a which means $U_i(0,0) \geq U_i\left(O_i^l, O_i^l - \mu_i\right)$ for any O_i^l in range 1 including O_i^k. The choice of the allocation compatible with state $1n$ under j reveals that $U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i(0,0)$. If this last inequality is strict a contradiction ensues. Otherwise, the woman must be indifferent under a among the allocation compatible with state $0n$, the allocation compatible with state $0r$, and the allocation entailing earnings O_i^k while off assistance. If however she did not choose O_i^k over $0r$ under a then she will make the same choice under j (Assumption 4), which implies a contradiction. Therefore $\pi_{0r,1n} = 0$. This concludes the proof of the zero entries in the matrix.

Turning to the unitary entries, by Lemma 1, the allowable states are given by $\mathcal{S} \equiv \{0n, 1n, 2n, 0r, 1r, 1u, 2u\}$. Hence, each row of matrix Π must sum to one. Therefore $\pi_{1r,1r} = 1$ and $\pi_{1u,1r} = 1$. ■

Proposition 2 *Given Assumptions 1-4, the "free" response probabilities in matrix* Π *given by* π_{s^a,s^j} *for all* $\left(s^a, s^j\right)$ *in the two collections:*

$$\left\{(s^a, 1r) : s^a \in \{0n, 1n, 2n, 2u\}\right\}, \tag{A.2}$$

$$\left\{\left(0r, s^j\right) : s^j \in \{0n, 2n, 1r, 2u\}\right\}, \tag{A.3}$$

are unrestricted, meaning they need not equal zero or one.

Proof. We start by considering the collection of state pairs in (A.2). The common feature of the states in $\{0n, 1n, 2n, 2u\}$ is that they are equally attractive under AFDC and JF. Instead, state $1r$ is no worse under AFDC than under JF. In light of Proposition 1, to prove that the response probabilities corresponding to the pairs of states in collection (A.2) need not equal zero or one it suffices to provide examples where two women occupy the same state $s^a \in \{0n, 1n, 2n, 2u\}$ under AFDC, but the first woman occupies state $s^j = s^a$ and the second woman occupies state $s_j = 1r$ under JF. We then turn to the collection of state pairs in (A.3). The common feature of the states in $\{0n, 2n, 1r, 2u\}$ is that they are no worse under AFDC than under JF. Instead, state $0r$ is no better under AFDC than under JF. To prove that the response probabilities corresponding to the pairs of states in collection (A.3) need not equal zero it suffices to provide examples of a woman who occupies state $0r$ under AFDC and state $s^j \in \{0n, 2n, 1r, 2u\}$ under JF. This also proves that these response probabilities corresponding to the pairs of states in (A.3) need not equal one because the rows of matrix Π sum to one, hence $\sum_{s \in \mathcal{S}} \pi_{0r,s} = 1$ and $\pi_{0r,s} > 0$ for any s implies $\pi_{0r,s'} < 1$ for all $s \neq s'$.

$\pi_{0n,1r}$ **is not restricted to zero or one**

Consider two women i' and i'' who both choose an allocation compatible with state $0n$ under a. Assume that each woman draws both earnings offers from range 1. Let woman i' have a non-positive net of stigma reward from assistance so that:

$$\overline{G}_{i'} - \phi_{i'} \leq 0. \tag{A.4}$$

Woman i'', by contrast, has a positive net of stigma reward from assistance obeying:

$$0 < \overline{G}_{i''} - \phi_{i''} \leq \eta_{i''}^a. \tag{A.5}$$

We now show that woman i' chooses an allocation compatible with state $0n$ under j while woman i'' may select an available allocation compatible with state $1r$ under j. For both women, the choice of the allocation compatible with state $0n$ under a reveals (Assumption 3) that this allocation yields as much utility as the available allocations compatible with states $\{0r, 1r, 1u, 1n\}$. Thus, for $i \in \{i', i''\}$:

$$U_i(0,0) \geq U_i\left(0, \overline{G}_i - \phi_i - \eta_i^a\right), \tag{A.6}$$

$$U_i(0,0) \geq U_i\left(O_i^k, O_i^k - \mu_i + G_i^a\left(O_i^k\right) - \phi_i\right) \ \forall O_i^k, \tag{A.7}$$

$$U_i(0,0) \geq U_i\left(O_i^k, O_i^k - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \ \forall O_i^k, \tag{A.8}$$

$$U_i(0,0) \geq U_i\left(O_i^k, O_i^k - \mu_i\right) \ \forall O_i^k. \tag{A.9}$$

Observe that (A.6) explicitly bounds from above the net of stigma reward from assistance:

$$\overline{G}_i - \phi_i \le \eta_i^a, \tag{A.10}$$

which agrees with both (A.4) and (A.5). Both women prefer state $0n$ under j to the available allocations compatible with states $\{1n, 0r, 1u\}$ (Proposition 1). Also, condition (A.4), (A.9) and Assumptions 1, 3, and 4 imply that woman i' prefers state $0n$ to the available allocations compatible with state $1r$ under j. Thus, woman i' occupies the same state under both regimes which proves that $\pi_{0n,1r}$ need not equal zero.

If woman i'' has an earnings draw $O_{i''}^l$ obeying:

$$U_{i''}(0,0) < U_{i''}\left(O_{i''}^l, O_{i''}^l - \mu_{i''} + \overline{G}_{i''} - \phi_{i''}\right), \tag{A.11}$$

she would have selected an allocation compatible with state $1r$ had the grant formula under regime a fully disregarded earnings (but not otherwise, as per (A.7)). As evidenced by (A.6)-(A.9), inequality (A.11) is enabled by (A.5) and $\overline{G}_{i''} \ge G_{i''}^a(E)\,\forall E$.[1] Condition (A.11) allows us to conclude that, under j, woman i'' prefers the available allocations compatible with $1r$ to those compatible with all other states. Thus, woman i'' occupies a different state under the two regimes which proves that $\pi_{0n,1r}$ need not equal one. Hence, $\pi_{0n,1r}$ is unrestricted.

$\pi_{1n,1r}$ is not restricted to zero or one

Consider two women i' and i'' who both choose an allocation compatible with state $1n$ under a. Assume that each woman draws both earnings offers from range 1. Let $O_{i'}^k$ and $O_{i''}^k$ denote the earnings offers chosen under a by woman i' and i'' respectively. Let woman i' have a non-positive net of stigma reward from assistance so that:

$$\overline{G}_{i'} - \phi_{i'} \le 0. \tag{A.12}$$

Woman i'', by contrast, has a positive net of stigma reward from assistance obeying:

$$0 < \overline{G}_{i''} - \phi_{i''} \le \min\left\{\overline{G}_{i''} - G_{i''}^a\left(O_{i''}^k\right), \kappa_{i''}\right\}. \tag{A.13}$$

We now show that woman i' chooses an allocation compatible with state $1n$ under j while woman i'' may select an available allocation compatible with state $1r$ under j. For both women, the choice of the allocation compatible with state $1n$ under a reveals (Assumption 3) that this allocation yields as much utility as the available allocations compatible with states $\{0n, 0r, 1r, 1u\}$ as well as the other available allocations compatible with state $1n$. Formally, for $i \in \{i', i''\}$:

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \ge U_i(0,0), \tag{A.14}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \ge U_i\left(0, \overline{G}_i - \phi_i - \eta_i^a\right), \tag{A.15}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \ge U_i\left(O_i^l, O_i^l - \mu_i\right)\forall O_i^l, \tag{A.16}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \ge U_i\left(O_i^l, O_i^l - \mu_i + G_i^a\left(O_l^i\right) - \phi_i\right)\ \forall O_i^l, \tag{A.17}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \ge U_i\left(O_i^l, O_i^l - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right)\ \forall O_i^l. \tag{A.18}$$

Observe that (A.17) and (A.18) evaluated at $O_i^l = O_i^k$ explicitly bound from above the net of stigma reward from assistance (Assumption 1):

$$\overline{G}_i - \phi_i \le \kappa_i,\ G_i^a\left(O_i^k\right) - \phi_i \le 0, \tag{A.19}$$

[1] Suppose, for example, that $U_{i''}(E,C) = C - E$ and that $G_{i''}^a\left(O_{i''}^k\right) - \phi_{i''} \le \mu_{i''} < \overline{G}_{i''} - \phi_{i''} \le \min\{\mu_{i''} + \kappa_{i''}, \eta_{i''}^a\}\,\forall O_{i''}^k$ which is enabled by $\overline{G}_{i''} \ge G_{i''}^a(E)\,\forall E > 0$ and $\kappa_{i''} > 0$ (Assumption 1). It is easy to check that for this woman the conditions in (A.6)-(A.9), associated with occupying state $0n$ under a, and condition (A.11), associated with occupying state $1r$ under j, hold.

which agrees with (A.12) and (A.13). Both women still prefer earning O_i^k off assistance (state $1n$) under j to the available allocations compatible with states $\{0r, 0n, 1u\}$ (Proposition 1) as well as to the other available allocation compatible with state $1n$. Condition (A.12) and (A.16), and Assumptions 1, 3, and 4, imply that woman i' also still prefers state $1n$ under j to the available allocations compatible with state $1r$. Thus, woman i' occupies the same state under both regimes which proves that $\pi_{1n,1r}$ need not equal zero. If woman i'''s utility function is strictly increasing in consumption, and by condition (A.13) and Assumptions 1, 3, and 4, woman i'' prefers earning $O_{i''}^k$ on assistance to earning the same amount off assistance under j. Hence, the available allocation entailing earnings $O_{i''}^k$ on assistance is preferred under j to the available allocations compatible with all states but $1r$. Thus, woman i'' occupies different states under the two regimes which proves that $\pi_{1n,1r}$ need not equal one. Hence, $\pi_{1n,1r}$ is unrestricted.

$\pi_{2n,1r}$ **is not restricted to zero or one**

Consider two women i' and i'' who both choose an allocation compatible with state $2n$ under a. Assume that each woman draws an earnings offer in range 1 and another offer in range 2. Let $O_{i'}^k$ and $O_{i''}^k$ denote the earnings offer in range 2 drawn by woman i' and i'' respectively. Likewise, denote woman i' and i'''s earnings draw in range 1 by $O_{i'}^m$ and $O_{i''}^m$ respectively. Let woman i' have a non-positive net of stigma reward from assistance so that:

$$\overline{G}_{i'} - \phi_{i'} \leq 0. \tag{A.20}$$

Woman i'', by contrast, has a positive net of stigma reward from assistance obeying:

$$0 < \overline{G}_{i''} - \phi_{i''} \leq \kappa_{i''}. \tag{A.21}$$

We now show that woman i' chooses an allocation compatible with state $2n$ under j while woman i'' may select an available allocation compatible with state $1r$ under j. For both women, the choice of the allocation compatible with state $2n$ under a reveals (Assumption 3) that this allocation yields as much utility as the available allocations compatible with states $\{0n, 0r, 1n, 1r, 1u, 2u\}$. Formally, for $i \in \{i', i''\}$:

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i\left(0, 0\right), \tag{A.22}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i\left(0, \overline{G}_i - \phi_i - \eta_i^a\right), \tag{A.23}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i\left(O_i^m, O_i^m - \mu_i\right), \tag{A.24}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i\left(O_i^m, O_i^m - \mu_i + G_i^a\left(O_i^m\right) - \phi_i\right), \tag{A.25}$$

$$U_i\left(O_i^k, O_i^k - \mu_i\right) \geq U_i\left(O_i^l, O_i^l - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \ \forall O_i^l. \tag{A.26}$$

Observe that (A.26) explicitly bounds from above the net of stigma rewards from assistance:

$$\overline{G}_i - \phi_i \leq \kappa_i, \tag{A.27}$$

which agrees with (A.20) and (A.21). Under j both women still prefer earning O_i^k off assistance (state $2n$) to the available allocations compatible with states $\{0r, 2u, 0n, 1n, 1u\}$ (Proposition 1). This fact, (A.20), (A.24) and Assumptions 1, 3, and 4 imply that woman i' still prefers the allocation compatible with state $2n$ to the available allocation compatible with state $1r$. Thus, woman i' occupies the same state under both regimes which proves that $\pi_{2n,1r}$ need not equal zero. If woman i'''s earnings offer in range 1 obeys:

$$U_{i''}\left(O_{i''}^m, O_{i''}^m - \mu_{i''} + \overline{G}_{i''} - \phi_{i''}\right) > U_{i''}\left(O_{i''}^k, O_{i''}^k - \mu_{i''}\right). \tag{A.28}$$

she would have selected the allocation compatible with state $1r$ had the grant formula under regime a fully disregarded earnings (but not otherwise, as per (A.25)). As evidenced by (A.24)-(A.26), inequality (A.28) may hold because of (A.21) and $\overline{G}_{i''} \geq G^a_{i''}(E)\,\forall E$ in range 1.[2] Together with (A.22)-(A.26), (A.28) allows us to conclude that under j woman i'' prefers the available allocation entailing truthfully reporting earning $O^m_{i''}$ on assistance to the available allocations compatible with all states. Thus, woman i'' occupies a different state under the two regimes which proves that $\pi_{2n,1r}$ need not equal one. Hence, $\pi_{2n,1r}$ is unrestricted.

$\pi_{2u,1r}$ **is not restricted to zero or one**

Consider two women i' and i'' who both choose an allocation compatible with state $2u$ under a. Assume that each woman draws an earning offer in range 1 and another offer in range 2. Let $O^k_{i'}$ and $O^k_{i''}$ denote the earnings offer in range 2 drawn by woman i' and i'' respectively. Likewise, denote woman i' and i'''s earnings offer in range 1 by $O^m_{i'}$ and $O^m_{i''}$ respectively. Let woman i' have a positive net of stigma reward from assistance obeying:

$$\kappa_{i'} \leq \overline{G}_{i'} - \phi_{i'} < \mu_{i'} - FPL_{i'}. \tag{A.29}$$

Woman i'' also has a positive net of stigma reward from assistance obeying:

$$\max\{\kappa_{i''}, \mu_{i''} - FPL_{i''}\} \leq \overline{G}_{i''} - \phi_{i''}. \tag{A.30}$$

We now show that woman i' may select an allocation compatible with state $2u$ under j while woman i'' may select an available allocation compatible with state $1r$ under j. For both women, the choice of the allocation compatible with state $2u$ under a reveals (Assumption 3) that this allocation yields as much utility as the available allocations compatible with states $\{0n, 0r, 1n, 2n, 1r, 1u\}$. Formally, for $i \in \{i', i''\}$:

$$U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \geq U_i(0,0), \tag{A.31}$$

$$U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \geq U_i\left(0, \overline{G}_i - \phi_i - \eta^a_i\right), \tag{A.32}$$

$$U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \geq U_i\left(O^l_i, O^l_i - \mu_i\right)\ \forall O^l_i, \tag{A.33}$$

$$U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \geq U_i\left(O^m_i, O^m_i - \mu_i + G^a_i\left(O^m_i\right) - \phi_i\right), \tag{A.34}$$

$$U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right) \geq U_i\left(O^m_i, O^m_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right). \tag{A.35}$$

Observe that (A.33) explicitly bounds from below the net of stigma reward from assistance, namely,

$$\kappa_i \leq \overline{G}_i - \phi_i, \tag{A.36}$$

which agrees with both (A.29) and (A.30). Both women still prefer state $2u$ under j to the available allocations compatible with states $\{0r, 0n, 1n, 2n, 1u\}$ (Proposition 1). Condition (A.29) implies that $\overline{G}_{i'} - \phi_{i'} < \mu_{i'} - FPL_{i'} \leq \mu_{i'} - O^m_{i'}$ because $O^m_{i'}$ is in range 1, hence $O^m_{i'} - \mu_{i'} + \overline{G}_{i'} - \phi_{i'} < 0$. Assumption 1 then implies that woman i' prefers the allocation compatible with state $0n$ to the available allocation compatible with state $1r$. By (A.31), this means that she still prefers the allocation compatible with state $2u$ to the available allocation compatible with state $1r$. Hence she prefers state $2u$ under j to the allocations available and compatible with all other states. Thus, woman i' occupies the same state under both regimes which proves that $\pi_{2u,2u}$ need not equal zero. If woman i'''s earnings offer in range 1 obeys:

$$U_{i''}\left(O^m_{i''}, O^m_{i''} - \mu_{i''} + \overline{G}_{i''} - \phi_{i''}\right) > U_{i''}\left(O^k_{i''}, O^k_{i''} - \mu_{i''} + \overline{G}_{i''} - \phi_{i''} - \kappa_{i''}\right), \tag{A.37}$$

she would have selected the allocation compatible with state $1r$ had the grant formula under regime a fully disregarded earnings (but not otherwise, as per (A.34)). As evidenced by (A.35), inequality (A.37) is enabled

[2] Suppose, for example, that $U_{i''}(E,C) = C - E$, $\mu_{i''} = 0$, $\eta^a_{i''} > 0$, and $G^a_{i''}\left(O^m_{i''}\right) - \phi_{i''} \leq 0 < \overline{G}_{i''} - \phi_{i''} \leq \min\{\kappa_{i''}, \eta^a_{i''}\}$ which may hold because $\overline{G}_{i''} \geq G^a_{i''}(E)\,\forall E$ in range 1. It is easy to check that condition (A.28), associated with occupying state $1r$ under j, and the conditions in (A.22)-(A.26), associated with occupying state $2n$ under a, all hold.

by $\kappa_{i''} > 0$ (Assumption 3) and it can agree with (A.34) because $\overline{G}_{i''} \geq G^a_{i''}(E)\,\forall E$ in range 1.[3] In such a case, under j woman i'' prefers the available allocation entailing truthfully reporting earning $O^m_{i''}$ on assistance to the available allocations compatible with all states. Thus, woman i'' occupies a different state under the two regimes which proves that $\pi_{2u,1r}$ need not equal one. Hence, $\pi_{2u,1r}$ is unrestricted.

$(\pi_{0r,0r}, \pi_{0r,1r}, \pi_{0r,0n}, \pi_{0r,2n}, \pi_{0r,2u})$ **are not restricted to zero or one**

Consider five women $i', i'', i^{III}, i^{IV}, i^{V}$ who all choose an allocation compatible with state $0r$ under a. Assume that each woman draws an earnings offer in range 1 and another in range 2. Let women i' and i'' have identical hassle from not working on assistance under both regimes, $\eta^a_i = \eta^j_i$, and a sufficiently large net of stigma reward from assistance obeying:

$$\eta^a_i \leq \overline{G}_i - \phi_i. \tag{A.38}$$

Women $i = \left\{i^{III}, i^{IV}, i^{V}\right\}$, by constrast, have strictly larger hassle from not working on assistance under j than a, $\eta^a_i < \eta^j_i$, and a net of stigma reward from assistance obeying:

$$\eta^a_{i^{III}} \leq \overline{G}_{i^{III}} - \phi_{i^{III}} \leq \min\left\{\kappa_{i^{III}}, \eta^j_{i^{III}}\right\}, \tag{A.39}$$

$$\eta^j_{i^{IV}} \leq \overline{G}_{i^{IV}} - \phi_{i^{IV}} \leq \kappa_{i^{IV}}, \tag{A.40}$$

$$\max\left\{\kappa_{i^{V}}, \eta^j_{i^{V}}\right\} \leq \overline{G}_{i^{V}} - \phi_{i^{V}}. \tag{A.41}$$

We now show that woman i' may select an allocation compatible with state $0r$ under j, woman i'' may select an available allocation compatible with state $1r$ under j, woman i^{III} may select an available allocation compatible with state $0n$ under j, woman i^{IV} may select an available allocation compatible with state $2n$ under j, and woman i^{V} may select an available allocation compatible with state $2u$ under j. For all women, the choice of the allocation compatible with state $0r$ under a reveals that such allocation yields as much utility as the available allocations compatible with all the other states, namely $\{0n, 1n, 2n, 1r, 1u, 2u\}$. Formally, for $i \in \left\{i', i'', i^{III}, i^{IV}, i^{V}\right\}$:

$$U_i\left(0, \overline{G}_i - \phi_i - \eta^a_i\right) \geq U_i\left(0, 0\right), \tag{A.42}$$

$$U_i\left(0, \overline{G}_i - \phi_i - \eta^a_i\right) \geq U_i\left(O^k_i, O^k_i - \mu_i\right)\ \forall O^k_i, \tag{A.43}$$

$$U_i\left(0, \overline{G}_i - \phi_i - \eta^a_i\right) \geq U_i\left(O^k_i, O^k_i - \mu_i + G^a_i\left(O^k_i\right) - \phi_i\right)\ \text{for } O^k_i \text{ in range 1}, \tag{A.44}$$

$$U_i\left(0, \overline{G}_i - \phi_i - \eta^a_i\right) \geq U_i\left(O^k_i, O^k_i - \mu_i + \overline{G}_i - \phi_i - \kappa_i\right)\ \forall O^k_i. \tag{A.45}$$

Inequality (A.42) explicilty bounds from below the net of stigma reward from assistance, namely,

$$\eta^a_i \leq \overline{G}_i - \phi_i, \tag{A.46}$$

which can agree with (A.38), (A.39), (A.40), and (A.41). In such a case, because $\eta^a_i = \eta^j_i$ for $i = i', i''$, state $0r$ has the same utility value under both regimes hence both women still prefer $0r$ under j to the available allocations compatible with states $\{0n, 1n, 2n, 1u, 2u\}$ (Lemma 4). If woman i' satisfies the additional requirement that:

$$U_{i'}\left(0, \overline{G}_{i'} - \phi_{i'} - \eta^a_{i'}\right) \geq U_i\left(O^k_{i'}, O^k_{i'} - \mu_{i'} + \overline{G}_{i'} - \phi_{i'}\right)\ \text{for } O^k_{i'} \text{ in range 1}, \tag{A.47}$$

which implies (A.44) because $G^a_{i'}(E) \leq \overline{G}_{i'}$ for all E in range 1.[4] In such a case woman i' still prefers $0r$ under j to the available allocation compatible with state $1r$, this proves that $\pi_{0r,0r}$ need not equal zero. If woman i'' has an earnings draw $O^k_{i''}$ in range 1 such that:

$$U_{i''}\left(O^k_{i''}, O^k_{i''} - \mu_{i''} + \overline{G}_{i''} - \phi_{i''}\right)\{1r, 2n\}\, under j. > U_{i''}\left(0, \overline{G}_{i''} - \phi_{i''} - \eta^a_{i''}\right), \tag{A.48}$$

[3]Suppose, for example, that $U_{i''}(E, C) = C - E$. Then (A.37) always holds since it requires $\kappa_{i''} > 0$ (Assumption 2).

[4]Suppose, for example, that $U_{i'}(E, C) = C - E$. Then (A.47) requires $\eta^a_{i'} < \mu_{i'}$ which agrees with the condition for optimality of $0r$ under a, namely (A.42)-(A.45).

which can agree with (A.44) because $G^a_{i''}(E) \leq \overline{G}_{i''}$ for all E in range 1.[5] In such a case, woman i'' prefers earning $O^k_{i''}$ compatible with state $1r$ under j to state $0r$, this proves that $\pi_{0r,1r}$ need not equal zero.

Consider now women i^{III}, i^{IV}, i^V. By Proposition 1, none of these women will occupy states $\{1n, 1u\}$ under j. Let woman i^{III} prefer non-employment off assistance to the available allocations compatible with states $\{1r, 2n\}$ under j. Formally,

$$U_{i^{III}}(0,0) \geq U_{i^{III}}\left(O^k_{i^{III}}, O^k_{i^{III}} - \mu_{i^{III}} + \overline{G}_{i^{III}} - \phi_{i^{III}}\right) \text{ for } O^k_{i^{III}} \text{ in range 1,} \tag{A.49}$$

$$U_{i^{III}}(0,0) \geq U_{i^{III}}\left(O^k_{i^{III}}, O^k_{i^{III}} - \mu_{i^{III}}\right) \text{ for } O^k_{i^{III}} \text{ in range 2.} \tag{A.50}$$

Then, by the upper bound $\overline{G}_{i^{III}} - \phi_{i^{III}} \leq \eta^j_{i^{III}}$ in (A.39), woman i^{III} also prefers non-employment off assistance to state $0r$ under j (which, thanks to $\eta^j_{i^{III}} > \eta^a_{i^{III}}$, is compatible with (A.42) hence with the optimality of $0r$ under a).[6] By the upper bound $\overline{G}_{i^{III}} - \phi_{i^{III}} \leq \kappa_{i^{III}}$ in (A.39) and (A.50), woman i^{III} also prefers non-employment off assistance to the allocations compatible with state $2u$ under j. In summary, woman i^{III} prefers an allocation compatible with state $0n$ under j to the allocations compatible with all the other states. This shows that $\pi_{0r,0n}$ need not equal zero.

Let woman i^{IV} have an earnings draw $O^k_{i^{IV}}$ in range 2 such that under j she prefers earning $O^k_{i^{IV}}$ off assistance to the available allocations compatible with states $\{0r, 1r\}$. Formally,

$$U_{i^{IV}}\left(O^k_{i^{IV}}, O^k_{i^{IV}} - \mu_{i^{IV}}\right) > U_{i^{IV}}\left(0, \overline{G}_{i^{IV}} - \phi_{i^{IV}} - \eta^j_{i^{IV}}\right), \tag{A.51}$$

$$U_{i^{IV}}\left(O^k_{i^{IV}}, O^k_{i^{IV}} - \mu_{i^{IV}}\right) \geq U_{i^{IV}}\left(O^l_{i^{IV}}, O^l_{i^{IV}} - \mu_{i^{IV}} + \overline{G}_{i^{IV}} - \phi_{i^{IV}}\right) \text{for } O^l_{i^{IV}} \text{ in range 1,} \tag{A.52}$$

where, because $\eta^j_{i^{IV}} > \eta^a_{i^{IV}}$, (A.51) is compatible with (A.42), and hence with the optimality of $0r$ under a.[7] Then, under j, by the lower bound $\overline{G}_{i^{IV}} - \phi_{i^{IV}} > \eta^j_{i^{IV}}$ in (A.40) and (A.51), woman i^{IV} also prefers $O^k_{i^{IV}}$ off assistance to state $0n$. By the lower bound $\overline{G}_{i^{IV}} - \phi_{i^{IV}} > 0$ implicit in (A.40) and (A.52), woman i^{IV} also prefers $O^k_{i^{IV}}$ off assistance to the allocation compatible with state $1n$ under j. By the upper bound $\overline{G}_{i^{IV}} - \phi_{i^{IV}} \leq \kappa_{i^{IV}}$ in (A.40), woman i^{IV} also prefers $O^k_{i^{IV}}$ off assistance to the allocation compatible with state $2u$ under j. In summary, woman i^{IV} prefers $O^k_{i^{IV}}$ off assistance under j to the allocations compatible with all the other states. This shows that $\pi_{0r,2n}$ need not equal zero.

Let woman i^V have an earnings draw $O^k_{i^V}$ in range 2 such that under j she prefers earning and misreporting $O^k_{i^V}$ to the available allocations compatible with states $\{0r, 1r\}$. Formally,

$$U_{i^V}\left(O^k_{i^V}, O^k_{i^V} - \mu_{i^V} + \overline{G}_{i^V} - \phi_{i^V} - \kappa_{i^V}\right) > U_{i^V}\left(0, \overline{G}_{i^V} - \phi_{i^V} - \eta^j_{i^V}\right), \tag{A.53}$$

$$U_{i^V}\left(O^k_{i^V}, O^k_{i^V} - \mu_{i^V} + \overline{G}_{i^V} - \phi_{i^V} - \kappa_{i^V}\right) \geq U_{i^V}\left(O^l_{i^V}, O^l_{i^V} - \mu_{i^V} + \overline{G}_{i^V} - \phi_{i^V}\right) \text{for } O^l_{i^V} \text{ in range 1.} \tag{A.54}$$

where, because $\eta^j_{i^V} > \eta^a_{i^V}$, (A.53) is compatible with (A.42), and hence with the optimality of $0r$ under a.[8] In such a case, under j, by the lower bound $\overline{G}_{i^V} - \phi_{i^V} \geq \eta^j_{i^V}$ in (A.41) and (A.53), woman i^V also prefers

[5]Suppose, for example, that $U_{i''}(E, C) = C - E$. Then (A.48) requires $\eta^a_{i'} > \mu_{i'}$ which agrees with the conditions for optimality of $0r$ under a, namely (A.42)-(A.45).

[6]Suppose, for example, that $U_{i^{III}}(E, C) = C - E$. Then (A.49)-(A.50) require $\mu_{i^{III}} \geq \overline{G}_{i^{III}} - \phi_{i^{III}}$ which agrees with the condition for optimality of $0r$ under a, namely (A.42)-(A.45) since $\overline{G}_{i^{III}} - \phi_{i^{III}} \geq \eta^a_{i^{III}}$ by our characterization of woman i^{III} in (A.39).

[7]Suppose, for example, that $U_{i^{iv}}(E, C) = C - \beta_{i^{iv}} E$ with $\beta_{i^{iv}} \in (0, 1)$. Then (A.51)-(A.52) bound from above the net of stigma reward from assistance: $O^k_{i^{IV}}(1 - \beta_{i^{iv}}) - \mu_{i^{IV}} + \eta^j_{i^{IV}} > \overline{G}_{i^{IV}} - \phi_{i^{IV}}$ and $\left(O^k_{i^{IV}} - O^l_{i^{IV}}\right)(1 - \beta_{i^{iv}}) \geq \overline{G}_{i^{IV}} - \phi_{i^{IV}}$ for $O^k_{i^{IV}}$ in range 2 and $O^l_{i^{IV}}$ in range 1. Given our characterization of woman i^{IV}, namely, $\eta^j_{i^{IV}} \leq \overline{G}_{i^{IV}} - \phi_{i^{IV}}$ from (A.40), this requires $O^k_{i^{IV}}(1 - \beta_{i^{iv}}) - \mu_{i^{IV}} \leq 0$ and $\left(O^k_{i^{IV}} - O^l_{i^{IV}}\right)(1 - \beta_{i^{iv}}) \geq \eta^j_{i^{IV}}$ which agree with the conditions for optimality of $0r$ under a, namely (A.42)-(A.45).

[8]Suppose, for example, that $U_{i^v}(E, C) = C - \beta_{i^v} E$ with $\beta_{i^v} \in (0, 1)$. Then (A.53)-(A.54) bound from above the cost of under-reporting: $O^k_{i^V}(1 - \beta_{i^v}) - \mu_{i^{IV}} + \eta^j_{i^V} > \kappa_{i^V}$ and $\left(O^k_{i^V} - O^l_{i^V}\right)(1 - \beta_{i^v}) > \kappa_{i^V}$ for $O^k_{i^V}$ in range 2 and $O^l_{i^V}$ in range 1. These bounds agree with the conditions for optimality of $0r$ under a, namely (A.42)-(A.45).

misreporting $O^k_{i^V}$ to state $0n$. By the lower bound $\overline{G}_{i^V} - \phi_{i^V} \geq 0$ implicit in (A.41) and (A.54), woman i^{IV} also prefers misreporting $O^k_{i^V}$ to the available allocation compatible with state $1n$. By the lower bound $\overline{G}_{i^V} - \phi_{i^V} \geq \kappa_{i^V}$ in (A.41) woman i^V also prefers misreporting $O^k_{i^V}$ to the allocation compatible with state $2n$. In summary, woman i^V prefers misreporting $O^k_{i^V}$ under j to the allocations compatible with all the other states. This shows that $\pi_{0r,2u}$ need not equal zero. ■

2 Propensity Score Reweighting

We use propensity score reweighting methods to adjust for the chance imbalances in baseline characteristics between the AFDC and JF groups. Following BGH (2006) we estimate a logit of the JF assignment dummy on: quarterly earnings in each of the 8 pre-assignment quarters, separate variables representing quarterly AFDC and quarterly food stamps payments in each of the 7 pre-assignment quarters, dummies indicating whether each of these 22 variables is nonzero, and dummies indicating whether the woman was employed at all or on welfare at all in the year preceding random assignment or in the applicant sample. We also include dummies indicating each of the following baseline demographic characteristics: being white, black, or Hispanic; being never married or separated; having a high-school diploma/GED or more than a high-school education; having more than two children; being younger than 25 or age 25-34; and dummies indicating whether baseline information is missing for education, number of children, or marital status.

Denote the predicted values from this model by $\widehat{p}_i$. The propensity score weights used to adjust the moments of interest are given by:

$$\omega_i = \frac{\frac{1[T_i=j]}{\widehat{p}_i}}{\sum\limits_n \frac{1[T_n=j]}{\widehat{p}_n}} + \frac{\frac{1-1[T_i=j]}{1-\widehat{p}_i}}{\sum\limits_n \frac{1-1[T_n=j]}{1-\widehat{p}_n}}.$$

These are inverse probability weights, re-normalized to sum to one within policy group. When examining subgroups we always recompute a new set of propensity score weights and re-normalize them.

3 Kolmogorov-Smirnov Tests for Equality of Distribution

We use a bootstrap procedure to compute the p-values for our reweighted Kolmogorov-Smirnov (K-S) tests for equality of distribution functions across treatment groups. Let $F_n^t(e)$ be the propensity score reweighted EDF of earnings in treatment group t. That is,

$$F_n^t(e) \equiv \sum_i \omega_i 1\left[E_i \leq e, T_i = t\right].$$

Define the corresponding bootstrap EDF as:

$$F_n^{t*}(e) \equiv \sum_i \omega_i^* 1\left[E_i^* \leq e, T_i^* = t\right].$$

where stars refer to resampled values (we resampled at the case level in order to preserve serial correlation in the data). The K-S test statistic is given by:

$$\widehat{KS} \equiv \sup_e |F_n^j(e) - F_n^a(e)|.$$

To obtain a critical value for this statistic, we compute the bootstrap distribution of the *recentered* K-S statistic:

$$KS^* \equiv \sup_e |F_n^{j*}(e) - F_n^{a*}(e) - \left(F_n^j(e) - F_n^a(e)\right)|.$$

Recentering is necessary to impose the correct null hypothesis on the bootstrap DGP (Giné and Zinn, 1990). We compute an estimated p-value $\widehat{\alpha}$ for the null hypothesis that the two distributions are equal as:

$$\widehat{\alpha} \equiv \frac{1}{1000} \sum_{b=1}^{1000} 1 \left[KS^{*(b)} > \widehat{KS} \right],$$

where b indexes the bootstrap replication.

4 Bounds on the Response Probabilities

List of Bounds

The analytical expressions for the bounds on the response probabilities are:

$$\pi_{2n,1r} \geq \max \left\{ 0, \frac{p_{2n}^a - p_{2n}^j}{p_{2n}^a} \right\},$$

$$\pi_{2n,1r} \leq \min \left\{ \begin{array}{c} 1, \frac{p_{2n}^a - p_{2n}^j + p_{0p}^a - p_{0p}^j}{p_{2n}^a}, \frac{p_{2n}^a - p_{2n}^j + p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j}{p_{2n}^a}, \\ \frac{p_{2n}^a - p_{2n}^j + p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j}{p_{2n}^a}, \frac{p_{2n}^a - p_{2n}^j + p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j + p_{2p}^a - p_{2p}^j}{p_{2n}^a} \end{array} \right\},$$

$$\pi_{0n,1r} \geq \max \left\{ 0, \frac{p_{0n}^a - p_{0n}^j}{p_{0n}^a} \right\},$$

$$\pi_{0n,1r} \leq \min \left\{ \begin{array}{c} 1, \frac{p_{0n}^a - p_{0n}^j + p_{0p}^a - p_{0p}^j}{p_{0n}^a}, \frac{p_{0n}^a - p_{0n}^j + p_{0p}^a - p_{0p}^j + p_{2n}^a - p_{2n}^j}{p_{0n}^a}, \\ \frac{p_{0n}^a - p_{0n}^j + p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j}{p_{0n}^a}, \frac{p_{0n}^a - p_{0n}^j + p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j + p_{2n}^a - p_{2n}^j}{p_{0n}^a} \end{array} \right\},$$

$$\pi_{2u,1r} \geq \max \left\{ 0, \frac{p_{2p}^a - p_{2p}^j}{p_{2p}^a} \right\},$$

$$\pi_{2u,1r} \leq \min \left\{ \begin{array}{c} 1, \frac{p_{2p}^a - p_{2p}^j + p_{0p}^a - p_{0p}^j}{p_{2p}^a}, \frac{p_{2p}^a - p_{2p}^j + p_{0p}^a - p_{0p}^j + p_{2n}^a - p_{2n}^j}{p_{2p}^a}, \\ \frac{p_{2p}^a - p_{2p}^j}{p_{2p}^a} + \frac{p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j}{p_{2p}^a}, \frac{p_{2p}^a - p_{2p}^j + p_{0p}^a - p_{0p}^j + p_{2n}^a - p_{2n}^j + p_{0n}^a - p_{0n}^j}{p_{2p}^a} \end{array} \right\},$$

$$\pi_{0r,1r} \geq \max \left\{ 0, \frac{p_{0p}^a - p_{0p}^j - p_{0n}^j - p_{2n}^j - p_{2p}^j}{p_{0p}^a} \right\},$$

$$\pi_{0r,1r} \leq \min \left\{ \begin{array}{c} \frac{p_{0p}^a - p_{0p}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{2n}^a - p_{2n}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j}{p_{0p}^a}, \\ \frac{p_{0p}^a - p_{0p}^j + p_{2n}^a - p_{2n}^j + p_{2p}^a - p_{2p}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j + p_{2n}^a - p_{2n}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j + p_{0n}^a - p_{0n}^j}{p_{0p}^a}, \\ \frac{p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j + p_{2n}^a - p_{2n}^j + p_{2p}^a - p_{2p}^j}{p_{0p}^a} \end{array} \right\},$$

$$\pi_{0r,2n} \geq \max \left\{ 0, \frac{-\left(p_{2n}^a - p_{2n}^j\right)}{p_{0p}^a} \right\},$$

$$\pi_{0r,2n} \leq \min \left\{ \begin{array}{c} \frac{p_{2n}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j}{p_{0p}^a}, \frac{p_{0p}^a - p_{0p}^j + p_{0n}^a - p_{0n}^j}{p_{0p}^a}, \\ \frac{p_{0p}^a - p_{0p}^j + p_{2p}^a - p_{2p}^j + p_{0n}^a - p_{0n}^j}{p_{0p}^a} \end{array} \right\},$$

$$\pi_{0r,2u} \geq \max\left\{0, \frac{-\left(p^a_{2p} - p^j_{2p}\right)}{p^a_{0p}}\right\},$$

$$\pi_{0r,2u} \leq \min\left\{\begin{array}{c} \frac{p^j_{2p}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}+p^a_{2n}-p^j_{2n}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}+p^a_{0n}-p^j_{0n}}{p^a_{0p}}, \\ \frac{p^a_{0p}-p^j_{0p}+p^a_{2n}-p^j_{2n}+p^a_{0n}-p^j_{0n}}{p^a_{0p}} \end{array}\right\},$$

$$\pi_{0r,0n} \geq \max\left\{0, \frac{-\left(p^a_{0n} - p^j_{0n}\right)}{p^a_{0p}}\right\},$$

$$\pi_{0r,0n} \leq \min\left\{\begin{array}{c} \frac{p^j_{0n}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}+p^a_{2p}-p^j_{2p}}{p^a_{0p}}, \frac{p^a_{0p}-p^j_{0p}+p^a_{2n}-p^j_{2n}}{p^a_{0p}}, \\ \frac{p^a_{0p}-p^j_{0p}+p^a_{2p}-p^j_{2p}+p^a_{2n}-p^j_{2n}}{p^a_{0p}} \end{array}\right\}.$$

Derivation of Bounds

A solution to any linear programming problem has to occur at one of the vertices of the problem's constraint space (see Murty, 1983). Recall that the linear constraints are:

$$\begin{aligned}
p^j_{0n} - p^a_{0n} &= -p^a_{0n}\pi_{0n,1r} + p^a_{0p}\pi_{0r,0n} \\
p^j_{1n} - p^a_{1n} &= -p^a_{1n}\pi_{1n,1r} \\
p^j_{2n} - p^a_{2n} &= -p^a_{2n}\pi_{2n,1r} + p^a_{0p}\pi_{0r,2n} \\
p^j_{0p} - p^a_{0p} &= -p^a_{0p}\left(\pi_{0r,1r} + \pi_{0r,2u} + \pi_{0r,2n} + \pi_{0r,0n}\right) \\
p^j_{2p} - p^a_{2p} &= p^a_{0p}\pi_{0r,2u} - p^a_{2p}\pi_{2u,1r}
\end{aligned} \tag{A.55}$$

To obtain the set of possible solutions to the linear programming problem

$$\max_{\pi} \lambda'\pi \text{ subject to (A.55) and } \pi \in [0,1]^7,$$

we enumerated all vertices of the convex polytope defined by the intersection of the hyperplane defined by the equations in $(A.55)$ with the hypercube defined by the unit constraints on the parameters. In practice, this amounted to setting all possible choices of three of the seven parameters in $(A.55)$ to 0 or 1 and solving for the remaining four parameters. There were $\binom{7}{3} = 35$ different possible choices of three parameters and $2^3 = 8$ different binary arrangements those parameters could take, yielding 280 possible vertices. However we were able to use the structure of our problem to rule out the existence of solutions at certain vertices – e.g., $\pi_{2n,1r}$ and $\pi_{0r,2n}$ cannot both be set arbitrarily because this would lead to a violation of the second equation in $(A.55)$. Such restrictions reduced the problem to solving the system at 160 vertices. We then enumerated the set of minima and maxima each parameter could achieve across the 160 relevant solutions. After eliminating dominated solutions, we arrived at the stated bounds.

Inference on Bounds

We begin with a description of the upper limit of our confidence interval. For each response probability π we have a set of possible upper bound solutions $\{ub_1, ub_2, ..., ub_K\}$. We know that:

$$\begin{aligned}
\pi \leq \overline{\pi} &\equiv \min\{\underline{ub}, 1\} \\
\underline{ub} &\equiv \min\{ub_1, ub_2, ..., ub_K\}.
\end{aligned}$$

A consistent estimate of the least upper bound $\underline{ub}$ can be had by plugging in consistent sample moments $\widehat{ub}_k \xrightarrow{p} ub_k$ and using $\widehat{\underline{ub}} \equiv min\left\{\widehat{ub}_1, \widehat{ub}_2, ..., \widehat{ub}_K\right\}$ as an estimate of $\underline{ub}$. This estimator is consistent by continuity of probability limits. We can then form a corresponding consistent estimator $\widehat{\overline{\pi}} \equiv \min\left\{\widehat{\underline{ub}}, 1\right\}$ of $\overline{\pi}$.

To conduct inference on π, we seek a critical value r such that:

$$P\left(\underline{ub} \leq \widehat{\underline{ub}} + r\right) = 0.95, \tag{A.56}$$

as such an r implies

$$\begin{aligned} P\left(\pi \leq \min\left\{\widehat{\underline{ub}} + r, 1\right\}\right) &\geq P\left(\overline{\pi} \leq \min\left\{\widehat{\underline{ub}} + r, 1\right\}\right) \\ &\geq P\left(\underline{ub} \leq \min\left\{\widehat{\underline{ub}} + r, 1\right\}\right) \\ &= P\left(\underline{ub} \leq \widehat{\underline{ub}} + r\right) 1\left[\widehat{\underline{ub}} + r < 1\right] + 1\left[\widehat{\underline{ub}} + r \geq 1\right] \\ &= 0.95 \times 1\left[\widehat{\underline{ub}} + r < 1\right] + 1\left[\widehat{\underline{ub}} + r \geq 1\right] \\ &\geq 0.95, \end{aligned}$$

with the first inequality binding when $\pi = \overline{\pi}$ and the second when $\underline{ub} < \overline{\pi}$.

We can rewrite $(A.56)$ as:

$$P\left(-\min\left\{\widehat{ub}_1 - \underline{ub}, \widehat{ub}_2 - \underline{ub}, ..., \widehat{ub}_K - \underline{ub}\right\} \leq r\right) = 0.95,$$

or equivalently

$$P\left(\max\left\{\underline{ub} - \widehat{ub}_1, \underline{ub} - \widehat{ub}_2, ..., \underline{ub} - \widehat{ub}_K\right\} \leq r\right) = 0.95.$$

It is well known that the limiting distribution of $\max\left\{\underline{ub} - \widehat{ub}_1, \underline{ub} - \widehat{ub}_2, ..., \underline{ub} - \widehat{ub}_K\right\}$ depends on which and how many of the upper bound constraints bind. Several approaches to this problem have been proposed which involve conducting pre-tests for which constraints are binding (e.g. Andrews and Barwick, 2012).

We take an alternative approach to inference that is simple to implement and consistent regardless of the constraints that bind. Our approach is predicated on the observation that:

$$P\left(\max\left\{ub_1 - \widehat{ub}_1, ..., ub_K - \widehat{ub}_K\right\} \leq r\right) \leq P\left(\max\left\{\underline{ub} - \widehat{ub}_1, ..., \underline{ub} - \widehat{ub}_K\right\} \leq r\right), \tag{A.57}$$

with equality holding in the case where all of the upper bound solutions are identical. We seek an r' such that:

$$P\left(\max\left\{ub_1 - \widehat{ub}_1, ..., ub_K - \widehat{ub}_K\right\} \leq r'\right) = .95. \tag{A.58}$$

From $(A.57)$,

$$P\left(\max\left\{\underline{ub} - \widehat{ub}_1, ..., \underline{ub} - \widehat{ub}_K\right\} \leq r'\right) \geq .95,$$

with equality holding when all bounds are identical.

A bootstrap estimate $r^* \xrightarrow{p} r'$ of the necessary critical value can be had by considering the bootstrap analog of condition (A.58) (see Proposition 10.7 of Kosorok, 2008). That is, by computing the 95th percentile of:

$$\max\left\{\widehat{ub}_1 - \widehat{ub}_1^*, ..., \widehat{ub}_K - \widehat{ub}_K^*\right\}$$

across bootstrap replications, where stars refer to bootstrap quantities. An upper limit U of the confidence region for π can then be formed as:

$$U = \min\left\{\widehat{\underline{ub}} + r^*, 1\right\}.$$

Note that this procedure is essentially an unstudentized version of the inference method of Chernozhukov et al. (2009) where the set of relevant upper bounds ($\mathcal{V}_0$ in their notation) is taken here to be the set of all upper bounds, thus yielding conservative inference.

We turn now to the lower limit of our confidence interval. Our greatest lower bounds are all of the form:

$$\pi \geq \underline{\pi} \equiv \max\{lb, 0\}.$$

We have the plugin lower bound estimator $\widehat{lb} \xrightarrow{p} lb$. By the same arguments as above we want to search for an r'' such that

$$P\left(lb \geq \widehat{lb} - r''\right) = 0.95.$$

Since $\widehat{lb}$ is just a scalar sample mean, we can choose $r'' = 1.65\sigma_{lb}$ where σ_{lb} is the asymptotic standard error of $\widehat{lb}$ in order to guarantee the above condition holds asymptotically. To account for the propensity score reweighting, we use a bootstrap standard error estimator $\widehat{\sigma}_{lb}$ of σ_{lb} which is consistent via the usual arguments. Thus, our conservative 95% confidence interval for π is:

$$\left[\max\left\{0, \widehat{lb} - 1.65\widehat{\sigma}_{lb}\right\}, \min\left\{\underline{\widehat{ub}} + r^*, 1\right\}\right].$$

This confidence interval covers the parameter π with asymptotic probability of at least 95%.

References

1. Andrews, Donald and Panle Jia Barwick. 2012. "Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure." *Econometrica* 80(6): 2805-2826.

2. Chernozhukov, Victor, Sokbae Lee, and Adam Rosen. 2009. "Intersection bounds: estimation and inference." working paper.

3. Giné, Evarist and Joel Zinn. 1990. "Bootstrapping General Empirical Measures." *Annals of Probability* 18(2): 851-869.

4. Kosorok, Michael. 2008. *Introduction to Empirical Processes and Semiparametric Inference* Springer: New York.

5. Murty, Katta. 1983. *Linear Programming* New York: John Wiley & Sons.

Appendix Table 1 - Summary of Policy Differences Between AFDC and Jobs First

	JF	AFDC
Eligibility	Earnings Below Poverty Line	Earnings level at which benefits are exhausted (see disregard parameters below)
Earnings disregard	All earned income disregarded up to poverty line	Months 1-4: $120 + 1/3 Months 5-12: $120 Month >12: $90 "Fill the gap" budgeting
Time Limit	21 months (6 month extensions possible)	None
Work requirements	Mandatory work first (exempt if child <1)	Education / training (exempt if child < 2)
Other	●Sanctions (moderate enforcement) ●Asset limit $3,000 ●Partial family cap (50 percent) ●Two years transitional Medicaid ●Child care assistance ●Child support: $100 disregard, full pass-through	●Sanctions (rarely enforced) ●Asset limit $1,000 ●100 hour rule and work history requirement for two-parent families ●One year transitional Medicaid ●Child support: $50 disregard, $50 maximum pass-through

Sources: Adams-Ciardullo et al. (2002) and Bitler, Gelbach, and Hoynes (2005).

Appendix Table 2: Cross Tabulation of grant-inferred AU size and kidcount

	kidcount				
	0	1	2	3	Total
Inferred AU Size					
1	0.17	0.08	0.04	0.01	0.05
2	0.53	0.84	0.19	0.06	0.42
3	0.17	0.06	0.72	0.17	0.29
4	0.11	0.01	0.05	0.53	0.17
5	0.00	0.00	0.00	0.14	0.04
6	0.00	0.00	0.00	0.01	0.00
7	0.03	0.00	0.00	0.07	0.02
8	0.00	0.00	0.00	0.00	0.00
# of monthly observations	840	11,361	8,463	8,043	28,707

Notes: Analysis conducted on Jobs First sample over quarters 1-7 post-random assignment. Kidcount variable, which gives the number of children reported in baseline survey, is tabulated conditional on non-missing. The AU size is inferred from (rounded) monthly grant amounts. Starting with AU size 5, the unique correspondence between AU size and rounded grant amount obtains only for units which do not receive housing subsidies. The size inferred during months on assistance is imputed forward to months off assistance and to months that otherwise lack an inferred size.

Appendix Table 3: Probability of Earnings / Participation States in AFDC Sample (Conditional on State=0p in Quarter Prior to Random Assignment)

Quarter post-RA:	1	2	3	4	5	6	7
Pr(State=0n)	0.022	0.062	0.086	0.093	0.114	0.136	0.136
Pr(State=1n)	0.021	0.045	0.058	0.079	0.084	0.101	0.112
Pr(State=2n)	0.006	0.021	0.024	0.033	0.048	0.044	0.074
Pr(State=0p)	0.786	0.723	0.675	0.631	0.584	0.563	0.539
Pr(State=1p)	0.160	0.160	0.145	0.160	0.157	0.150	0.143
Pr(State=2p)	0.002	0.001	0.004	0.004	0.004	0.002	0.005

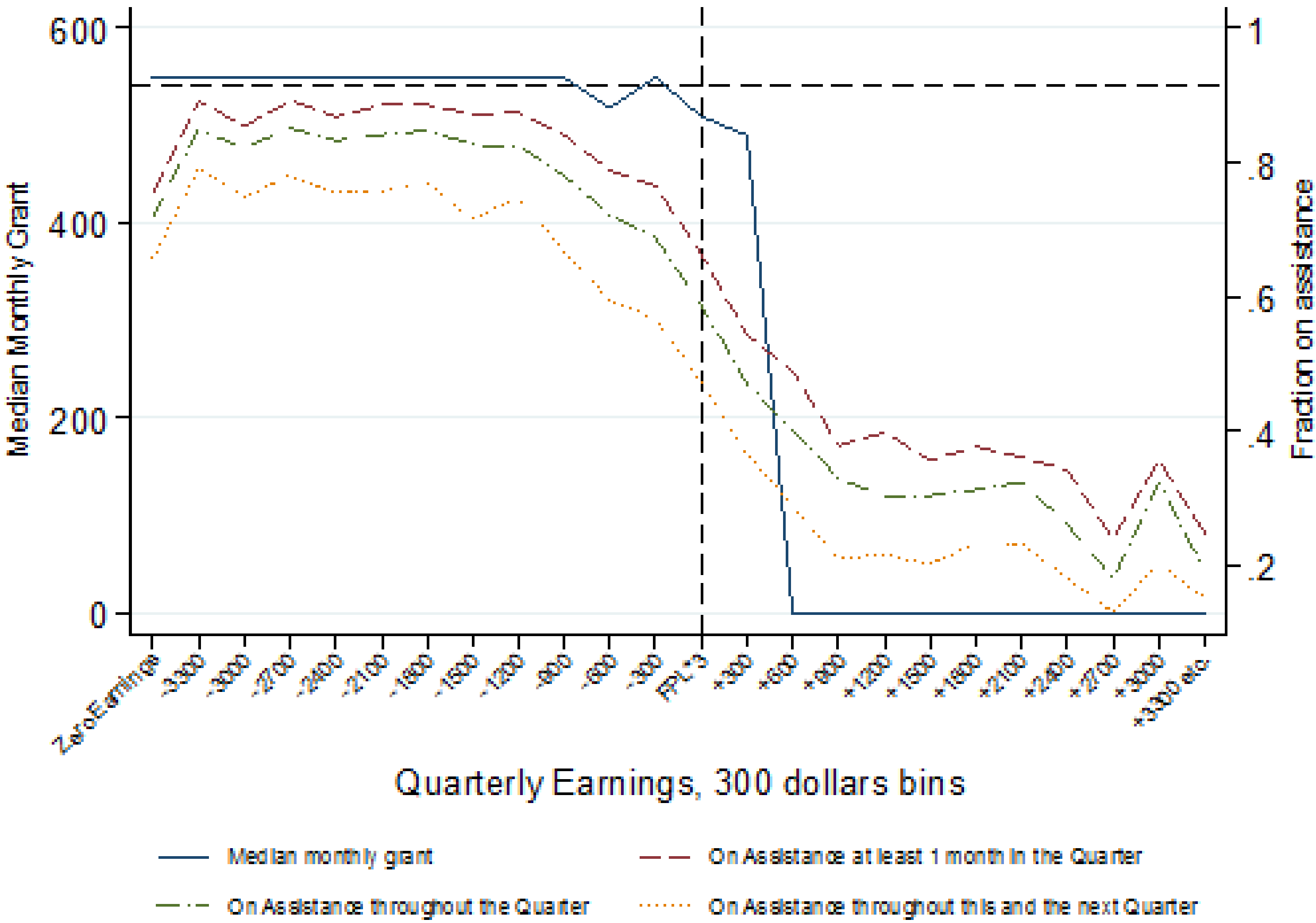
Notes: Sample consists of 902 AFDC cases that were not working in the quarter prior to random assignment and were on welfare. Sample units with kidcount missing are excluded. Numbers give the reweighted fraction of sample in specified quarter after random assignment occupying each earnings / welfare paticipation state. Number of state refers to earnings level, with 0 indicating no earnings, 1 indicating earnings below 3 times the monthly FPL, and 2 indicating earnings above 3FPL. The letter n indicates welfare nonparticipation throughout the quarter while the letter p indicates welfare participation throughout the quarter. Poverty line computed under assumption AU size is one greater than amount implied by baseline kidcount variable. Probabilities are adjusted via the propensity score reweighting algorithm described in the Appendix.

Appendix Table 4: Fraction of Months on Welfare by Experimental Status and Age of Youngest Child

Age of Youngest Child at Baseline:	16 or 17	15 or less
AFDC	0.441	0.651
	(0.038)	(0.008)
JF	0.508	0.740
	(0.039)	(0.007)
Impact	0.089	0.066
	(0.010)	(0.055)
Difference in Impacts	**-0.022**	
	(0.056)	

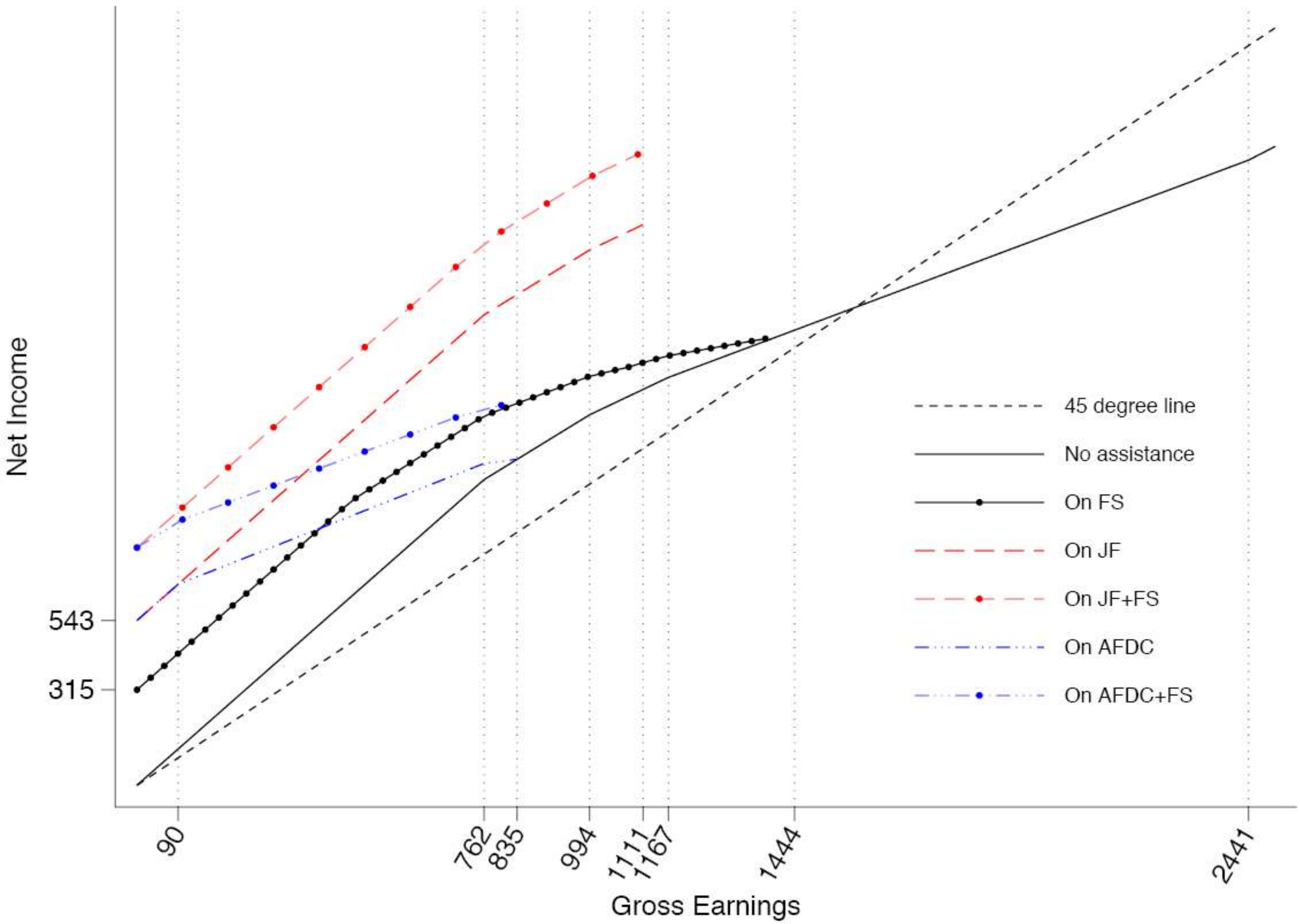
Notes: Sample consists of 87,717 person-months: 21 months of data on each of 4,177 cases with non-missing baseline information on age of youngest child. Table gives reweighted fraction of person-months that women participated in welfare by experimental status and age of youngest child at baseline. Standard errors computed using 1,000 block bootstrap replications (resampling at case level).

Appendix Figure 1: Benefits as a Function of Earnings in JF Sample



Notes: The sample includes JF cases in all months in Q1–Q7 post-RA such that imputed AU size is between 2 and 5. The horizontal line is drawn at (unrounded) maximum grant of a size 3 AU in 1997. Grant amounts and earnings are rescaled using, respectively, the maximum grant and FPL of size 3 AU in 1997. Median monthly grant computed across all cases including those not on welfare in the quarter.

Appendix Figure 2: Hypothetical Budget Sets Under AFDC and JF, Accounting for Food Stamps and Taxes



Notes: Figure depicts hypothetical budget set faced by assistance unit of size 3 under AFDC and JF policy rules. Illustration assumes household only has access to fixed $90 disregard under AFDC and faces $366 in monthly rental expenses. Net income is earnings net of federal taxes and inclusive of EITC and transfers (given participation). No assistance corresponds to earnings net of payroll taxes and federal income taxes and inclusive of EITC. *Vertical lines*: at the AFDC fixed earning disregard and break-even level ($90 and $835), at the end of the EITC phase-in and start and end of the phase-out regions ($762, $994 and $2,441), at the minimum taxable earnings ($1,167), at the FPL ($1,111), and at 1.3FPL ($1,444) which is a FS eligibility threshold under AFDC. *Horizontal ticks*: at maximum FS and welfare grants.