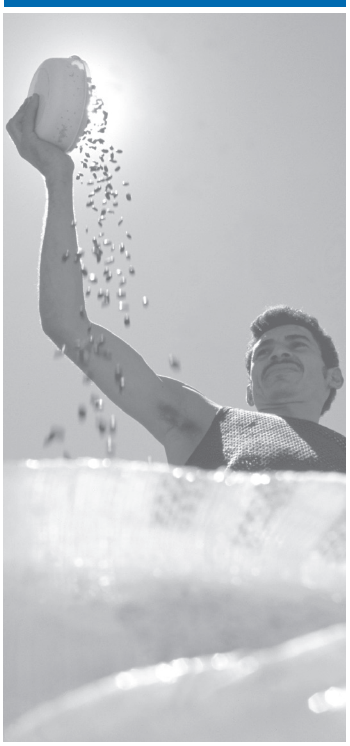
Social Protection & Labor

DISCUSSION PAPER NO. 1702

The Impacts of Cash Transfers on Women's Empowerment: Learning from Pakistan's BISP Program

Kate Ambler and Alan de Brauw



February 2017

WORLD BANK GROUP
Social Protection & Labor

The Impacts of Cash Transfers on Women's Empowerment: Learning from Pakistan's BISP Program

Kate Ambler and Alan de Brauw[1]

February 2017

Abstract

Large-scale government cash transfer programs have become an important element of social protection and poverty reduction strategies throughout the developing world. Pakistan is no exception; in 2008, Pakistan established the Benazir Income Support Program (BISP) as an unconditional cash transfer targeted at the poorest of the poor. The primary goal of the BISP program is to provide the poorest households in Pakistan with unconditional transfers in order to improve their consumption and investments in children. To attain this goal, it is believed important that the transfers are provided directly to women to ensure the funds are spent as intended. Beyond changes in consumption and investment, directing these transfers to women can also serve to empower women by increasing household resources under their control. We analyze the impacts of Pakistan's BISP program on women's decision-making power within households using data collected between 2011 and 2013 as the program was rolling out. Using fuzzy regression discontinuity methods to statistically identify impacts, the BISP transfer is found to have substantial, positive impacts on some variables measuring women's decision-making power and empowerment.

JEL classification: I38, J16, O12

Key words: Pakistan, cash transfer, women's empowerment, regression discontinuity design

---

[1] Kate Ambler is a Research Fellow and Alan de Brauw is a Senior Research Fellow in the Markets, Trade, and Institutions Division at IFPRI. Please direct correspondence to Alan de Brauw, a.debrauw@cgiar.org. We thank the support of the World Bank for primarily funding the research and the CGIAR Program on Policies, Institutions, and Markets for additional funding, and we thank Jill Bernstein and Cheng Qiu for research assistance. We further thank Benedicte de la Briere, Yanyan Liu, Lucian Pop, and Dominique van de Walle for comments that have improved the manuscript. The results do not reflect the opinions of either the World Bank or the International Food Policy Research Institute.

Contents

Figures

Tables

The Impacts of Cash Transfers on Women's Empowerment: Learning from Pakistan's BISP Program

Starting with programs in Brazil (Lindert et al., 2007) and *Progresa* in Mexico (Levy, 2006), large scale cash transfer programs have been established in many developing countries in the last two decades to attempt to improve human capital outcomes among the poor. Recent innovative program designs combined with rigorous impact evaluations have justified programs to both donors and taxpayers, as benefits have become clear (Fiszbein and Schady, 2009). Conditional cash transfer programs have demonstrated strong positive impacts on schooling in Latin America (e.g. Schultz, 2004; Attanasio et al., 2005; Schady and Araujo, 2008; de Brauw and Gilligan, 2012); several programs have also reduced poverty headcounts and poverty gaps (Fiszbein and Schady, 2009). As transfers are typically given to women, the increase in the diet quality of food consumption in Mexico (Hoddinott, Skoufias, and Washburn, 2000), Colombia (Attanasio and Mesnard, 2006), and Nicaragua (Maluccio and Flores, 2005) led Schady and Rosero (2008) to hypothesize and indirectly test whether women's bargaining power increases when households receive cash transfers.

Progresa as well as more recently designed programs have the improvement of women's empowerment as an explicit program goal (Doss, 2013). Women's empowerment can be defined as *"the expansion in the ability to make strategic life choices in a context where this ability was previously denied to them"* (Kabeer, 2001). Yet this concept is difficult to quantify, as it encompasses both decision making and the context in which those decisions are made.

In this paper, we study the impacts of Pakistan's Benazir Income Support Program (BISP) on measures related to women's empowerment. As in many other cash transfer programs, BISP transfers are distributed solely to women, with the broad goal of improving women's status within the household. We explicitly test whether outcomes beyond just decision making are affected by the receipt of a BISP transfer, examining impacts on variables related to gender norms, women's mobility, and on the ability to vote. Whereas these variables do not totally encompass women's empowerment, they go beyond the types of questions often asked in evaluation surveys of social protection programs which largely measure spheres of decision making (e.g. de Brauw et al., 2014; Bonilla et al., 2016). Moreover, due to the survey design we are able to study both men's and women's answers to a subset of these questions.

The BISP is a unique case in social protection for several reasons. First, it is an unconditional transfer, similar to the types of transfers given in sub-Saharan Africa (Bonilla et al., 2016). Whereas there is evidence that conditions matter for some outcomes (e.g. schooling in Mexico; de Brauw and Hoddinott, 2011), more recent papers that have randomized conditionality find that the conditions do not affect impacts on all outcomes (e.g. Akresh et al., 2013), and moreover, the impacts of conditional cash transfers can depend on the form of conditions (Baird et al., 2013).

Second, the BISP is a nationwide program that has expanded quickly. Since 2011, the BISP has been targeted using a proxy means test (PMT), which is used to attempt to reach 15 percent of households nationwide with regular, unconditional cash transfers. One of the challenges in evaluating a large social program such as the BISP is the ability to credibly identify impact estimates, as such programs are typically not randomized. Because BISP beneficiaries are determined through the poverty score, there is a change in the probability of receiving the BISP transfer at a pre-determined threshold. Therefore, we can use fuzzy regression discontinuity design to identify causal impacts of the program at that threshold. While the estimates are therefore theoretically unbiased, they should be interpreted as local average treatment effects, local to the threshold and not as average impacts over all beneficiaries.

To meet the objective of the paper, we proceed as follows. First, we describe the history of the BISP program and its primary objectives in more detail. Second, we describe the literature on impacts of cash transfers on women's empowerment. Third, we introduce the data and fourth, we describe the estimation strategy and present evidence that the methodology used should lead to low bias impact estimates for the local average treatment effect. We present results in the fifth section, and the final section concludes.

1. Program Background

Cash transfers are somewhat rare in South Asia relative to East Asia and Latin America. Whereas 70 and 60 percent of the poorest quintile in Latin America as East Asia receive cash transfers, respectively, only 27 percent of the poorest quintile receive cash transfers in South Asia (FAO, 2015). Pakistan is an exception; it first developed a Social Protection Strategy in 2007, and announced the BISP as its main social safety net program in 2008. The BISP initially aimed to help the "poorest of the poor" through unconditional cash transfers. It has three main policy goals. First, it aims to eradicate extreme and chronic poverty. Second, it aims to empower women, and third, it aims to achieve universal primary education. The first goal is meant to be achieved through regular cash transfers, and the second goal by specifically by giving transfers to women. The government is attempting to meet the third goal through other means.

As the Pakistani (and the world) economy was characterized by high food price inflation when the BISP began, there was some urgency to increase the declining purchasing power among the poorest members of society. Consequentially, initial program targeting took place through Parliamentarians, who were each asked to identify 8,000 beneficiary households on a prescribed form, on which names, national ID card, and household income information was collected. Under this system of targeting, the initial rollout led to disbursement to over 2 million eligible families.

As a result of concerns over the effectiveness and transparency of Parliamentarian targeting, payments to beneficiaries were stopped and a new national targeting mechanism based on a Proxy Means Test (PMT) was developed. Weights for the PMT were developed using the 2007/8 Pakistan Living Standards Measurement Survey, and the PMT uses 23 variables to compile a poverty score. To target the BISP, a Poverty Scorecard survey was initiated in 2010/11, collecting information on those 23 variables. Upon completion of data collection, a PMT score was generated for every household. A PMT threshold (cut-off score) of 16.17 was established to attempt to reach the targeted population, though a few exceptions were also allowed. Households could appeal and receive transfers if their score was between 16.17 and 21.17 and there was either 1) at least one disabled member of the household; 2) the presence of at least one senior citizens (65 years of age or older) and fewer than three household members; or 3) The household includes 4 or more children under 12 years of age. All beneficiaries selected under the parliamentarian targeting system were also recertified; those who had PMT scores above 16.17 and did not qualify for one of the four exceptions were excluded from further benefits.

To protect the poorest families, the BISP aims to deliver cash transfers to each ever-married female in eligible households as identified by the PMT, subject to the females possessing a valid Computerized National Identity Card (CNIC). Women have to register at their local BISP office for the transfers, where their PMT score or poverty score is verified for eligibility, and they become registered in the system for transfers. Some households include as members more than one eligible woman; for the purposes of analysis, we count any household with at least one eligible woman as a beneficiary household. During the period under analysis for this paper (2011-2013), the benefit level is of PKR 1,000 per month, to be paid in quarterly installments of PKR 3,000. Originally, the vast majority of beneficiaries were paid through the Pakistan Post. However, following reports of misappropriation of funds, other options were considered, and BISP began to replace payments through the Pakistan Post with ATM cards (Benazir Debit Card, or BDC), which allow beneficiaries to collect money directly from ATMs or Point-of-Sale (POS) machines. The transition from Pakistan Post payments to ATM card payments began in 2012, and the majority (around 75 percent) of beneficiaries that we identify in the administrative data were receiving payments through ATM cards; some beneficiaries continue to receive payments through Pakistan Post because they are quite geographically isolated, and ATM cards are issued through the local BISP office.

2. Do Cash Transfers Affect Women's Empowerment?

The notion that making women the direct recipients of cash transfers will improve their influence in households' resource allocation decisions and will empower them in general is at the heart of BISP. The notion that women should receive cash transfers derives from the original plan for *Progresa/Oportunidades*, which was designed under the working hypothesis that by providing women with more external income, they would have a larger share of overall household income, and therefore their agency within the household would increase (Schultz, 2004). From a theoretical perspective, the transfers represent a change in the share of income earned by each adult within the household, which lasts so long as the household receives transfers; the changes are not expected to last beyond the receipt of the transfers.

Yoong, Rabinovich, and Diepenveen (2012) systematically review the literature on the relative efficacy of giving transfers to women; they suggest that although the bargaining power of an individual within the household depends upon their income share, social norms or a lack of formal legal rights can reduce the impact of making social protection payments to women on their bargaining power. Making a similar argument, Handa et al. (2009) argue that cash transfers could crowd out any intrahousehold transfers from men to women, rather than working to increase women's bargaining power within households. This argument is consistent with evidence from the *Progresa* impact evaluation, which found that women retain agency over the transfers, but little else within households (Attanasio and Lechene, 2002).[2]

Consequently, the literature on cash transfer programs shows mixed evidence on impacts on women's empowerment or decision making, if any (de la O Campos, 2015). Whereas Attanasio and Lechene (2002) and Handa et al. (2009) both show little evidence of quantitative impacts of *Progresa* on women's empowerment, Adato et al. (2000) in qualitative work on the same program find evidence of increased self-confidence and self-esteem. Moreover, evaluations of cash transfer programs in other countries show positive impacts. For example, de Brauw et al. (2014) find positive impacts of *Bolsa Familia* on some spheres of women's decision making power in Brazil. Ambler (2016) finds that the receipt of the pension in South Africa increases the likelihood of women becoming the primary decision maker in the household. The size of the transfer may be particularly important in the latter

[2] Moreover, if transfers end, unless women have made investments that increase their bargaining power their share of income would drop, reducing their decision making power or other measures of women's empowerment.

case; as described by both Duflo (2003) and Ambler (2016), the pension represented a very substantial increase in income for households in South Africa. In Pakistan, the BISP transfers only represent 6 percent of household consumption expenditures, implying that the situation might differ there.

A major challenge in demonstrating that cash transfers have impacts on women's empowerment or decision making power is that these concepts are difficult to quantify. Fortunately, the BISP evaluation data has several sets of indicators that relate to women's empowerment, and a set of questions about gender norms were asked both among men and women.[3] As measures of empowerment, we use the full set of these questions as measures of gender norms, the set of questions about female mobility reported by women, and a separate question that was asked about female voting behavior. The indicators we study, and their source in the questionnaire, can be found in Table 1.

2.1. Women's Status in Pakistan

Overall levels of women's autonomy are generally low in South Asia (Jejeebhoy and Sathar, 2001). In Pakistan, however, South Asian norms combine with Islamic norms to give women even less control over their lives; Jejeebhoy and Sathar (2001) show that women in Punjab, Pakistan, have less mobility than women in Uttar Pradesh or Tamil Nadu in India. In describing the gendered division of labor in Pakistan, Akram-Lodhi (1996) finds that women perceive virtually no mobility rights outside of the village. Although Pakistan has fairly poor statistics related to reproductive health (UNDP, 2013), Mumtaz and Salway (2005) caution that mobility may not be linearly related to women's reproductive health outcomes. Their ethnographic study shows that if poor women are mobile it may reflect negatively on their status, whereas richer women have more freedom to move around. Whereas Mumtaz and Salway (2005) cannot establish a correlation between mobility and reproductive health outcomes, in a nationally representative data set Hou and Ma (2013) find that women's decision-making power—a broader concept—is positively correlated with the use of

[3] In Pakistan, women and men would not sit together to respond to an interview, so they were interviewed separately in each household in the data set about topics they would each be best placed to answer.

maternal health services in Pakistan, whereas men's decision-making power is negatively correlated with maternal health service use. These papers suggest that women's empowerment should not be studied in isolation, but alongside thoughts about how men perceive women's empowerment; and that in a relative sense women in particular lack status in Pakistan.

3. Data

For the impact evaluation, we use the combination of two data sources. The primary data source are panel household data collected specifically for the evaluation of the BISP transfer. Those data are combined with administrative records of both eligibility for payments and the receipt of BISP payments, so that we are able to use administrative data on payments rather than self-reports.

3.1. Household Survey Data

We use the baseline and follow-up surveys conducted by Oxford Policy Management (OPM) between April and July, 2011, and between April and July, 2013, respectively. The baseline survey included households in the four target provinces (Punjab, Sindh, Khyber Pakhtunkhwa or KPK, and Baluchistan), and the sample was chosen to be representative of households close to the pre-determined poverty threshold of 16.17 in the four provinces. The evaluation sample was created with the intention of exploiting the eligibility criterion (poverty score of 16.17 or less) to create a sample of "control" (households with poverty score just above 16.17) and "treatment" (households with poverty score below 16.17). Thus the evaluation sample has beneficiaries or potential beneficiaries as well as comparable non-beneficiaries.

Ideally, the baseline survey for the impact evaluation would have taken place after the BISP poverty census had taken place. However, the poverty census did not take place at the same time in all of the evaluation provinces, and so payments took place in some districts before the census was complete in others. As a result, in order to conduct the baseline survey before payments were made to any beneficiaries, the survey was conducted while the BISP poverty census was still taking place. Consequently, a sampling scheme was required to ensure that the baseline would include as many beneficiaries and valid controls as possible.

The baseline sample was constructed as follows.[4] The sample frame was constructed in four phases. First, primary sampling units (PSUs) or clusters within the 2007-8 Pakistan Social and Living Standards Measurement (PSLM) Survey were stratified at the provincial and rural/urban level. Evaluation PSUs were then sampled directly from these strata using simple random sampling. In phase 2, a household listing exercise was conducted in all evaluation PSUs to form the basis of the sampling frame of households within evaluation PSUs. The household listing exercise was conducted in all sampled clusters, and implemented by the Federal Bureau of Statistics (FBS) on behalf of OPM, and included the variables necessary to reconstruct the proxy means test (PMT) used for targeting the BISP transfer. In phase 3, from the listing exercise a predetermined number of households were randomly selected using simple random sampling, on which the PMT was applied. The PMT from the listing exercise was a mimic of the BISP Poverty Scorecard described above, and an average of 100 household PMTs were applied per PSU. Finally, in the fourth phase, households to whom the PMT was applied were split into two groups. Group A households are defined as those households with PMT scores equal and below the cut-off score of 16.17. Group B households are defined as those households with a PMT score greater than the cut-off score of 16.17 and within a predetermined range up to a score of 21.17, making them hypothetically valid counterfactuals for the RDD methodology. Respective samples of Group A and Group B households were chosen from within each PSU using simple random sampling.

The baseline survey was conducted in 488 PSUs, with average sample sizes per PSU of 19 households per rural cluster and 15 households per urban cluster. The total baseline sample size was 8,675 households. Sample size by group, province, and survey round is reported in Table 2. Some households were already receiving BISP transfers at the time of the baseline, and households were asked to self-report about their experiences with receiving the BISP. The follow-up survey in 2013 attempted to reach the baseline households again, as

[4] This section relies heavily on data collection reports by Oxford Policy Management (2014).

well as households that split off from the original households. Consequently, the follow-up survey attempted to reach 9,119 households, including split households. In total, 8,221 households were reached in the 2013 follow-up, implying an attrition rate of 9.8 percent.

3.2. Constructing the Panel with Administrative Poverty Scores

We use administrative data from the BISP Management Information System (MIS) to locate the PMT of each household and its beneficiary status. As described above, the PMT is a number that tells us whether or not the household is eligible, while the beneficiary status reflects whether or not a specific woman received payments. By the 2013 survey, not all eligible women had registered to receive payments. A drawback in the way that the baseline data was collected is that individuals and households had to be matched to the BISP MIS, largely using CNIC numbers which had to be collected, written onto survey forms, and then entered into databases. If no women in the household had a CNIC card (either because they actually do not have one or because it was not recorded properly on the survey form), it is not possible to match the household to the baseline (or follow-up) survey data. From a matching perspective, an additional complication is that some households had multiple women with CNIC cards, who may have also been given different poverty scores when the BISP census took place.[5] We treat such households by using the lowest possible poverty score, as we are interested in overall household eligibility for transfers.

Given the chances for errors in entering CNIC numbers and the general ambiguity in the way multiple CNIC numbers should be treated, there were several chances for errors to take place. Moreover, women with CNIC cards may not have attempted to register for BISP payments, which would preclude their inclusion in the BISP MIS. Using all possible CNIC numbers, of the 8,221 panel households we matched 5,650 households to the BISP MIS, so 2,583 households drop from the sample (Figure 1). Of these 5,650 households, 2,924 are eligible to be beneficiaries (have a PMT below 16.17) and 2,726 are not eligible (have a PMT

[5] Less than 25 percent of matched households have more than one resident married woman with a CNIC card.

above 16.17).[6] Of those households that we could not match, in 606 households no woman reported owning a CNIC card, and in 53 households it is reported as "don't know" or missing whether any married women have a CNIC card. Hence, there are 1,912 households that could not be matched to the BISP in which married women reported having CNIC cards.

Consequently, there are effectively two levels of attrition from the data set. First, there are households interviewed in the baseline that could not be found in the second round; these households again could have had CNIC cards or not. Additionally, though we have information on the PMT for 5,650 households, we do not have information on whether or not of 55 of those households received transfers, so they further drop from the sample. For the purposes of the paper, we provide some comparisons between statistics generated using the full sample and the estimation sample.[7]

Because the CNIC card is required for BISP receipt and not all women had one at baseline, one may be concerned that those who were eligible for the transfer were more likely to obtain a CNIC card and therefore more likely to show up in our sample, potentially biasing our estimates. Therefore, for our estimates we drop all households in which women report obtaining a CNIC card between rounds, resulting in an estimation sample of 4,981 households.

Finally, we note that by combining these data sources, we are largely measuring the household level impact of receiving BISP transfers, as any recipient within a household might affect outcomes among all. One might be concerned that some households receive additional benefits (e.g. those with multiple transfer eligibility) and thus we do not properly account for these multiple transfers with an indicator variable at the household level measuring BISP transfer receipt. That said, the BISP transfer is constant at 1000 rupees per beneficiary, and does not account at all for household size. So the per capita benefits in a

---

[6] We also considered dropping beneficiaries at baseline from the sample. We do not, because the estimation strategy we use—regression discontinuity design—is theoretically unbiased for the difference at endline, so there is no need to remove beneficiaries at baseline.

[7] The increase in the number of households in which married women had obtained CNICs can be considered a result of the BISP program, though we cannot make a causal estimate of the impact. The CNIC effectively ensures the citizenship of these women, allowing them access to any social services as well as voting eligibility.

single transfer household are much smaller for an 8 person household than for a four person household (for example). One should interpret our results as averaged across single and multiple beneficiary households; potentially taking into account the number of transfers (or the per capita amount of transfers) is a possible extension of this work and a question of interest in the general literature on cash transfers.

4. Methodology

Although BISP receipt is not random, its implementation was designed to allow for causal identification of its impacts. The PMT was used as a cutoff for the BISP specifically to be able to use regression discontinuity design (RDD) as an evaluation strategy. Under RDD, marginally ineligible households serve as a comparison group for marginally eligible households (above and below the PMT cut-off of 16.17). However, the discontinuity is not sharp, meaning that some households below the threshold do not receive transfers, and some households just above the threshold do receive transfers. Therefore we use fuzzy RDD design to measure impacts of the BISP transfer, which we briefly explain below.

4.1. Regression Discontinuity Design

RDD was first introduced by Thistlewaite and Campbell (1960) and an excellent review of RDD in economics can be found in Lee and Lemieux (2010). Following Lee and Lemieux's notation, consider a variable X that is used to determine program participation. Observations with a value of $X \geq c$, where *c* is the threshold, are eligible for the program, while individuals with a $X < c$ are not eligible. We can therefore define a dummy variable for treatment such that $D = 1$ if $X \geq c$ and $D = 0$ if $X < c$.

Now, consider an outcome Y that that the program above may affect. Prior to program participation, there is no reason to expect that the relationship between Y and X would be discontinuous at the value $X = c$. So long as *Y* is continuous over the range of *X* near $X = c$, one can consider the average treatment effect at *c* as:

$$\tau = \lim_{X \downarrow c} E[Y_i | X_i = x] - \lim_{X \uparrow c} E[Y_i | X_i = x] \qquad (1)$$

where *i* indexes individuals.

Although the average treatment effect estimated using RDD is limited because it is local, as Lee and Lemieux (2010) point out this design is actually akin to a localized random experiment. Given that the treatment assignment *D* is defined solely based on a specific value of *X*, it is a somewhat trivial assumption that unobservables are not related to *D*, so the estimate of τ is unbiased by design. Further, one can test whether other variables that should not be affected by the treatment are indeed continuous in *X* at the threshold, as one would expect. And finally, as in a randomized experiment, it is not necessary to also control for any other baseline covariates *W* in a linear regression framework. However, one can add covariates to the regression to attempt to explain some of the variation in the dependent variable. Theoretically, doing so can reduce standard errors on estimated coefficients since the unexplained variance decreases.

The RDD explained to this point is a sharp RDD, meaning that the probability that someone receives the program is zero if the variable *X* is below *c* and one if the individual has a value of *X* above *c*. In the case of Pakistan BISP, the "forcing" variable *X* is the PMT and the cutoff *c* is 16.17, but some households below the threshold receive transfers and some above the threshold receive them. In terms of RDD, we can call this occurrence imperfect compliance to the forcing rule. Trochim (1984) termed this occurrence as fuzzy RDD design, and showed that all that is required to demonstrate impacts at the threshold is that

$$\lim_{X \downarrow c} E[D = 1|X = x] \neq \lim_{X \uparrow c} E[D = 1|X = x] \tag{2}$$

Since the probability of treatment at the threshold *c* no longer increases from zero to one, the jump in the relationship between *Y* and *X* cannot be interpreted as a local average treatment effect. However, the treatment effect can be estimated as in an instrumental variables framework. Theoretically, to obtain the local average treatment effect one can divide the jump in the relationship between *Y* and *X* at *c* by the jump in probability of treatment at *c*:

$$\tau_F = \frac{\lim_{x \downarrow c} E(Y|X = x) - \lim_{x \uparrow c} E(Y|X = x)}{\lim_{x \downarrow c} E(D|X = x) - \lim_{x \uparrow c} E(D|X = x)} \tag{3}$$

where, again following Lee and Lemieux (2010), the F subscript denotes "fuzzy." Imbens and Angrist (1994) demonstrate that this ratio can be interpreted as a causal effect if we assume monotonicity and excludability. Monotonicity implies that when *X* crosses the threshold value *c*, it does not also cause some observations to take up the treatment and others to reject it. Excludability implies that crossing the threshold cannot affect the outcome *Y* except through the impact on the receipt of treatment.

To estimate equation (3), we assume that we can write the probability of treatment as:

$$\Pr(D_i = 1|X_i = x) = \gamma + \delta T_i + g(x - c) \quad (4)$$

where $T_i = 1(X_i \geq c)$ indicates whether the forcing variable exceeds the threshold, and *g(.)* is a function of the distance from the threshold; δ represents the increase in probability of treatment at the threshold. We can theoretically write that$\mathrm{D}_i = \Pr(D_i = 1|X_i = x) + \omega_i$, where ω_i is a disturbance term independent of X_i, and then we can characterize the fuzzy RDD with a two equation system:

$$Y_i = \alpha + \tau D_i + f(X_i - c) + \varepsilon_i \quad (5)$$

$$D_i = \gamma + \delta T_i + g(X_i - c) + \omega_i \quad (6)$$

Estimation of equations (5) and (6) can be performed with local linear regression and two stage least squares (2SLS). An important choice in RDD analysis is selecting the range of values of the running variable which are used to conduct estimation, termed the bandwidth. While the estimates are unbiased in the limit at the threshold, one must use data to actually estimate the treatment effect τ_F using data, which implies that as data farther away from the threshold is used in estimation, estimates for τ_F become more susceptible to bias, as observations on either side of the threshold become less and less comparable as data are added. On the other hand, the inclusion of additional data on either side of the threshold allows for more precise estimates. To choose a bandwidth for estimation, then, one must balance the bias of including more observations against the variance of treatment effect estimates. A second complication in fuzzy RDD design is whether to focus on the bias-variance tradeoff in equation (5) or equation (6); Imbens and Lemieux (2008) suggest focusing on the outcome equation (5) for selecting the bandwidth and using the same bandwidth for the treatment equation (6). We test the sensitivity of results to alternative bandwidths.

In estimation, we use procedures developed by Calonico, Cattaneo, and Titunik (2014a, 2014b) to determine the optimal bandwidth. The estimator uses a local linear regression on either side of the threshold and includes a data driven bias correction as well as bias corrected confidence interval estimation. The estimator uses a triangular kernel for data included in regressions, as also suggested by Lee and Lemieux (2010).[8]

We provide two sets of estimates for each outcome. First, we provide estimates using equations (5) and (6) as written in a local linear regression framework. However, these estimates ignore available information on both baseline covariates and the baseline value of the outcome variable. While there is no theoretical need to include either for unbiased estimates, as discussed above including such variables can reduce variance of estimates and improve precision; furthermore, they take advantage of the panel nature of the data. Consequently, we also estimate versions of equations (5) and (6) that include the baseline outcome and a set of baseline control variables; for these estimates, we follow the procedure defined by Calonico et al. (2016). Other than the baseline value of the outcome, baseline controls include the logarithm of per capita consumption, household size, the years of schooling of the household head, and whether or not the head was female.

4.2. Demonstrating the Discontinuity

Given the theoretical discussion, within the estimation sample our first goal is to show that there is a discontinuous jump in the probability of transfer receipt at the PMT threshold of 16.17. If no such obvious jump exists, then the poverty score is generally not a valid forcing variable. First, we visually examine the discontinuity among the sample of households in the panel. We visually examine the discontinuity in Figure 2 by plotting the probability of transfer

[8] We have tested other kernels (e.g. rectangular, Epanachikov) and results shown in the paper are consistent under a variety of kernels.

receipt on the *y*-axis and the PMT on the *x*-axis, with the probability of transfer receipt measured administratively. The graph shows average BISP beneficiary status within bins of 0.2 points of the PMT score. The data is fit with a fourth order polynomial estimated separately on either side of the eligibility cut off.

The graphs illustratesa smooth, nearly flat relationship between the probability of receiving payment and the poverty score on the left hand side of the graph. At the threshold, the probability of receiving transfers at the threshold among eligible households is above 80 percent. We observe a fairly substantial percentage of households receiving transfers just above the threshold (around 25 percent), but the probability quickly drops off as the poverty score increases. This finding is consistent with the program rules, which award transfers to certain classes of households with poverty scores above 16.17, but by a poverty score of 21.17 no household should receive transfers. The probabilities of receiving transfers generated from the administrative data are largely consistent with these rules. Further, this break should be large enough to identify the impacts of the transfers.

Second, we estimate a version of equation (5), which specifies the relationship between the probability of receiving payments and the poverty score as a function. We estimate the equation with a linear functional form (Table 3), allowing for different relationships on either side of the threshold with interaction terms. Using all of the data on either side of the threshold, and without weighting observations, we estimate that the discontinuity leads to a jump of 54.8 to 55.1 percentage points (columns 1 and 3), depending upon whether we use the full sample of beneficiaries (column 1) or those who had CNIC cards at baseline (column 2).[9] Given the strong indication of a discontinuity, we proceed to use local linear regression to estimate impacts of the BISP transfer.

4.3. Manipulation Testing

Prior to examining the results it is important to further test the validity of the RDD estimation strategy. As described by McCrary (2008), an important concern is that subjects might be able

[9] Note that when we use this equation as something akin to a first stage with an optimal bandwidth calculation as well as triangular weights, the threshold estimate depends upon both.

to manipulate the forcing variable, thereby invalidating the assumption that those immediately to the left of the cutoff are directly comparable to those immediately to the right of the cutoff. First, it should be noted that the nature of the PMT makes manipulation difficult if not impossible. The PMT is constructed from 23 different variables and because households did not know either how the score was constructed or the cutoff used for eligibility, it is not clear how they could have directly manipulated their score to ensure eligibility. Despite this point, we can still use the data to directly test for such a manipulation. If individuals were indeed able to manipulate their scores we would expect to observe a discontinuity would exist in the density of the forcing variable at the threshold. We test this possibility in this section.

Initially, we want to visually inspect the data to determine whether or not the poverty score appears to have been manipulated. We therefore initially show a histogram of administrative poverty scores, in 50 bins, for households with poverty scores between 0 and 32.34 (or double the threshold; Figure 3). If there was manipulation of the dependent variable, we would expect to observe a bunching of the data just below the threshold, or more density in the histogram just below the threshold. We might expect then a concurrent drop in the probability just above the threshold. While the histogram does not appear to be smooth, neither does it suggest such a discontinuity; the lack of smoothness may be a result of the somewhat non-standard sampling procedure that was followed to develop the sample.

Cattaneo, Jansson and Ma (2015) develop an alternative test that is based on a more flexible, fully data driven approach than the more traditionally estimated McCrary (2008) density test, which requires the data to be initially placed in "bins," hence adding parameters to the estimation problem. We estimate the Cattaneo et al. manipulation test for both the overall sample (e.g. all 5,650 matched households) as well as the estimation subsample (households in which women had a CNIC card at baseline). We prefer what Cattaneo et al. deem the undersmoothed version of the test statistic; their preferred version includes a bias correction, and the bias correction involves a computation of the bias that would be affected by the sampling procedure.[10] Nonetheless, we report both versions of the test statistic. We find that we fail to reject the null hypothesis of no discontinuity in three of the four cases (Table 4); in the fourth, when examining the overall sample with the bias-corrected statistic, it rejects at the 10 percent level. We only use that sample in estimating consumption results, and the undersmoothed version of the test does not reject, so manipulation of the running variable does not appear to be a concern in most of our estimates.

4.4 Continuity at the Threshold: Covariates

A second common test for the validity of RDD is to test the density of baseline variables that we might consider "control" variables, or variables that we would not expect to be affected by the program. The density of such variables should appear continuous through the threshold, else we might be concerned that impact estimates will be biased due to selection on such characteristics. One could use a direct statistical test here; one way to do so is to estimate an RDD regression as noted above using baseline characteristics as false "outcomes." However, one can also simply visually inspect the data initially to look for breaks, and if a break appears to occur one can then conduct more rigorous statistical tests.

We illustrate continuity at the threshold for four sets of variables measured at baseline. First, we examine variables related to household demographics (Figure 4); specifically, we consider household size, whether or not the household has a female head, whether or not the head is married, and the years of schooling of the household head. In all four cases, the distributions on either side of the threshold appear to match nicely, and have expected patterns; household size declines with the poverty score, while the head's education level increases with the poverty score.

[10] In work not shown here, we demonstrate with a Monte Carlo experiment that the optimal bandwidth as defined by Cattaneo et al. (2014a) is perhaps not surprisingly dependent upon the sampling procedure, and that the optimal bandwidth provided by the accompanying software is actually only truly optimal for a simple random sample. Since the bias correction provided by their software depends upon this bias calculation, for this particular sampling procedure it is not necessarily right, and as such we prefer the undersmoothed rather than the bias corrected statistic.

Second, we consider variables related to consumer durables and housing owned at baseline (Figure 5). We examine whether or not households owned a TV, a cooking stove, their residence, and a washing machine. All four variables show the expected pattern; that is, they become more likely as the poverty score increases. Important for the validity of impact of impact estimates, we observe no difference in the pattern of data around the threshold for any of the variables.

Third, we examine some housing characteristics (Figure 6). Specifically, we examine whether households have a mud floor, the number of rooms in the house, whether or not the household has a toilet, and whether or not the household has a thatched roof. We again find sensible patterns in the data, in terms of a negative correlation between a mud floor, a toilet, and a thatched roof versus the poverty score, and no evidence of a structural break at the threshold for any of the four variables.

As a final check, one might be concerned that households with more married women in them might have been targeted for transfers. If there was a structural break in the number of married women per household, then the probability of receiving multiple transfers in the household would also change and confound the average impact estimate (averaged over the number of transfers actually received). We plot the number of married women against the normed poverty score, and find no evidence of a structural break (Figure 7), suggesting this bias is not present in impact estimates. Consequently, we conclude that continuity in the distribution of a wide variety of baseline characteristics do not differ substantially at the threshold, and there is no evidence of this type of bias. Combining this evidence, we feel comfortable that baseline characteristics are continuous at the poverty score threshold, which implies that we do not need to be concerned about this type of potential bias in program impact estimates.

5. Impacts on Women's Empowerment

As the BISP cash transfer is specifically targeted to women, and women's labor force participation is relatively low, a guaranteed income stream specifically for women may have important impacts on women's empowerment. Therefore, it is important to test whether the BISP transfer has impacts on women's empowerment. We examine potential impacts in several ways, and we group variables that relate to empowerment into two broad categories. First, we look at a set of outcomes related to opinions about gender norms reported by both men and women, and then we examine a set of outcomes related primarily to female mobility, but also to voting.

Our initial list of variables relates to gender norms that may have changed over time as a consequence of receiving the BISP transfer. The variables are all as indicated in Table 1 (disagree that only male should make important decisions, agree male should help with household chores, agree that female should work, agree that wife should express opinion, disagree that wife should tolerate being beaten, and disagree that it's better to send son to school than daughter). These variables were constructed from ratings scales stating "strongly agree", "agree", "disagree", or "strongly disagree;" so each answer was collapsed first to a binary zero/one variable. In each case, the "positive" outcome is always coded as one.

Table 5 summarizes male and female responses in the estimation sample for all six variables. These are averages across all households who are eligible for the transfer, hence any changes between surveys reflect changes averaged across eligible and ineligible households. We find that in most cases, men actually suggest more positive answers than women do. For example, whereas only 25 percent of men in each survey disagree with the statement that a wife should tolerate being beaten by her husband, 18 percent of women disagree at baseline (and only 13 percent disagree in 2013). The only exception is the statement about women being able to work outside the home; women are 20 percentage points more likely to agree with that statement.

To study whether the BISP transfer has an impact on these norms, we present both estimates using the 2013 outcome as the dependent variable alone, while also estimating a model using a set of control variables, including the baseline outcome. Other control variables in those regressions include the logarithm of monthly consumption per capita, the household size, the head's education level, and whether or not the head was female, all measured at baseline. In each case, the PMT score is used as an instrument for beneficiary status as reported in the administrative data, and the analysis is on the individual level.

As discussed above, to estimate impacts we must choose a bandwidth that trades off bias and precision. For both specifications, we show results using three bandwidths: a narrow bandwidth of 3, a larger bandwidth of 5, and the optimal bandwidth resulting from the calculation in Calonico et al. (2014b). We do not solely rely on the optimal bandwidth computation because it implicitly assumes a simple random sample, and the sample here clearly attempted to measure more households near the threshold than away from it. Specifically, the program tried to sample within the bandwidth of 5 from the threshold, and we choose 3 as well as the calculated optimal bandwidth often falls between 3 and 5. For all variables, we run separate regressions for women and men.[11]

We first discuss the impacts on the female reports (Table 6). Across specifications and outcomes, the estimated impacts are generally positive; in the cross-sectional regressions (Panel A), all estimated coefficients are positive. However, a lack of statistical power means that the results fall short of statistical significance in most cases. The strongest results are in column 5, for the outcome of women disagreeing that the wife should tolerate being beaten, where there is a strong and significant positive impact for five of the six specifications, and they are significant at the 5 percent level for the optimal bandwidth, whether or not we control for covariates. Figure 8 shows the column 5 results graphically, using the cross-sectional outcome. The scatterplot is the average of the outcome variable by bins of size 0.5 points, with a quadratic regression line and 90 percent confidence intervals fit on either side of the cutoff. It also is suggestive that women who receive the transfer are more likely to disagree with the statement.

When we estimate impacts of the BISP on the same set of questions about gender norms among men, results show a similar pattern (Table 7). Though almost all of the estimated coefficients are positive, only a few are consistently significantly different from

[11] We also ran regressions using agreement between men and women on the answer as a dependent variable; however, we did not find any significant impacts so we do not show them here.

zero. The strongest results for men are in column 2 (agreement that males should help with household chores) and column 4 (agreement that wives should be able to express opinions). The latter is only significant at the five percent level when we control for covariates. We again show cross-sectional results for the agreement on helping with household chores are shown graphically in Figure 9; the graph is suggestive of a discontinuity at the threshold, or an impact on male answers to that question. Interestingly, the strongest result found for women is one in which overall agreement is quite low among women (the sample mean is 13 percent) while for men the two stronger results occur for outcomes on which agreement is quite high among men (the sample mean is about 74 percent for both variables).

We next examine outcomes related to female mobility. These variables are reports by women about where they are able to go alone: local market, local health facility, friends' home, and nearby shrine or mosque. We add one variable to this analysis; specifically, whether or not the woman reports voting either sometimes or all the time (Table 8). As with the estimates on gender norms, all estimated impacts for mobility measures across specifications are positive in the cross-section, but again they are not consistently statistically significant. The results in column 3 for women being permitted to go to friends' homes alone are significant, showing a strong increase of the BISP transfer in women being allowed to visit their friends. There is also compelling evidence that women are more likely to vote when they receive the transfer, as the impact is significant in three of six specifications, and coefficients are larger and statistically significant when we control for the baseline value in Panel B. The optimal bandwidth suggests a 15 percentage point increase in the proportion of women who are allowed to vote, from 73 percent in the control group. The results for both of these outcomes are shown graphically in Figures 10 and 11, and again are quite suggestive of positive impacts of the BISP transfer.

Given the positive coefficients that are found for all of the variables in the cross-section, but only a handful were statistically different from zero, we next create an index of these variables. First, we add up a "score" consisting of the six gender norms as reported by women; then we sum up the measures of women's mobility, both with and without the voting measure, and finally we consider the sum of the gender norms as reported by men. [12]

We estimate these four regressions in Table 9. The results in Panel A with fixed bandwidth, regardless of the bandwidth, suggest positive and significant impacts on all four indices (at the 10 percent level or better). However, all of the indices among women lose their statistical significance in Panel A when we use the optimal bandwidth, which tends to be higher than 5. However, from the men's answers about gender norms, all of the coefficients are significant at the 5 percent level or better, suggesting an increased value of 0.6 relative to a mean of 3.1, suggesting meaningful improvements in women's empowerment from the perspective of men.

These results only differ slightly when we add the baseline outcome and control variables as covariates (Panel B). The overall index, including gender norms and the mobility variable, does exhibit significant at the 10 percent level when we use the optimal bandwidth (column 3). Among men, the results become even stronger, suggesting a 0.8 point increase in the men's index. These results suggest that although the outcome by outcome measures of empowerment may appear to be only suggestive, when combining a variety of measures the evidence is at least strongly supportive that the BISP results in an increase in men's perception of women's empowerment.

In sum, there is reasonably strong evidence that the BISP transfer has had some impacts on women's empowerment. Along the lines of some specific gender norms, men's and women's attitudes appear to change, and women feel more freedom to visit their friends. In all, the coefficients we estimate in the cross-section are all positive, suggesting a movement towards more women's empowerment. Giving cash, even small amounts, to women appears to have had at least a small impact on their empowerment within households.

[12] We have also experimented with, for example, examining the first principal component of the matrix of outcomes; the results are similar but more difficult to interpret, so we stick with the simple sum.

6. Conclusion

Using fuzzy regression discontinuity design, we find some evidence of impacts on outcomes related to women's empowerment in Pakistan as a result of the receipt of the BISP transfer. Specifically, we find that the transfer appears to cause wives to report they are less likely to tolerate being beaten, and we find that men are more likely to agree that they should be expected to help around the house. Women also become more likely to report that they can visit friends without permission, and they become more likely to vote. Finally, it must be mentioned that the proportion of women with a CNIC card substantially increased between 2011 and 2013, particularly for women below the poverty score threshold; in the data set, the percentage of married women reporting having a CNIC number increased from 82 percent in 2011 to 89 percent in 2013. While we cannot statistically attribute that increase to the presence of the BISP transfer, it is highly unlikely that the increase would have been so large without the BISP.

Given that women's status is perceived to be low in Pakistan relative to many other countries, and that women's empowerment is correlated with economic growth (Duflo, 2011), any increases in women's empowerment are a good sign for Pakistan's economy in the future. As transfers have increased in size since the data for this paper were collected, and the transfers have increased in regularity, some of the impacts may have also become stronger since 2013. That said, given the limits of the questions asked as it stands in the data set, it would be interesting to take qualitative work already completed (OPM, 2014) on the BISP transfer to develop women's empowerment modules further to be able to understand how women's empowerment is affected by the BISP in a more nuanced, subtle manner.

References

Adato, M., D. Mindek, and A. Quisumbing. 2000. “The Impact of PROGRESA on Women’s Status and Intra-household Relations.” Washington: International Food Policy Research Institute.

Akram-Lodhi, A. H. 1996. “You Are Not Excused From Cooking: Peasants and the Gender Division of Labor in Pakistan,” *Feminist Economics* 2 (2): 87-105.

Akresh, Richard, de Walque, Damien, and Harounan Kazianga. 2013. “Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality,” World Bank Policy Research Working Paper no. 6340., World Bank, Washington, DC.

Ambler, Kate. 2016. “Bargaining with Grandma: The Impact of the South African Pension on Household Decision Making,” *Journal of Human Resources*.

Attanasio, Orazio, and V. Lechene. 2002. “Testing of Income Pooling in Household Decisions,” *Review of Economic Dynamics* 5(4): 720-748.

Attanasio, Orazio, Emla Fitzsimmons, and Ana Gómez. 2005. ``The Impact of a Conditional Education Subsidy on School Enrollment in Colombia,'' Report Summary: Familias 01, Institute for Fiscal Studies, London.

Attanasio, Orazio, and Alice Mesnard, 2006, “The Impact of a Conditional Cash Transfer Programme on Consumption in Colombia,” *Fiscal Studies* 27(4): 421-442.

Baird, Sarah, Ferreira, Francisco, Ozler, Berk, and Michael Woolcock. 2013. “Relative Effectiveness of Conditional and Unconditional Cash Transfers for Schooling Outcomes in Developing Countries: A Systematic Review,” *Campbell Collaboration Systematic Review* no. 8.

Bonilla, Juan, Rosa Castro Zarzur, Sudhanshu Handa, Claire Nowlin, Amber Peterman, Hannah Ring, and David Seidenfeld. 2016. “Cash for Women’s Empowerment? A

Mixed-Methods Evaluation of the Government of Zambia's Child Grant Programme," UNICEF Office of Research- Innocenti working paper no. WP-2016-01.

Calonico, Sebastian, Matias Cattaneo, Max H. Farrell, and Rocio Titunik. 2016. "Regression Discontinuity Designs using Covariates," working paper, University of Michigan.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titunik. 2014(a). "Robust Non-Parametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica* 82(6): 2295-2326.

_____. 2014(b). "Robust Data-Driven Inference in the Regression-Discontinuity Design," *Stata Journal* 14(4): 909-946.

Cattaneo, Matias, Michael Jansson, and Xinwei Ma. 2016. "rdrobust: Manipulation Testing Based on Density Discontinuity," University of Michigan working paper.

de Brauw, Alan, Daniel Gilligan, John Hoddinott, and Shalini Roy. 2014. "The Impact of *Bolsa Família* on Women's Decision Making Power," *World Development* 59: 487-504.

de Brauw, Alan, and Daniel Gilligan. 2012. "Using Regression Discontinuity Design with Implicit Partitions: The Impacts of *Comunidades Solidarias Rurales* on Schooling in El Salvador," IFPRI Discussion Paper no. 1116.

de Brauw, Alan, and John Hoddinott. 2011. "Must Conditional Cash Transfer Programs be conditioned to be effective? The impact of conditioning transfers on school enrollment in Mexico," *Journal of Development Economics* 96(2): 359-370.

de la O Campos, Ana. 2015. "Do Cash Transfers Empower Women? Impacts on Economic Advancement, Decision-Making, and Agency," PtoP (From Protection to Production) policy brief. Rome: Food and Agriculture Organization.

Doss, Cheryl. 2013. "Intrahousehold Bargaining and Resource Allocation in Developing Countries," *World Bank Research Observer*, 28(1).

Duflo, Esther. 2011. “Women’s Empowerment and Economic Development,” *Journal of Economic Literature* 50(4): 1051-1079.

_____. 2003. “Grandmothers and granddaughters: old age pension and intra-household allocation in South Africa,” *World Bank Economic Review* 17: 1-25.

Fiszbein, Ariel, and Norbert Schady. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*, World Bank: Washington, DC.

FAO (Food and Agriculture Organization). 2015. *The State of Food and Agriculture: Social Protection and Agriculture: Breaking the Cycle of Rural Poverty*. Rome: Food and Agriculture Organization.

Handa, Sudhanshu, Amber Peterman, Ben Davis, and Marco Stampini. 2009. “Opening Up Pandora's Box: The Effect of Gender Targeting and Conditionality on Household Spending Behavior in Mexico's *Progresa* Program,” *World Development* 37(6): 1129-1142.

Hoddinott, John, Emmanuel Skoufias, and Ryan Washburn. 2000. “The Impact of PROGRESA on Food Consumption,” International Food Policy Research Institute, Washington DC.

Hou, Xiaohui, and Ning Ma. 2013. “The Effect of Women’s Decision Making Power on Maternal Health Services: Evidence from Pakistan,” *Health Policy and Planning* 28(2): 176-184.

Imbens, Guido, and Josh Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects,” *Journal of Econometrics* 61(2): 467-476.

Jeejebhoy, Shireen, and Zeba A. Sathar. 2001. “Women’s Autonomy in India and Pakistan: The Influence of Religion and Region,” *Population and Development Review* 27(4): 687-712.

Kabeer, N. 2001. “Reflections on the Measurement of Women’s Empowerment.” In A. Sisask (Series Ed.), *Sida studies no. 3: Discussing Women’s Empowerment: Theory and practice* (pp. 17–57). Novum Grafiska AB: Stockholm, Sweden.

Lee, David, and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature* 48(2): 281-355.

Levy, Santiago. 2006. *Progress against Poverty: Sustaining Mexico’s Progresa-Oportunidades Program*. Washington, DC: Brookings Institution Press.

Lindert, K., A. Linder, J. Hobbs, and B. de la Brière. 2007. “The nuts and bolts of Brazil’s *Bolsa Família* program: Implementing conditional cash transfers in a decentralized context,” World Bank Social Protection Discussion Paper 0709, World Bank, Washington, D.C.

Maluccio, John, and Rafael Flores. 2005. “Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan *Red de Proteccion* Social,” Research Report 141, International Food Policy Research Institute, Washington, DC.

McCrary, Justin. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test,” *Journal of Econometrics* 142(2): 698-714.

Mumtaz, Zubia, and Sarah Salway. 2005. “’I Never Go Anywhere’: Extricating the Links Between Women’s Mobility and Uptake of Reproductive Health Services in Pakistan,” *Social Science and Medicine* 60(8): 1751-1765.

Oxford Policy Management. 2014. “Benazir Income Support Programme: First Follow-Up Impact Evaluation Report,” report prepared for the Government of Pakistan.

Quisumbing, A. and J. Maluccio. 2003. Resources at marriage and intrahousehold allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa” *Oxford Bulletin of Economics and Statistics* 65(3): 283-327.

Schady, Norbert, and Maria Caridad Araujo. 2008. ``Cash Transfers, Conditions, and School Enrollment in Ecuador,'' *Economía* 8(2): 43-70.

Schady, Norbert, and Jose Rosero. 2008. "Are Cash Transfers Made to Women Spent Like Other Sources of Income?" *Economics Letters* 101(3): 246-248.

Schultz, T. Paul. 2004. ``School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program,'' *Journal of Development Economics* 74(1): 199-250.

Thistlewaite, Donald, and Donald Campbell. 1960. "Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment," *Journal of Educational Psychology* 51: 309-317.

Thomas, D. 1990. Intrahousehold Resource Allocation: An Inferential Approach, *Journal of Human Resources* 25(4): 635-664.

Thomas, D. 1997. Incomes, expenditures and health outcomes: Evidence on intrahousehold allocation. In L. Haddad, J. Hoddinott, & H. Alderman (Eds.), *Intrahousehold resource allocation in developing countries*. Baltimore: Johns Hopkins University Press for IFPRI.

Trochim, William. 1984. *Research Design for Program Evaluation: The Regression-Discontinuity Approach*. Sage Publications: Beverly Hills, CA.

UNDP (United Nations Development Programme). 2010. *Country Profile of Human Development Indicators*, Accessed 16 May 2011. http://hdrstats.undp.org/en/countries/profiles/PAK.html

Yoong, J. Rabinovich, L. & Diepeveen, S. 2012. *The impact of economic resource transfers to women versus men: a systematic review*. Technical report. London, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Figure 1. Description of Sample by Beneficiary and Match Status, BISP Evaluation Follow-Up Survey, 2013

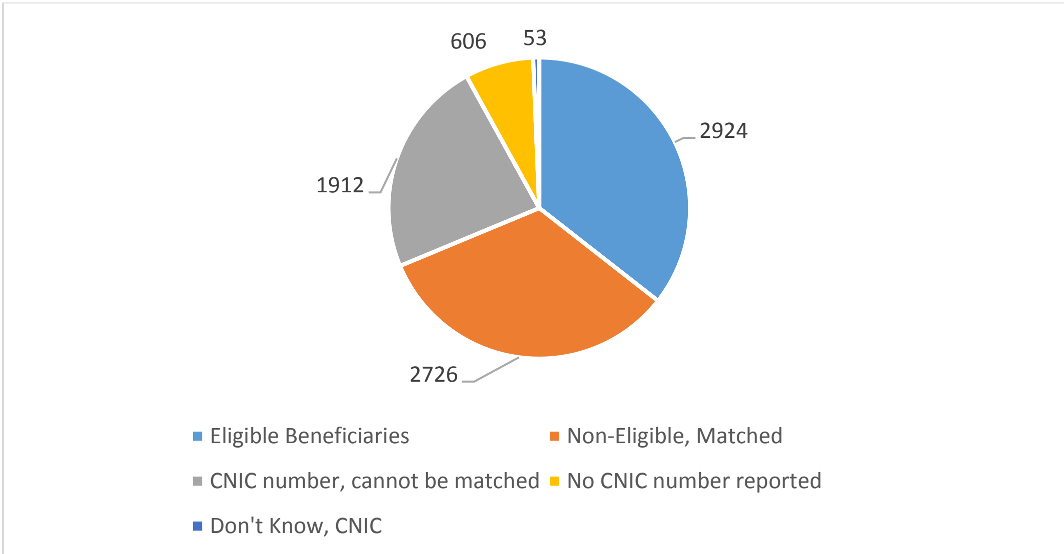


Figure 2. Probability of Receiving BISP Transfer versus Administrative Poverty Score, All Households with CNIC Match, 2013 Survey and BISP Administrative Data, Pakistan, 2013

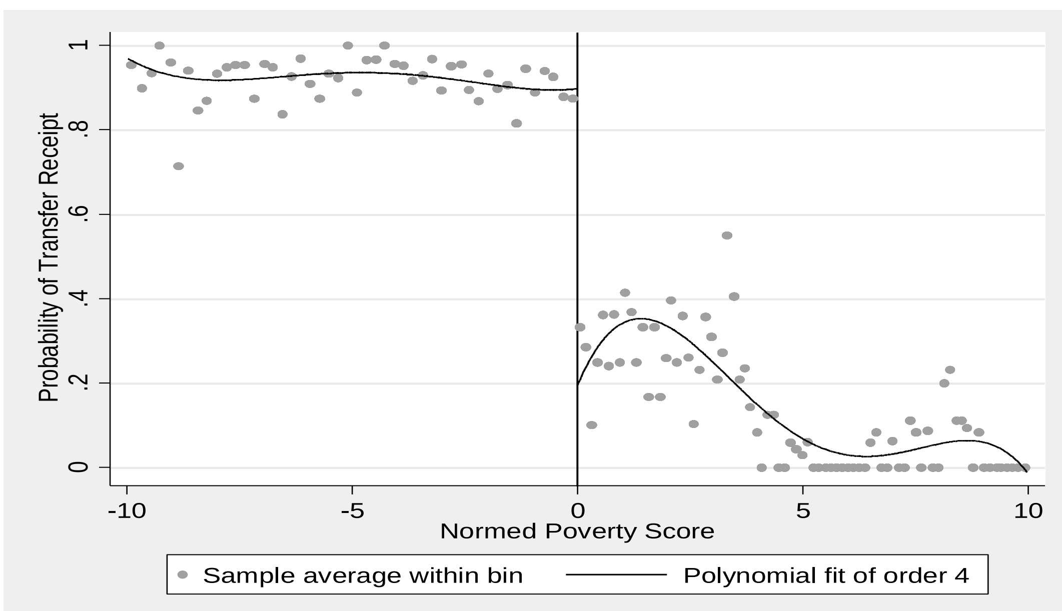


Figure 3. Histogram of Poverty Scores, 2013 Survey and BISP Administrative Data, Pakistan, 2013

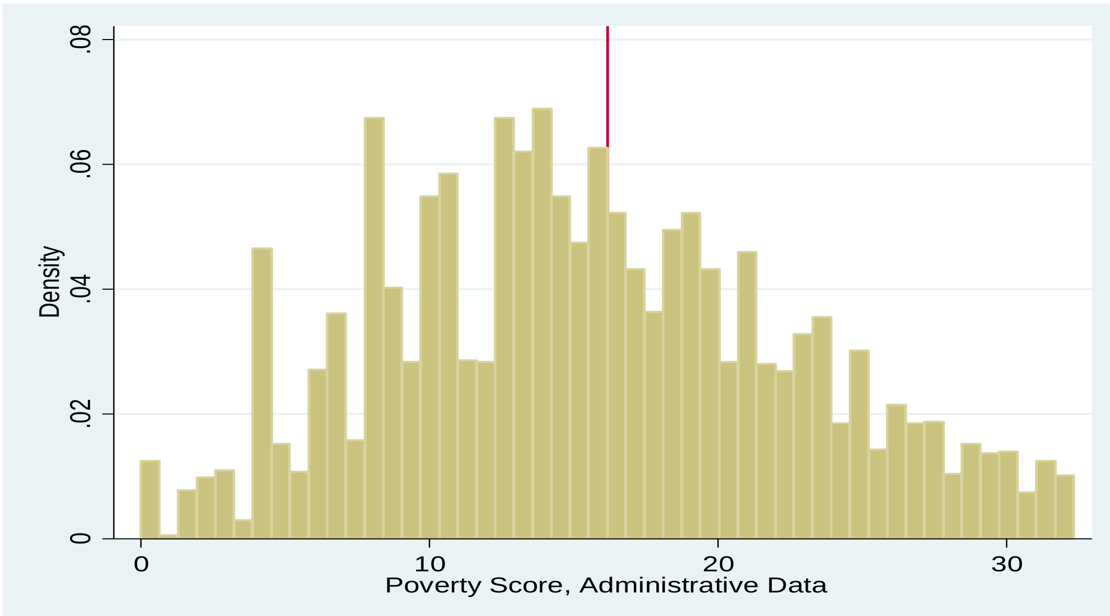


Figure 4. Household Demographic Characteristics, Baseline, versus Normed Poverty Score

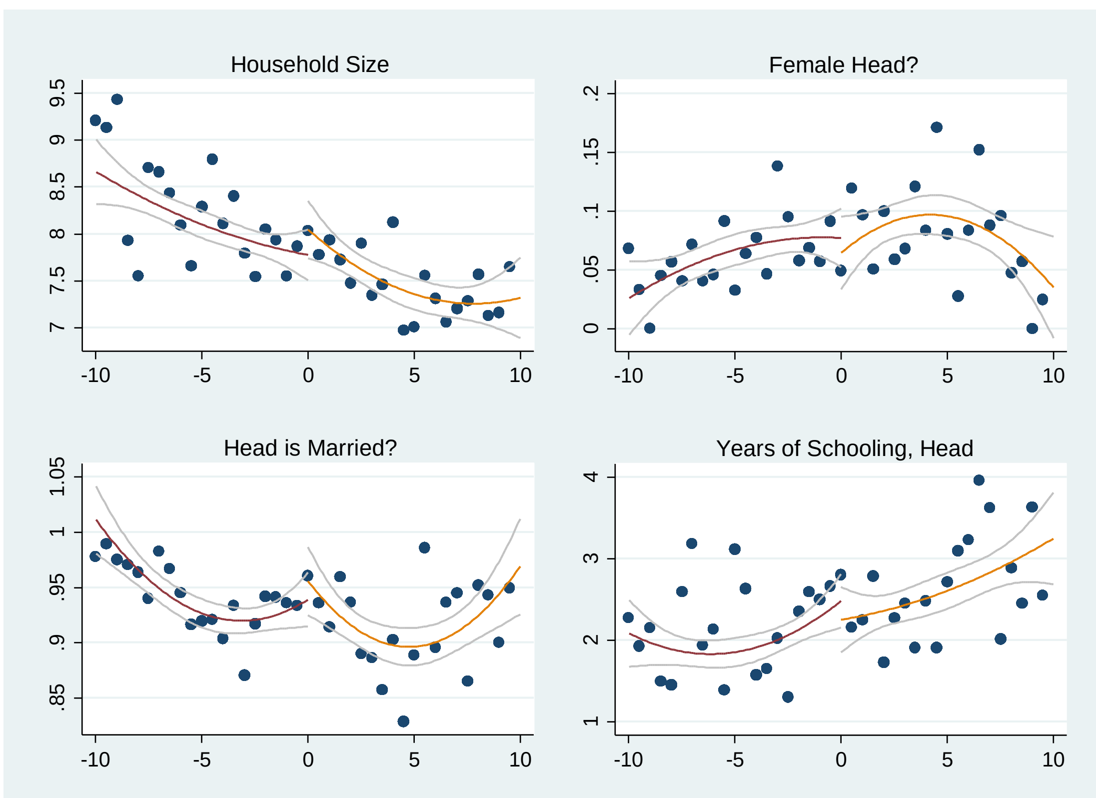


Figure 5. Selected Consumer Items Owned at Baseline, by Normed Poverty Score

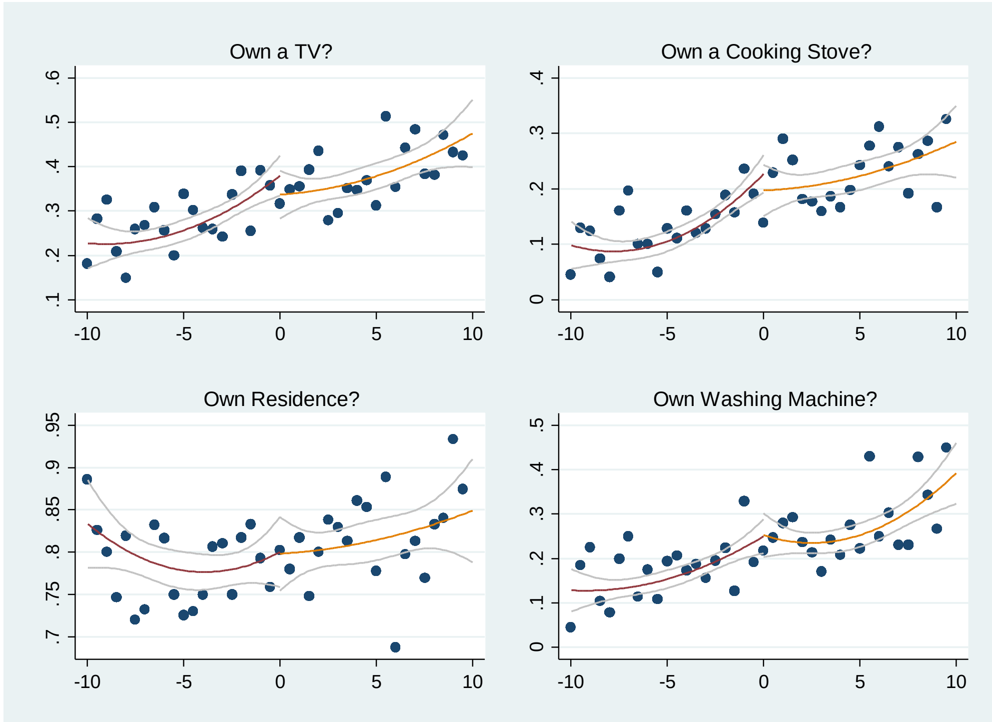


Figure 6. Housing Characteristics at Baseline, by Normed Poverty Score

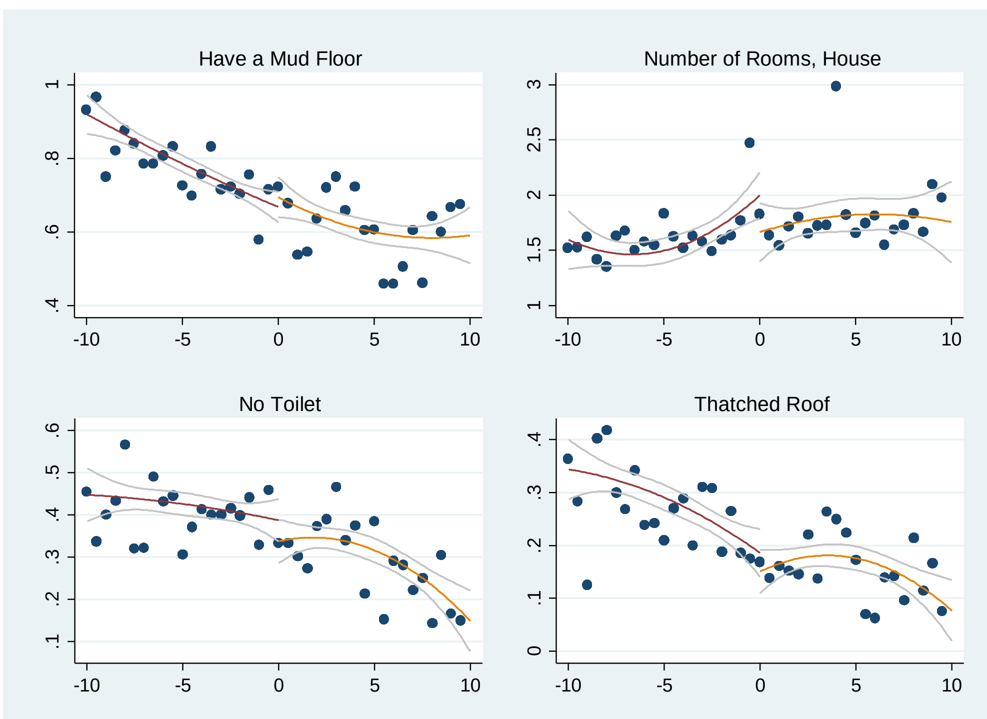


Figure 7. Number of Married Women in Household at Baseline, by Normed Poverty Score

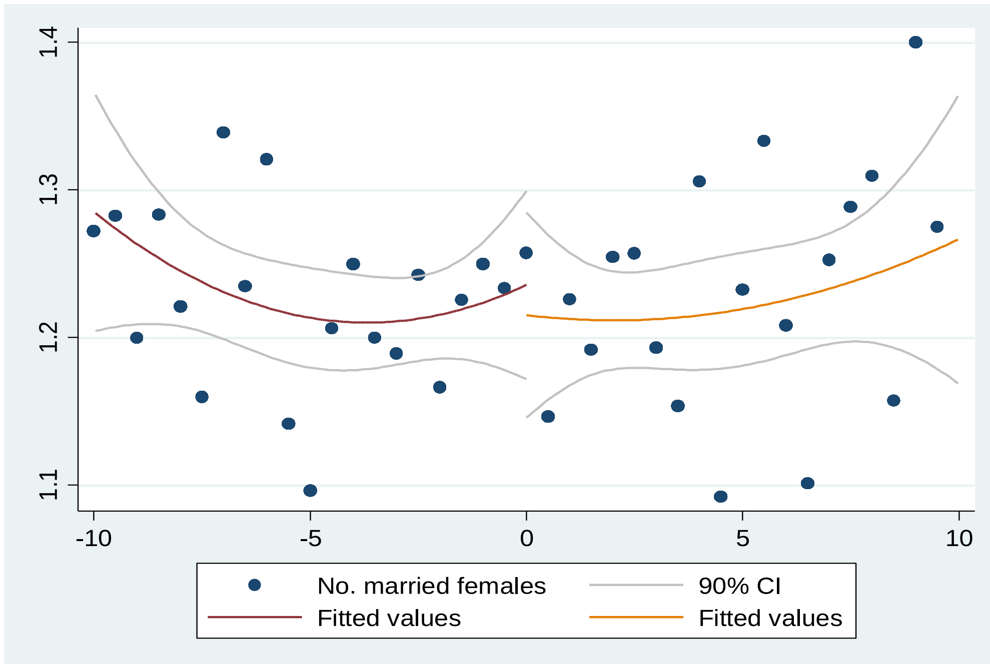


Figure 8. Graphical Representation of the Impact of BISP Transfer on Proportion of Women Disagreeing with Statement that they should tolerate being beaten, 2013 BISP Survey

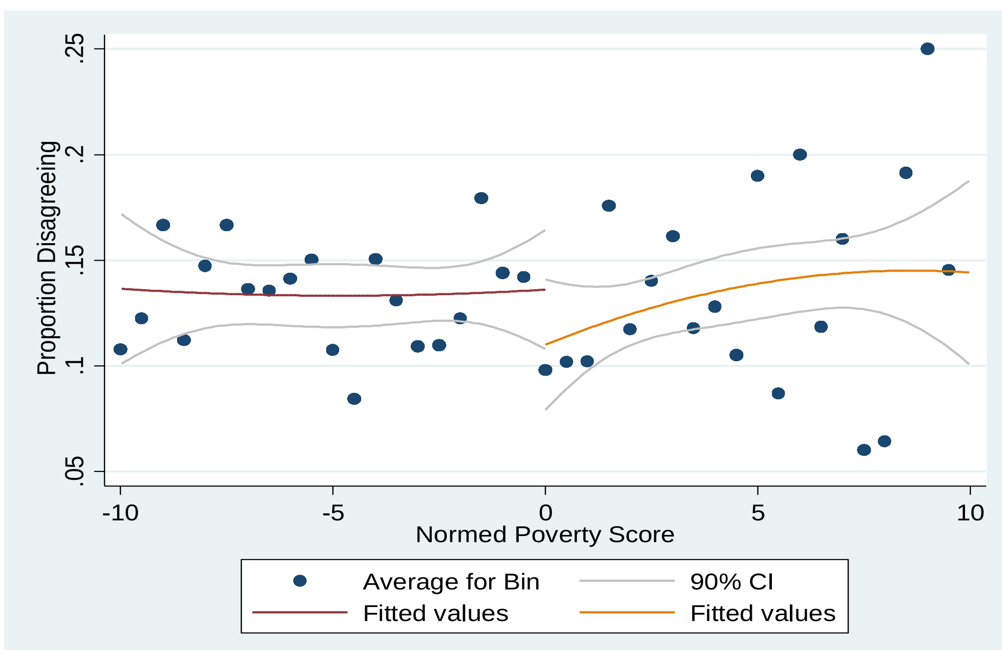


Figure 9. Graphical Representation of the Impact of BISP Transfer on Proportion of Men agreeing that men should help with chores, 2013 BISP Survey

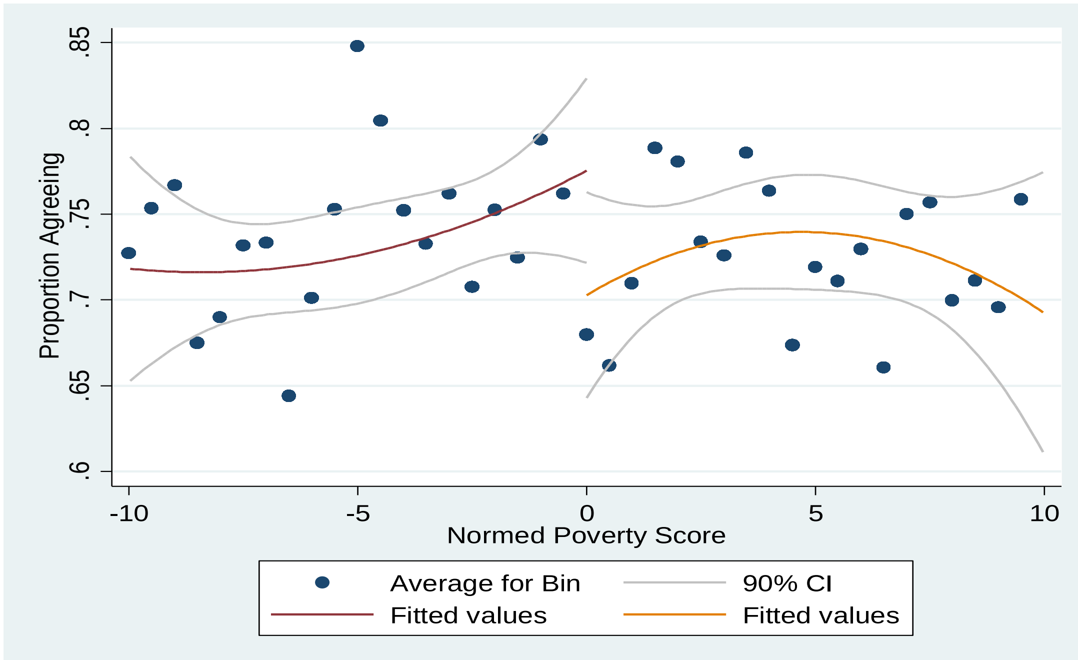


Figure 10. Graphical Representation of the Impact of BISP Transfer on Proportion of Women Stating they can visit a friend without permission, 2013 BISP Survey

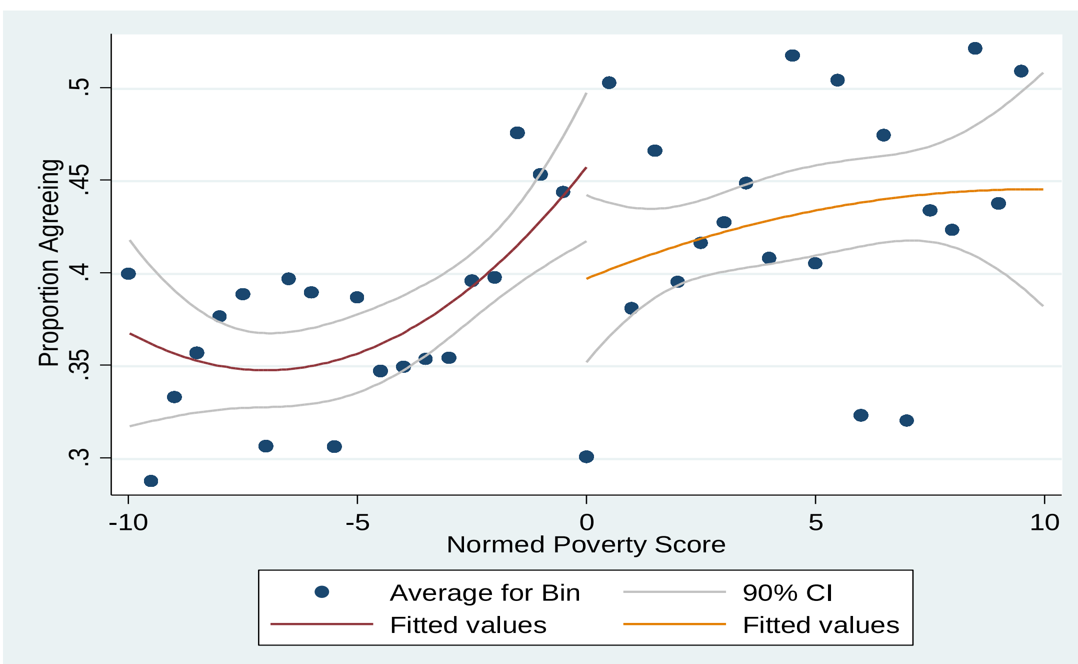


Figure 11. Graphical Representation of the Impact of BISP Transfer on Proportion of Women who can Vote, 2013 BISP Survey

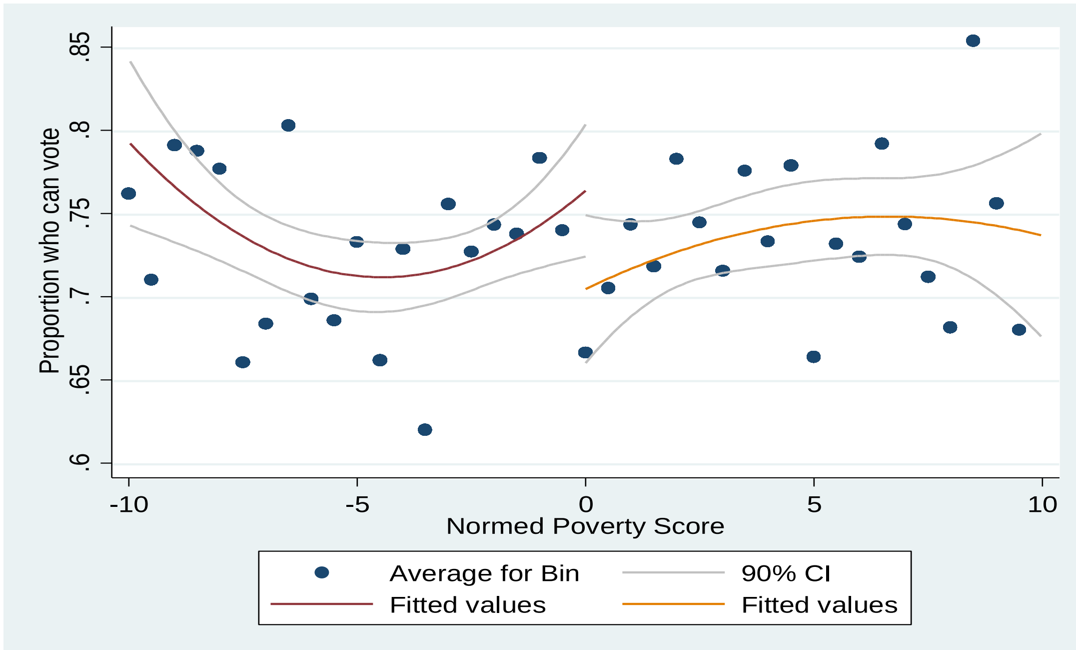


Table 1. Indicators of Women's Empowerment

Indicator	Source in Data Set
Gender Norms	
Disagree that male only should make important decisions	Women's Form and Men's Form. Combine "Agree/Strongly Agree" or "Disagree/Strongly Disagree" to construct variables.
Agree that male should help with household chores	
Agree that female should work outside	
Agree on wife to express opinion	
Disagree on wife to tolerate being beaten	
Disagree on it's better to send son to school than daughter	
Women's Mobility	
Female permitted to go to local market alone	Women's Form only.
Female permitted to go to local health facility alone	
Female permitted to go to friends' home alone	
Female permitted to go to a nearby shrine or mosque alone	
Voting	
Women report they can vote	Women's Form only.

Table 2. Final Household Level Sample Size, by Province, Oxford Policy Management BISP Impact Evaluation Surveys, 2011-2013

	Baseline Survey			**Follow-Up Survey**
Province	Group A	Group B	Total	Total
Punjab	1,456	1,706	3,162	3,017
Sindh	1,256	1,078	2,334	2,327
KPK	957	1,097	2,054	1,908
Baluchistan	489	636	1,125	969

Table 3. Estimates of Effect of Crossing Threshold on Probability of Receiving Payments from BIS Program, Administrative Data and 2013 Follow-Up Survey, Pakistan, 2013

	Full Sample (1)	Restricted Sample (2)
Indicator, Poverty Score Below 16.17	0.551***	0.548***
	(0.029)	(0.030)
Poverty Score	--0.010***	-0.010***
	(0.001)	(0.001)
Poverty Score interacted with Indicator	0.008***	0.009***
	(0.002)	(0.002)
Number of Obs.	5588	4939
R^2	0.604	0.638

Note: Standard errors clustered at the PSU level in parentheses. ***- indicates significance at the 1 percent level; **- indicates significance at the 5 percent level; *- indicates significance at the 10 percent level.

Table 4. Manipulation Tests, Based on Cattaneo et al. (2015a) method, Overall Sample and Estimation Subsample

	Full Sample (5650 Households)	Estimation Sample (4981 Households)
Undersmoothed Test Statistic	-1.43	-0.98
	0.153	*0.328*
Bias Corrected Test Statistic	-1.66	-1.52
	0.097	*0.127*

Note: p-values associated with test statistics in italics. Undersmoothed version of test reduces bandwidth to reduce bias associated with the parametric approximation of the test statistic, whereas the bias corrected version instead uses a higher order polynomial (than used in estimating densities) to attempt to model the bias in the test statistic.

Table 5. Baseline and Follow-Up Averages, Questions Related to Gender Norms, Estimation Sub-sample, Pakistan BISP Impact Evaluation Sample

	Men		Women	
	Baseline	2013 Survey	Baseline	2013 Survey
Disagree: Important decisions in the family should be made only by men	0.171	0.175	0.129	0.160
Agree: If wife works outside the home, husband should help with household chores	0.693	0.725	0.636	0.641
Agree: A married woman should be allowed to work outside the home if she wants to	0.476	0.454	0.660	0.683
Agree: Wife has right to express opinion even if she disagrees with husband	0.667	0.725	0.702	0.762
Disagree: A wife should tolerate being beaten by her husband	0.251	0.248	0.179	0.129
Disagree: Better to send a son to school than a daughter	0.673	0.760	0.696	0.652

Note: Numbers represent the proportion of positive answers among the estimation sample.

Table 6. Impacts of the BISP Transfer on Gender Norms from Female Perspective, Using both 2013 Survey and Differenced Outcome Variables and Varying Bandwidth

	(1) Disagree that only male should decide	(2) Agree that male should help with household chores	(3) Agree that female should work	(4) Agree on wife to express opinion	(5) Disagree on wife to tolerate being beaten	(6) Disagree that it is better to send son to school than daughter
Panel A: Estimates using 2013 survey outcomes						
Fixed bandwidth = 3						
BISP recipient	0.0758**	0.0830	0.0695	0.023	0.109	0.0593*
p-values for RD estimate	0.00588	0.107	0.106	0.597	0.257	0.0576
Sample Size, Left	1189	1189	1189	1189	1189	1189
Sample Size, Right	978	978	978	978	978	978
Fixed bandwidth = 5						
BISP recipient	0.0136	0.0375	0.034	0.004	0.084**	0.014
p-values for RD estimate	0.115	0.173	0.232	0.680	0.0437	0.276
Sample Size, Left	1850	1850	1850	1850	1850	1850
Sample Size, Right	1475	1475	1475	1475	1475	1475
Optimal bandwidth						
BISP recipient	0.0112	0.0486	0.039	0.001	0.077**	0.001
p-values for RD estimate	0.943	0.416	0.539	0.956	0.0465	0.940
Sample Size, Left	1895	1755	1787	2382	2175	2095
Sample Size, Right	1572	1356	1381	1806	1740	1660
Sample mean	0.161	0.651	0.684	0.765	0.131	0.663
Bandwidth	5.222	4.471	4.627	6.395	5.951	5.634
Panel B: Estimates using baseline covariates, including baseline value						
Fixed bandwidth = 3						
BISP recipient	0.0362	0.0398	0.0307	-0.0385	0.124	-0.0298
p-values for RD estimate	0.0851	0.391	0.192	0.849	0.170	0.346
Sample Size, Left	1001	1001	1001	1001	1001	1001
Sample Size, Right	809	809	809	809	809	809
Fixed bandwidth = 5						
BISP recipient	0.00147	0.00261	-0.00133	-0.0559	0.0879**	-0.0642
p-values for RD estimate	0.427	0.460	0.529	0.672	0.0393	0.947
Sample Size, Left	1543	1543	1543	1543	1543	1543
Sample Size, Right	1233	1233	1233	1233	1233	1233
Optimal bandwidth						
BISP recipient	0.0294	0.0466	0.0348	-0.0229	0.124**	-0.0293
p-values for RD estimate	0.405	0.426	0.492	0.812	0.0270	0.992
Sample Size, Left	1086	986	990	898	994	995
Sample Size, Right	854	771	784	681	797	809
Sample mean	0.162	0.650	0.679	0.760	0.127	0.671
Bandwidth	3.171	2.843	2.916	2.639	2.944	2.986

Note: Estimates of impact of receiving BISP transfers are calculated using the robust fuzzy regression discontinuity estimator developed in Calonico et al. (2015). The p-values reported are associated with a robust local-polynomial estimate, and are bias corrected, and estimation uses the triangular kernel. Sample sizes used in estimation to the right and left of the threshold are reported, as is the “optimal” bandwidth when it is calculated. Sample means in this table only refer to the observations within the optimal bandwidth.

Table 7. Impacts of the BISP Transfer on Gender Norms from Male Perspective, using both 2013 Survey and Differenced Outcome Variables and Varying Bandwidth

	(1) Disagree that only male should decide	(2) Agree that male should help with household chores	(3) Agree that female should work	(4) Agree on wife to express opinion	(5) Disagree on wife to tolerate being beaten	(6) Disagree that it is better to send son to school than daughter
Panel A: Estimates using 2013 survey outcomes						
Fixed bandwidth = 3						
BISP recipient	0.0034	0.218**	0.197*	0.0825	0.119	0.144
p-value	0.849	0.0555	0.251	0.257	0.447	0.0301
Sample Size, Left	520	520	520	520	520	520
Sample Size, Right	446	446	446	446	446	446
Fixed bandwidth = 5						
BISP recipient	0.0462	0.152*	0.135	0.109	0.0922	0.0955
p-value	0.870	0.0285	0.108	0.439	0.287	0.129
Sample Size, Left	818	818	818	818	818	818
Sample Size, Right	664	664	664	664	664	664
Optimal bandwidth						
BISP recipient	0.0726	0.113	0.137	0.0886	0.104	0.0934
p-value	0.372	0.119	0.133	0.254	0.118	0.235
Sample Size, Left	1059	1098	1035	1038	1077	967
Sample Size, Right	788	815	776	780	800	748
Sample mean	0.181	0.738	0.456	0.739	0.235	0.763
Bandwidth Used	6.262	6.553	6.138	6.214	6.376	5.806
Panel B: Estimates using baseline covariates, including baseline value						
Fixed bandwidth = 3						
BISP recipient	0.0523	0.306**	0.168	0.254*	0.117	0.215*
p-value	0.915	0.0992	0.312	0.0478	0.730	0.105
Sample Size, Left	264	264	264	264	264	264
Sample Size, Right	229	229	229	229	229	229
Fixed bandwidth = 5						
BISP recipient	0.0708	0.221*	0.136	0.279**	0.109	0.123
p-value	0.749	0.0284	0.288	0.101	0.497	0.0937
Sample Size, Left	435	435	435	435	435	435
Sample Size, Right	335	335	335	335	335	335
Optimal bandwidth						
BISP recipient	0.0529	0.174	0.137	0.251**	0.120	0.116
p-value	0.384	0.0804	0.309	0.0139	0.237	0.192
Sample Size, Left	600	515	582	572	507	582
Sample Size, Right	409	371	392	389	367	392
Sample Mean	0.179	0.742	0.451	0.738	0.240	0.761
Bandwidth Used	6.829	5.849	6.507	6.382	5.769	6.507

Note: Estimates of impact of receiving BISP transfers are calculated using the robust fuzzy regression discontinuity estimator developed in Calonico et al. (2015). The p-values reported are associated with a robust local-polynomial estimate, and are bias corrected, and estimation uses the triangular kernel. Sample sizes used in estimation to the right and left of the threshold are reported, as is the "optimal" bandwidth when it is calculated. Sample means in this table only refer to the observations within the optimal bandwidth.

Table 8. Impacts of the BISP Transfer on Perceived Mobility of Women, using both 2013 Survey and Differenced Outcome Variables and Varying Bandwidth

	(1) Female permitted to go to local market alone	(2) Female permitted to go to local health facility alone	(3) Female permitted to go to friends' home alone	(4) Female permitted to go to a nearby shrine or mosque alone	(5) Female votes always or sometimes
Panel A: Estimates using 2013 survey outcomes					
Fixed bandwidth = 3					
BISP recipient	0.0703	0.115*	0.167**	0.0760	0.109
p-value	0.281	0.0944	0.0372	0.163	0.401
Sample Size, Left	1189	1189	1189	1189	996
Sample Size, Right	978	978	978	978	808
Fixed bandwidth = 5					
BISP recipient	0.0387	0.0880	0.153**	0.0472	0.114
p-value	0.258	0.113	0.0431	0.208	0.195
Sample Size, Left	1850	1850	1850	1850	1544
Sample Size, Right	1475	1475	1475	1475	1222
Optimal bandwidth					
BISP recipient	0.0272	0.0600	0.133***	0.0332	0.0687*
p-value	0.544	0.168	0.00961	0.373	0.0892
Sample Size, Left	2113	2479	2445	2436	2176
Sample Size, Right	1676	1914	1853	1837	1727
Sample mean	0.264	0.296	0.402	0.230	0.728
Bandwidth Used	5.726	6.912	6.688	6.593	7.587
Panel B: Estimates using baseline covariates, including baseline value					
Fixed bandwidth = 3					
BISP recipient	0.0708	0.103	0.138*	0.0536	0.186
p-value	0.367	0.236	0.0911	0.352	0.147
Sample Size, Left	1001	1001	1001	1001	795
Sample Size, Right	809	809	809	809	621
Fixed bandwidth = 5					
BISP recipient	0.0393	0.0803	0.118*	0.0342	0.192**
p-value	0.290	0.157	0.0836	0.381	0.0321
Sample Size, Left	1543	1543	1543	1543	1224
Sample Size, Right	1233	1233	1233	1233	947
Optimal bandwidth					
BISP recipient	0.0740	0.106	0.139*	0.0545	0.153***
p-value	0.264	0.150	0.0766	0.409	0.0013
Sample Size, Left	986	989	995	989	1677
Sample Size, Right	783	783	806	783	1229
Sample Mean	0.278	0.303	0.413	0.228	0.728
Bandwidth Used	2.882	2.900	2.962	2.894	7.034

Note: Estimates of impact of receiving BISP transfers are calculated using the robust fuzzy regression discontinuity estimator developed in Calonico et al. (2015). The p-values reported are associated with a robust local-polynomial estimate, and are bias corrected, and estimation uses the triangular kernel. Sample sizes used in estimation to the right and left of the threshold are reported, as is the “optimal” bandwidth when it is calculated. Sample means in this table only refer to the observations within the optimal bandwidth.

Table 9. Impacts of Pakistan BISP Transfer on Simple Sum Indices of Women's Empowerment, using both 2013 Survey and Differenced Outcome Variables and Varying Bandwidth

	(1) Simple sum of Gender Norms, female responses	(2) Simple sum of Women's Mobility Measures	(3) Simple sum of Gender Norms, Mobility	(4) Simple sum of Gender Norms, Male Responses
Panel A: Estimates using 2013 survey outcomes				
Fixed bandwidth = 3				
BISP recipient	0.419*	0.428*	0.847**	0.847**
p-value	0.0157	0.0885	0.0115	0.0115
Sample size, left	1189	1189	1189	1189
Sample size, right	978	978	978	978
Fixed bandwidth = 5				
BISP recipient	0.187*	0.327*	0.514**	0.514**
p-value	0.0602	0.0991	0.0300	0.0300
Sample size, left	1850	1850	1850	1850
Sample size, right	1475	1475	1475	1475
Optimal bandwidth				
BISP recipient	0.193	0.233	0.418	0.632**
p-value	0.369	0.134	0.183	0.0230
Sample size, left	1847	2476	2140	816
Sample size, right	1444	1904	1685	652
Sample mean	3.048	1.195	4.242	3.136
Bandwidth used	4.890	6.889	5.784	4.878
Panel B: Estimates using baseline covariates, including baseline value				
Fixed bandwidth = 3				
BISP recipient	0.174*	0.366	0.537*	1.105*
p-value	0.0940	0.193	0.0592	0.065
Sample size, left	1001	1001	1001	264
Sample size, right	809	809	809	229
Fixed bandwidth = 5				
BISP recipient	-0.0113	0.273	0.261	0.929**
p-value	0.365	0.158	0.149	0.026
Sample size, left	1543	1543	1543	435
Sample size, right	1233	1233	1233	335
Optimal bandwidth				
BISP recipient	0.180	0.377	0.632*	0.827***
p-value	0.288	0.149	0.0961	0.006
Sample size, left	995	986	907	572
Sample size, right	806	783	691	389
Bandwidth used	3.058	1.220	4.305	3.101

Note: Estimates of impact of receiving BISP transfers are calculated using the robust fuzzy regression discontinuity estimator developed in Calonico et al. (2015). The p-values reported are associated with a robust local-polynomial estimate, and are bias corrected, and estimation uses the triangular kernel. Sample sizes used in estimation to the right and left of the threshold are reported, as is the "optimal" bandwidth when it is calculated. Sample means in this table only refer to the observations within the optimal bandwidth.

Social Protection & Labor Discussion Paper Series Titles
2015-2017

No.	Title
1702	The Impacts of Cash Transfers on Women's Empowerment: Learning from Pakistan's BISP Program by Kate Ambler and Alan de Brauw, February 2017
1701	Social Protection and Humanitarian Assistance Nexus for Disaster Response: Lessons Learnt from Fiji's Tropical Cyclone Winston, February 2017 by Aisha Mansur, Jesse Doyle, and Oleksiy Ivaschenko
1614	Urban Social Assistance: Emerging Insights from Three African Countries by Vanessa Moreira and Ugo Gentilini, December 2016
1613	Issues for Civil Service Pension Reform in Sub-Saharan Africa by Anita M. Schwarz and Miglena Abels, November 2016
1612	How to Target Households in Adaptive Social Protection Systems? Relative Efficiency of Proxy Means Test and Household Economy Analysis in Niger by Pascale Schnitzer, October 2016
1611	Pensions for Public-Sector Employees: Lessons from OECD Countries' Experience by Edward Whitehouse, October 2016
1610	Pension Systems in Sub-Saharan Africa: Brief Review of Design Parameters and Key Performance Indicators by Miglena Abels and Melis U. Guven, October 2016
1609	Household Enterprises in Fragile and Conflict-Affected States: Results from a Qualitative Toolkit Piloted in Liberia, Volume 2 – Annexes by Emily Weedon and Gwendolyn Heaner, August 2016
1608	Household Enterprises in Fragile and Conflict-Affected States: Results from a Qualitative Toolkit Piloted in Liberia, Volume 1 – Report by Emily Weedon and Gwendolyn Heaner, August 2016
1607	Benefits and Costs of Social Pensions in Sub-Saharan Africa by Melis U. Guven and Phillippe G. Leite, June 2016
1606	Assessing Benefit Portability for International Migrant Workers: A Review of the Germany-Turkey Bilateral Social Security Agreement by Robert Holzmann, Michael Fuchs, Seçil Paçacı Elitok and Pamela Dale, May 2016
1605	Do Bilateral Social Security Agreements Deliver on the Portability of Pensions and Health Care Benefits? A Summary Policy Paper on Four Migration Corridors Between EU and Non-EU Member States

by Robert Holzmann, May 2016

1604 Assessing Benefit Portability for International Migrant Workers: A Review of the France-Morocco Bilateral Social Security Agreement
by Robert Holzmann, Florence Legro and Pamela Dale, May 2016

1603 Assessing Benefit Portability for International Migrant Workers: A Review of the Belgium-Morocco Bilateral Social Security Agreement
by Robert Holzmann, Jacques Wels and Pamela Dale, May 2016

1602 Assessing Benefit Portability for International Migrant Workers: A Review of the Austria-Turkey Bilateral Social Security Agreement
by Robert Holzmann, Michael Fuchs, Seçil Paçaci Elitok and Pamela Dale, May 2016

1601 The Greek Pension Reform Strategy 2010-2016
by Georgios Symeonidis, July 2016

1507 Integrating Disaster Response and Climate Resilience in Social Protection Programs in the Pacific Island Countries
by Cecilia Costella and Oleksiy Ivaschenko, September 2015

1506 Effectiveness of Targeting Mechanisms Utilized in Social Protection Programs in Bolivia
by Ignacio Apella and Gastón Blanco, September 2015

1505 Kyrgyz Republic: Social Sectors at a Glance
by João Pedro Azevedo, Paula Calvo, Minh Nguyen and Josefina Posadas, August 2015

1504 Entering the City: Emerging Evidence and Practices with Safety Nets in Urban Areas
by Ugo Gentilini, July 2015

1503 Pension Patterns in Sub-Saharan Africa
by Mark Dorfman, July 2015

1502 Social Protection in Fragile and Conflict-Affected Countries: Trends and Challenges
by Mirey Ovadiya, Adea Kryeziu, Syeda Masood and Eric Zapatero, April 2015

1501 Defining, Measuring, and Benchmarking Administrative Expenditures of Mandatory Social Security Programs
by Oleksiy Sluchynsky, February 2015

To view Social Protection & Labor Discussion papers published prior to 2013, please visit www.worldbank.org/spl.

Abstract

Large-scale government cash transfer programs have become an important element of social protection and poverty reduction strategies throughout the developing world. Pakistan is no exception; in 2008, Pakistan established the Benazir Income Support Program (BISP) as an unconditional cash transfer targeted at the poorest of the poor. The primary goal of the BISP program is to provide the poorest households in Pakistan with unconditional transfers in order to improve their consumption and investments in children. To attain this goal, it is believed important that the transfers are provided directly to women to ensure the funds are spent as intended. Beyond changes in consumption and investment, directing these transfers to women can also serve to empower women by increasing household resources under their control. We analyze the impacts of Pakistan's BISP program on women's decision-making power within households using data collected between 2011 and 2013 as the program was rolling out. Using fuzzy regression discontinuity methods to statistically identify impacts, the BISP transfer is found to have substantial, positive impacts on some variables measuring women's decision-making power and empowerment.

About this series

For more information, please contact the Social Protection Advisory Service, The World Bank, 1818 H Street, N.W., Room G7-803, Washington, DC 20433 USA. Telephone: (202) 458-5267, Fax: (202) 614-0471, E-mail: socialprotection@worldbank.org or visit us on-line at www.worldbank.org/spl.