HHS Public Access

Author manuscript

Am Econ Rev. Author manuscript; available in PMC 2023 May 17.

Published in final edited form as:

Am Econ Rev. 2021 July ; 111(7): 2342–2375. doi:10.1257/aer.20200705.

On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms†

Erica Field,
Duke University

Rohini Pande,
Yale University

Natalia Rigol,
Harvard Business School

Simone Schaner,
University of Southern California

Charity Troyer Moore
Yale University

Abstract

Can increasing control over earnings incentivize a woman to work, and thereby influence norms around gender roles? We randomly varied whether rural Indian women received bank accounts, training in account use, and direct deposit of public sector wages into their own (versus husbands') accounts. Relative to the accounts only group, women who also received direct deposit and training worked more in public and private sector jobs. The private sector result suggests gender norms initially constrained female employment. Three years later, direct deposit and training broadly liberalized women's own work-related norms, and shifted perceptions of community norms.

JEL

G51; G53; J16; J31; O12; O16; Z13

Female labor force participation (FLFP) remains low and stagnant in many emerging economies, and India is a particularly stark example. Despite robust economic growth, India's FLFP declined from 32 percent in 2005 to 21 percent in 2018, making Indian women some of the least employed in the world (ILO 2020). Yet, nearly one-third of Indian housewives express an interest in working (Fletcher, Pande, and Moore 2018). Simply bringing these latent workers into the labor force would effectively double Indian FLFP.[1] What stops so many women who want to work from joining the labor force?

†Go to https://doi.org/10.1257/aer.20200705 to visit the article page for additional materials and author disclosure statements.

emf23@duke.edu .

One possibility is conservative gender norms around work roles, a phenomenon prevalent in India and beyond. In World Values Surveys spanning 60 countries between 2010 and 2014, one-third of respondents stated that when women earn more than husbands it causes problems in the household, and nearly one-half state that children suffer when their mother works. In many countries, a wife who works outside the home is a source of social stigma or shame for her husband, who is expected to be the primary breadwinner (Boudet et al. 2012, Bernhardt et al. 2018). When internalized by women, such norms directly lower their utility of working (Akerlof and Kranton 2000). When internalized by men, these norms may also reduce women's work through intrahousehold channels (Bertrand, Kamenica, and Pan 2015).

In this paper we look for evidence that norms constrain rural Indian women's labor supply by studying the impact of an exogenous increase in a woman's control over earned income. Under the canonical collective household model, this change should strengthen her bargaining power and, thereby, her consumption of both goods and leisure (i.e., a shift in bargaining power has an income effect). We show that this prediction of reduced labor supply can be flipped if husbands bear norm costs when their wives work. Specifically, increases in a woman's bargaining power can, by reducing the weight placed on her husband's preferences, induce her to *enter* the labor market. Thus, a rise in women's work in response to an exogenous increase in women's control over earnings offers an indirect means of empirically identifying norm-based barriers to female labor supply.

To study this prediction, we leverage a randomized controlled trial covering 197 village clusters (gram panchayats (GPs)) in northern Madhya Pradesh.[2] We focused on the government workfare program, the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS). The status quo was for female workers' MGNREGS wages to be deposited in the male household head's bank account. Hence, to increase women's control over earnings, we worked with banks to open individual accounts for women in all treatment GPs, and in one-half of the treatment GPs, coordinated with state authorities to designate these accounts to receive direct deposit of MGNREGS wages. In addition, we cross-randomized a short training on how to use the local bank kiosks that serviced these accounts. This generated five intervention arms: pure control, own account ("accounts only"), own account and training, own account and direct deposit, and own account, direct deposit, and training ("direct deposit and training," from now on, D^2T).

While our analysis reports separate impacts by intervention arm, we are most interested in the comparison of outcomes between accounts only and D^2T GPs, which isolates the effect of increasing a woman's control of her MGNREGS income holding financial inclusion constant. Given women's level of MGNREGS participation, D^2T gave women control over a sizable endowment, making shifts in bargaining power plausible: over the three-year study

[1]Female employment has also been shown to delay marriage, increase female work aspirations, improve child health, and reduce the male:female sex ratio (Qian 2008, Atkin 2009, Jensen 2012, Heath and Mobarak 2015). In the United States, rapid growth in FLFP preceded important changes in gender role norms (Goldin 2006).

[2]MP is the sixth largest and eighth poorest of India's 29 states, with particularly restrictive gender norms. Adult male and female rural labor force participation rates are 84 percent and 29 percent (close to the national averages) (RBI 2016). GPs, the lowest level of government, typically comprise 2–5 villages in MP.

period, women in D^2T who received MGNREGS direct deposits were paid an amount roughly equal to their annual private sector earnings.

We analyze impacts of the intervention on labor supply utilizing a combination of administrative data and two waves of household survey data conducted one and three years after intervention implementation. Pooling across survey rounds, treated women in D^2T GPs scored 0.11 standard deviation units higher on a labor market engagement index, with significant gains in both public and private sector work. The observed increase in female work in the private sector, where wages were never deposited to bank accounts, is inconsistent with standard efficient household models.

To reconcile the puzzle, we incorporate norms costs into the model: suppose conservative norms cause a woman and her husband to incur (potentially different) utility losses when she works. In a collective household, a gain in her bargaining power now has the additional effect of putting more weight on her labor supply preferences relative to her husband's. Under D^2T, a woman who previously stayed at home due to norm costs that her husband faced may now enter the labor market. Hence, we predict larger effects among "constrained" women who, absent the intervention, stay out of the labor market to avoid norms costs.

In our empirical analysis, our best proxy for being constrained is never having worked for MGNREGS at baseline. Absent intervention, these women are less likely to work, rank lower on a bargaining and agency index, and their husbands associate having a wife who works with greater social stigma. Treatment effects for constrained women (at 0.21 standard deviation units for the overall work index and 0.29 standard deviation units for the private sector work index) exceed those for unconstrained women. Further, consistent with our theoretical prediction that D^2T only increases women's private sector work if norm costs bind, we only see significant effects on private sector work among constrained women.

While labor supply effects persist in the long run for constrained women, they attenuate among unconstrained women. It could be that unconstrained women experienced a stronger income effect or that, within the accounts only arm, unconstrained women were better placed to respond to independent government efforts to enable MGNREGS direct deposit to individual accounts, described below.

Greater earnings control not only encouraged women to work but, in doing so, led to 0.15 and 0.12 standard deviation unit increases in indices of account use and banking autonomy, respectively. The latter captures important gains in female agency, including whether a woman goes to the bank on her own, is comfortable transacting independently, and prefers receiving wage payments into her account. While other markers of female agency and empowerment did not respond to earnings control on average, D^2T led to significant gains in the empowerment index for constrained women, paralleling our labor supply results.

Finally, we explore impacts on norms, as measured by long-run survey data on male and female attitudes toward women's work and their perceptions of community members' attitudes. Our interventions did not seek to directly alter these and we, therefore, do not anticipate norm changes as a mechanism underlying short-run labor supply changes. In the longer run, we acknowledge that norm changes could amplify the impacts of D^2T

on female labor supply. While our framework does not explicitly model how D^{2}T would influence gender norms, the existing literature suggests norms are, in part, shaped by the economic environment (Alesina, Giuliano, and Nunn 2013) and transmitted through social learning (Fernández, Fogli, and Olivetti 2004; Fernández and Fogli 2009). In our setting, the experience of working or having a spouse who works may lead individuals to update their beliefs about the propriety of female work. Moreover, if a woman and her husband face lower than expected stigma when she works then they may update their perceptions about community norms. Finally, individuals who see more women in the community working may update both their own norms and perceived norms of others.

D^{2}T influenced own and perceived attitudes towards female employment, with the patterns differing by gender. Among women, D^{2}T liberalized the own norms index by 0.10 standard deviation units and the perceived norms index by 0.08 standard deviation units. While the male own norms index did not change, D^{2}T increased the male perceived norms index by 0.09 standard deviation units. This shift is largely driven by a reduction in perceived social stigma falling on husbands of working women.

In investigating longer-run changes in norms, we depart from our pre-analysis plan in two ways. First, the long-run survey expanded beyond prespecified norms measures ("male attitudes toward female work") to include measures of both women's and men's own beliefs and perceptions of community members' beliefs about women and work.[3] Second, as a test of our theory, we evaluate heterogeneous treatment impacts based on whether a woman is likely "norms constrained." These departures reflect the salience of norms around women's work we observed in field-based interactions during the intervention and short-run survey, and our resulting interest in testing whether empirical data supported the underlying model posited above.

Our analysis considers multiple treatment arms and families of outcomes. Our pre-analysis plan outlined main families of outcomes and an empirical approach without completely tying our hands in terms of final analysis. Importantly, our PAP stated we would compare impacts of treatments to the control group and one another, implying 10 hypothesis tests per outcome. Our main tables feature 7 of these 10 tests. As guidance for assessing multiple comparison concerns with our subsequent analysis, we estimate sharpened q-values that control the false discovery rate (FDR) (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008). The adjustment pools all 10 hypothesis tests and all outcomes into a single family to account for the PAP's generality.[4] The q-values for D^{2}T estimates related to female labor supply and female own norms remain significant at 5 percent or less after this adjustment, while female perceived norms and account use are significant at the 10 percent level. Given this, we consider our findings that D^{2}T impacted women's labor supply, account use and, in the longer run, their norms as reasonably robust.

[3]As we did not update the pre-analysis plan between survey waves, an earlier paper focused on the short-run results (Field et al. 2016) and a grant application (both written prior to long-run data collection) provide the best ex ante plans for the norms analysis. Field et al. (2016) concludes by highlighting an interest in studying long-run impacts on norms. Other than norms, the other families of outcomes we consider, financial inclusion, labor supply, and empowerment, are listed in our pre-analysis plan. For details, see the registry and associated documents at https://www.socialscienceregistry.org/trials/115.

[4]FDR cannot account for discretion in constructing and featuring outcomes. For this we rely on the PAP and disclosing departures from it.

Turning to the related literature, our work complements Bursztyn, González, and Yanagizawa-Drott (2020), who show that in settings with mis-perceived norms, information provision can change labor market behaviors. We show that without directly targeting norms or norm-linked perceptions, large-scale policies that alter women's incentives to work can shift own beliefs over a relatively short time horizon. We also find evidence that perceived norms move, though in light of FDR adjustments we view results for men as more speculative.

Our paper also contributes to a large and growing literature on the gendered impacts of cash transfer programs. Many that focus on women's empowerment and bargaining power do not consider FLFP (Bobonis 2009, 2011; Attanasio and Lechene 2014; Almås et al. 2018). Those that do consider FLFP typically do not find increases, possibly because the income transfers are sizable enough to reduce labor supply (Hasan 2010; Skoufias, Unar, and de Cossio 2013). A second, related, set of papers demonstrates that productive asset transfers (coupled with additional support) can increase the labor supply of women in very poor households across a range of country contexts (Bandiera et al. 2017, Banerjee et al. 2015, Bedoya et al. 2019). Different from these papers, we study an intervention that only varied women's *control* over potential earnings, not the resources made available to the household. Separating the impact of control on women's economic lives from that achieved by providing more resources is key for deepening our understanding of how households make decisions. From a policy perspective, these concepts map to distinct policy strategies, and highlight an opportunity to improve the design of existing programs, such as MGNREGS.

More broadly, our paper contributes to recent research on social protection program design, which typically focuses on delivery efficiency (e.g., Muralidharan, Niehaus, and Sukhtankar 2016; Aker et al. 2016; Banerjee et al. 2020; Bachas et al. 2020 examines impacts on financial inclusion). We show that gender targeting can impact not only program outcomes (e.g., work days provided through MGNREGS) but also broader economic outcomes that have the potential to outlive the program (e.g., private sector work and gender norms).

The rest of the paper proceeds as follows. Section I describes the study context and experimental design. Section II provides a conceptual framework for evaluating treatment effects, and Section III the empirical strategy. Section IV discusses treatment impacts and Section V concludes.

I. Experimental Context and Design

We first describe work opportunities and gender norms in rural India and then outline the intervention design.

A. Work Opportunities and Gender Norms in Rural India

In rural India, work opportunities for both men and women with low levels of education (like those in our sample) are typically limited to unskilled labor. In the private sector, this includes self-employment in agriculture and seasonal casual wage labor on others' land or construction sites. In the public sector, MGNREGS entitles rural households to up to 100

days of unskilled work per year, although in practice, MGNREGS work opportunities are sufficiently scarce that the day limit rarely binds. Nevertheless, MGNREGS is one of the largest household-level redistribution programs in India and, indeed, the world (Subbarao et al. 2013), with annual participation frequently topping 50 million households.

In terms of rural Indian women's willingness to take advantage of work opportunities, survey data suggest that spousal preferences matter, even over and above personal beliefs. The 2011–2012 India Human Development Survey, for instance, found that 52 percent of adult women stated that her husband has the most say as to whether she works. Long-run survey data for our control (status quo) group in Table 1 show that our study sample is even more conservative: only 28 percent of women stated they have a say in whether they work. Men report more conservative beliefs about the appropriateness of women's work and report greater social stigma from female work: while 74 percent of women agree that "women can work," just 66 percent of their husbands agree. The average woman reports that 39 percent of community members would speak badly about a woman who works, while the average man reports that 56 percent of community members would think the husband of a working woman is a bad provider. Further, these views correlate with economic outcomes: Bernhardt et al. (2018) found male perceived norms to be predictive of women's work in this sample.

More broadly, the four districts of MP covered by our study are marked by severe gender inequities; for instance, sex ratios in these districts range from 0.84 to 0.90 females to every male (India Census 2011). Our sample consisted of poor, middle aged women with very limited education (less than one year on average), see Table 1. Their husbands are on average 5 years older and have 3 more years of schooling. Over 40 percent of households belong to India's most disadvantaged social groups, scheduled castes or tribes. At the time of the long-run survey, at Rs 980 per capita, average monthly household income in our sample was below the rural MP poverty line of Rs 1,036.[5]

We are interested in identifying women who face labor supply constraints due to, e.g., conservative gender norms. Our proxy is based on the (only) female labor supply measure collected in the baseline survey: we consider a woman to be unconstrained if she ever worked for MGNREGS. To confirm its relevance, Table 1 compares labor force participation and gender norm measures across constrained and unconstrained women, using long-run survey data for the control group. Relative to unconstrained women, constrained women were 11 percentage points (18 percent) less likely to have worked for pay in the past month. Constrained women scored 0.07 standard deviations lower on our empowerment index, which captures the woman's control over purchases, self-reported decision making power, mobility, and freedom from gender-based violence. We observe no significant differences across constrained and unconstrained women or their husbands on whether they report women can work. However, men married to constrained women perceive greater social stigma associated with having a wife who works; they report nearly a 7 percentage point (12 percent) higher fraction of the community would think the husband of a working woman is a bad provider. Among women, we observe smaller and noisier differences in

[5]We calculate the poverty line by taking the latest poverty line for rural MP from the Reserve Bank of India (Rs 771 in 2011/2012) and inflating it to 2017 terms using the IMF CPI for India.

the same direction. The final table row provides some insight on the caste origins of these norms: constrained women are more likely to belong to castes with stronger norms against women's work, as identified using the 2005–2006 Demographic and Health Survey (DHS) Demographic and Surveys (2006) (online Appendix Section E.3 provides details on the DHS analysis).

B. Experimental Design

Our experiment builds on the 2008 policy reform to MGNREGS payment architecture, wherein states were required to transition from cash to electronic payment of MGNREGS wages into beneficiary-owned bank accounts. Since women typically lacked their own accounts, the status quo was electronic payment of all household member wages into a single account, almost always owned by the male head of household. In 2012 the Minister of Rural Development explicitly mandated that a woman's MGNREGS wages should be deposited into her individual bank account (UNWOMEN 2012).

Also relevant for our study was the community banking initiative launched in our study state, Madhya Pradesh (MP), in 2011. This initiative sought to ensure that citizens had access to a "last-mile" bank kiosk within 5 kilometers of their residence. Importantly, kiosk bank accounts could only be accessed with an authenticated fingerprint.

Together, these reforms made it possible for women to have MGNREGS wages directly deposited into private, easy-to-access, secure accounts. However, officials were slow to target women. For instance, in our study areas, rates of payment into individual bank accounts among female workers remained below 20 percent until 2016. This provided us ample opportunity to experimentally vary women's access to individual bank accounts and whether those accounts received direct deposits of MGNREGS wages.

In our study districts, we identified and then randomly assigned all 199 GPs with functional bank kiosks to one of three groups: 66 GPs formed the control group, 68 GPs were to receive bank accounts for eligible women, and 65 GPs were to receive bank accounts and direct deposit of MGNREGS wage into their new accounts. Using MGNREGS administrative data we identified households in these GPs that were listed as having worked for MGNREGS between October 2012 and October 2013. Between November 2013 and January 2014, we conducted a rapid screening of these 14,088 households. A married couple entered our sample if at least one household member reported having ever worked for MGNREGS and the wife lacked an individual bank account. We identified 5,851 eligible couples and two GPs without any eligible couples. These two GPs (both assigned to the control group) were dropped, leaving us with 197 GPs. Online Appendix Figure A2 and online Appendix Section E.1 provide a timeline of experimental activities and randomization details, respectively.

In all 133 treatment GPs, our team individually informed eligible women of an upcoming account opening drive where they could open a bank account at the kiosk, free of charge. On the day of the drive, our team returned to the household to invite the woman to visit the kiosk with her documents (proof of address and a passport-sized photo) and open an account. The team facilitated the account opening process at the kiosk.

In GPs assigned to the direct deposit treatment arm, our team additionally informed eligible women of the option to have their MGNREGS payments deposited in their (newly opened) bank account. Conditional on consent, our team submitted a request to enter their newly opened individual bank account into the MGNREGS administrative system ensuring that her wages would be directed into her new account rather than her husband's account.

Training was randomized as a third, cross-cutting treatment in one-half of the GPs selected for bank accounts or bank accounts and direct deposit. In GPs assigned to the training intervention, following the account opening camps, eligible women were invited to a group-based information session. The sessions familiarized women with procedures for deposits and withdrawals at the kiosk. They also provided women information such as account uses (including saving and receiving benefit transfers), why kiosk deposits were safe, and the time and cost savings of kiosk transactions.

To summarize, we created five intervention arms: control (64 GPs), accounts only (32 GPs), accounts and direct deposit (34 GPs), accounts and training (33 GPs), and accounts, direct deposit, and training (34 GPs), which we refer to as D^2T going forward.

II. Conceptual Framework

As a precursor to the empirical analysis, we modify a simple collective household model to examine how the presence of gender norms against women working moderates the impact of D^2T on FLFP. Among our intervention arms, D^2T maximized a woman's control over her earnings: her MGNREGS wages were deposited in her own account (instead of her husband's) and the training strengthened her ability to use that account.

A. Setup

Endowments and Wages: The household consists of a husband and wife, $i \in \{F, M\}$. Each has nonlabor income y^i. a time endowment of 1, and can supply labor h_s^i in sectors $s = P$ (private) and N (public/MGNREGS), for wages w_s^i. Consistent with program implementation, spouse i's MGNREGS labor supply is capped at $\overline{N}$ units.[6]

Preferences: Each spouse values private consumption c^i and leisure l^i according to the function $u^i(l^i, c^i)$.[7] A woman working can violate norms such as "the wife takes care of the household" and "the husband is the breadwinner." We capture such norm costs by a fixed utility cost $\gamma^i \geq 0$ which could include either, or both, "own norms" costs (i.e. the psychic cost to individuals of violating personal beliefs about gender roles), and "perceived norms" costs, (i.e., expected social stigma cost imposed by community members who disapprove of women working). We focus on a fixed cost because gender norms in India have a strong caste component, and there is empirical evidence of fixed norms costs related to caste and

[6]The MGNREGS act specifies a 100-day cap at the household rather than at the individual level. However, in practice, there are sufficiently scarce MGNREGS work opportunities available to households such that the cap is more appropriately modeled as an individual limit that is determined by the number of available projects. To streamline analysis we omit, without loss of generality, hours constraints for private sector work.

[7]Throughout, we assume that each $u^i(l^i, c^i)$ is a twice differentiable, increasing and concave utility function, that the cross derivative $\partial^2 u^i / \partial l^i \partial c^i$ is null, and that the standard Inada conditions are satisfied

labor supply (Oh 2020). A broader interpretation of γ^i would include other fixed costs associated with a woman working, such as the time and hassle of securing childcare.

Norms constrain labor force participation for two categories of women. These categories are not necessarily mutually exclusive, especially when both spouses bear norms costs. First, those kept of the labor force by self-internalized norms.

DEFINITION 1: *A woman is personally constrained if* $\gamma^F > 0$ *and she does not work but, holding other parameters constant, she would work if* $\gamma^F = 0$.

Second, those for whom husband preferences bind.

DEFINITION 2: *A woman is spousally constrained if* $\gamma^M > 0$ *and she does not work but, holding other parameters constant, she would work if* $\gamma^M = 0$.

Decision-Making: Households allocate consumption and leisure efficiently. Specifically, labor supply decisions maximize a Pareto-weighted average of husband and wife utilities, subject to the household budget constraint. We assume the wife's Pareto weight, $\mu \in (0, 1)$, depends on nonlabor income and other "distribution factors" that affect a woman's outside option but do not enter the budget constraint (Blundell, Chiappori, and Meghir 2005). Given the rarity of divorce, we anticipate the relevant outside option to be a noncooperative equilibrium where spouses do not share resources (Lundberg and Pollak 1993).

We build on Chiappori's (1992) two-stage representation of the household allocation problem. In the first stage, a lump-sum transfer ϕ^F between husband and wife effectively chooses a point on the Pareto frontier. This transfer could be positive or negative and is generically increasing with μ, which captures a woman's bargaining power/outside option.[8] In the second stage of the canonical model, each spouse maximizes own utility subject to an individual budget constraint. Online Appendix Section D shows that with norm costs, the woman's allocation instead solves the following problem in the second stage:

$$\max_{h_N^F,\, h_P^F,\, c^F} u^F\left(1 - h_N^F - h_P^F, c^F\right) - \left(\gamma^F + \frac{1-\mu}{\mu}\gamma^M\right)\mathbf{1}\left(h_P^F + h_N^F > 0\right), \tag{1}$$

subject to

$$c^F \leq w_N^F h_N^F + w_P^F h_P^F + \phi^F, \quad h_s^F \geq 0, \quad h_N^F \leq \overline{N},$$

where $\mathbf{1}(\cdot)$ is the indicator function. The key difference between our setup and the standard collective model comes from the norms externality that a woman imposes on her husband should she work. Program (1) shows that she internalizes her husband's preference that she not work in a manner proportional to her relative Pareto weight.

[8] With fixed costs, when a woman enters the labor force the household switches ϕ^F schedule. If a higher μ caused labor force entry then ϕ^F may decline, partially compensating the husband for norm costs (see online Appendix Section D).

To solve program (1), a woman compares the value of the objective function if she doesn't work to the value if she pays the norms costs and chooses labor supply optimally. A woman will work in both sectors only if MGNREGS work is more remunerative and the MGNREGS hours constraint is binding. By increasing a woman's control over her earnings, D^2T raises her outside option and, therefore, μ. This increases her net transfer, ϕ^F and creates an income effect that will lower her willingness to work. However, a higher μ also *lowers* the weight she places on her husband's norms cost, γ^M, making work more attractive.

While these opposing effects make the predicted impact of D^2T on overall labor supply ambiguous, we can identify subgroups for whom the impact is clear. A first group is already-working women. As these women are neither personally, nor spousally, constrained the reduced weight on γ^M is irrelevant for their labor force participation decision. For this group, the only relevant force is the increase in ϕ^F, which will lead to a reduction in labor supply. By similar logic women who are personally, but not spousally, constrained will not enter the labor force: by definition a reduction of γ^M to 0 will not induce them to work, while the income effect makes working even less attractive.

The picture differs for spousally constrained women. By definition, a spousally constrained woman will work if $\gamma^M = 0$. For some parameter values, the reduced weight on γ^M will therefore induce work, despite the income effect. Proposition 1 formalizes this logic.

PROPOSITION 1: *An increase in a woman's outside option can increase FLFP only if prior to the change she is spousally constrained.*

PROOF:

See online Appendix Section D.

Proposition 1 tells us that if D^2T increases FLFP then norms costs to work exist. Further D^2T can increase both public and private labor supply, even though D^2T only affects MGNREGS wage payments.

This insight, that when female work imposes fixed norms costs, D^2T can increase labor supply, does not require household efficiency. In online Appendix Section D we outline an alternative model where the household is inefficient, in that a portion of a woman's wage is directly appropriated by her husband. It is reasonable to assume that D^2T reduces the "spousal tax" on MGNREGS earnings. As a consequence, we show that D^2T can lead to personally constrained women working more in both the public and private sectors. Intuitively, higher post-tax MGNREGS wages act as a "carrot" that may induce personally constrained women to pay fixed norms costs and enter the labor force; once they have incurred this cost they may decide to also undertake private sector work.

A final possibility is that D^2T directly reduces γ^F and/or γ^M; by reducing costs to FLFP, this could increase female work in both public and private sectors. We view this channel as unlikely in the short term, since the intervention did not target norms or communicate information that would shift perceived norms. In the medium to longer run, it is certainly

possible that women's choice to work reduces norm costs and this would amplify the female labor supply impacts highlighted above.

B. Empirical Predictions

We use this framework to interpret the observed labor supply impact of D^2T: if D^2T increases women's private sector work then fixed costs to their work exist. In our setting, we anticipate norms around women's work to be a primary cause of such costs.

Motivated by the observation that D^2T unambiguously increases female labor supply only among women who do not work absent the intervention (Proposition 1), we separate impacts by a woman's prior working status. In the empirical analysis we reference this group as "constrained," acknowledging multiple reasons for not working that include own norm costs, spousal norm costs, and low wages.

We also examine impacts on male labor supply. For an efficient household, an increase in M increases work among men whose wives are *not* spousally constrained (see online Appendix Section D). This is as these men incur a negative income effect while their wives work weakly less. Given this, we examine impacts separately for husbands of constrained and unconstrained women.

We also evaluate intervention impacts on two additional sets of outcomes. The first set includes proxies of women's bargaining power and empowerment. Here, we anticipate impacts for two distinct reasons. First, our model predicts that impacts of D^2T are mediated by increases in female bargaining power. Second, D^2T may alter the relative incomes of husbands and wives and, thereby, further influence empowerment outcomes.

The second set of outcomes includes own and perceived norms regarding women's work. We conjecture that direct exposure to a proscribed counter-stereotypical behavior, here, working women, may in the longer run reduce norm costs associated with women's work (Bertrand 2020). Own norms may liberalize among households with new female workers; and perceived norms (stigma from the broader community) may ease as people see more women in the labor market.

III. Data and Empirical Strategy

A. Data

Our evaluation uses multiple data sources.[9] First, a short screening questionnaire conducted prior to the intervention: This baseline identified the study sample by collecting data on presence of a married couple in the household, whether either spouse had ever worked for MGNREGS, and whether the wife had an individual bank account. Given time constraints, this survey did not record any other detail on women's financial inclusion, labor force participation, empowerment, or norms.

[9]To replicate our analyses, see Field et al. (2021).

Second, two follow-up surveys conducted roughly one and three years after account openings (between August and December 2015 and April and October 2017, referenced as short-run and long-run surveys respectively): We sampled 4,500 eligible women and their husbands from the baseline screening (stratified by GP) who could be matched to MGNREGS administrative data as of August 2015. Attrition did not differ by treatment arm: we interviewed 93 and 91 percent of sampled women during the first and second survey waves, respectively (online Appendix Table A1). Both female and male surveys included modules on bank account ownership, banking activities, and labor market outcomes. The female survey also collected data on proxies of female bargaining power and empowerment, including self-reported decision-making power, mobility, and experiences of gender-based violence, drawn from the Indian Demographic and Health Survey questions (see online Appendix Section E.3 for details).

Based on extensive qualitative work, we introduced norms-related survey modules in the long-run survey. We designed three modules to capture beliefs about whether women should work and gender-specific costs stemming from own and perceived community norms.

Third, administrative data from two sources: First, we have data from the MGNREGS program management information system (MIS) through mid-November 2017 (Government of India 2013–2017). The data tell us when an individual worked for MGNREGS, how much s/he was paid, and what account the wages were deposited into. We assume a woman was paid into her individual account if no other household member shares that account number.[10] Second, we have data from one of our two banking partners, which serves 81 percent of our sample. For accounts opened through this bank, we have data from date of account opening until April 30, 2018. This includes a record of every transaction posted to 1,603 female-owned accounts.

B. Empirical Approach and Balance Check

Our main analysis uses the following regression specification:

$$y_{igt} = \beta_0 + \beta_1 D^2 T_g + \beta_2 D_g^2 + \beta_3 T_g + \beta_4 C_g + \mu_s + \lambda_d + \eta_t + \mathbf{x}_{ig}'\boldsymbol{\delta} + \varepsilon_{igt}, \quad (2)$$

where y_{igt} is the outcome of interest for individual i in GP g at survey round t. Here, $D^2 T_g$ indicates that GP g was selected to receive accounts, direct deposit, and training; D_g^2 indicates a GP was selected for accounts and direct deposit; T_g indicates GPs selected for accounts and training; C_g indicates a control GP that received no treatment. All regressions control for strata and district fixed effects (μ_s, λ_d) and survey month × year fixed effects (η_t). We also control for the predetermined variables used to assess balance in online Appendix Table A2 ($\mathbf{x}_{ig}$). The error term (ε_{igt}) is clustered at the GP level.

To focus analysis on impacts of increasing financial control through D^{2}T holding (initial) account ownership constant, we set the omitted group to be GPs that received "accounts

[10]These data were scraped in 2016 and 2017 from the public MGNREGS website. The data structure capturing account numbers changed between the 2016 and 2017 scrapes. Online Appendix Section E.3 provides additional detail on how we infer individual account ownership from account number data in the two scrapes.

only." Coefficients on the control group dummy (β_4) are informative of the extent to which financial inclusion alone moves our outcomes of interest. Throughout, we report *p*-values to test differences between the other treatment groups and the control group.

Overall, individual, household, and GP-level characteristics specified in our pre-analysis plan are balanced across treatment arms (online Appendix Table A2).[11] The *p*-values from *F*-tests of whether the treatment group coefficients are jointly equal to zero (column 6) show imbalance on 5 out of 23 characteristics at the 10 percent level or less.

Motivated by our conceptual framework, we also estimate an augmented version of equation (2) which includes a proxy for whether a woman is unconstrained (i.e., she would work absent intervention), and interactions of this dummy variable with treatment dummies:[12]

$$\begin{aligned} y_{igt} = {} & \gamma_0 + \gamma_1 D^2 T_g + \gamma_2 D_g^2 \times Unconst_{ig} + \gamma_3 D_g^2 + \gamma_4 D_g^2 \times Unconst_{ig} \\ & + \gamma_5 T_g + \gamma_6 T_g \times Unconst_{ig} + \gamma_7 C_g + \gamma_8 C_g \times Unconst_{ig} \\ & + \gamma_9 Unconst_{ig} + \mu_s + \lambda_d + \eta_t + \mathbf{x}_{ig}'\boldsymbol{\delta} + \varepsilon_{igt}. \end{aligned} \tag{3}$$

We cannot observe constraint status directly, as it requires knowledge of counter-factual work behavior. We therefore proxy for a woman being unconstrained by the (only) female labor supply measure collected in the baseline survey: whether she ever worked for MGNREGS (recall Table 1). To the extent that this variable misclassifies women's true constraint status, we expect differences in treatment effects to be biased toward zero.

Our pre-analysis plan was general in that it specified main families of outcomes and laid out our intent to "evaluate the effect of the treatments—opening bank accounts, opening bank accounts and linking them to [MG]NREGS payments, and financial capability building—relative to the control and to one another," leaving us with some discretion in terms of how to aggregate outcomes within families and which statistical tests of the 10 suggested by the PAP to emphasize. Moreover, as discussed earlier, we did not prespecify our intent to study heterogeneity with respect to constraint status.

We address concerns related to ex post multiple testing in two ways. To address testing *within* families of outcomes, we aggregate variables into subfamilies (e.g., "public sector work," "private sector work," etc.), constructing standardized indices per Kling, Liebman, and Katz (2007). For each family, we average subindices to create a summary index. For indices and subindices measured in both waves, we report pooled analysis and then separate results by wave; these indices only include outcomes with comparable data in both waves.

Next, to address concerns related to multiple families of outcomes and multiple hypothesis tests, we report sharpened *q*-values that control for the expected share of rejections that are Type I errors, or false discovery rate (FDR) for our major hypotheses. We use the FDR approach outlined in Anderson (2008), based on the methodology in Benjamini, Krieger, and

[11] We lack data on two PAP-listed controls: GP median income and below poverty line ratio. In addition, we include district fixed effects: First, district governments facilitate access to MGNREGS work. Second, there is slight (district-level) imbalance between D^2T and accounts only. Our results are similar when we omit these controls.

[12] Online Appendix Tables A3 and A4 verify balance among the constrained and unconstrained subsamples.

Yekutieli (2006). This procedure converts p-values into q-values, which control the share of rejections that are Type I errors: specifically we expect 5 percent of rejections based on $q \leq 0.05$ to be Type I errors and so on.

Our primary "main effects" adjustment pools p-values from all ten hypothesis tests implied by the PAP across all summary indices and time periods. Thus, it pools 210 tests into a single family. A second "heterogeneous effects" adjustment includes p-values related to heterogeneous treatment effects, with the caveat that we cannot rely on the PAP to dictate which tests to include.[13] This adjustment pools 294 tests into one family.

Online Appendix Section C reports p-values with corresponding q-values for all hypotheses (including $\beta_1 = \beta_2$, $\beta_1 = \beta_3$, and $\beta_2 = \beta_3$) and outcomes. Figure 1 summarizes main results, graphing treatment effects relative to the Accounts Only mean for key families of outcomes. Whiskers graph 90 and 95 percent confidence intervals based on conventional standard errors; we report sharpened q-values above each bar. Online Appendix Figure A6 summarizes results by constraint status.

IV. Results

We first describe treatment take-up, and then evaluate intervention impacts on indices measuring labor supply, financial inclusion and agency, other domains of empowerment/agency, and norms. Online Appendix Section B presents impacts on index components.

A. Take-Up

Field administrative records show high take-up of our treatments (online Appendix Table A5). We opened accounts for 73 percent of eligible women, with no significant differences across treatment arms. Roughly three quarters of women in GPs selected for training were trained, and over half of women in direct deposit GPs were signed up for direct deposit.

Figure 2 is based on MGNREGS administrative data and shows the cumulative share of women receiving wage deposits into an individual account (panel A) and the value of those deposits (panel B), beginning at the start of our study period. (Note this figure does not tell us about *overall* rates of MGNREGS work, which we study in the next subsection). By the time of our long-run survey over 40 percent of women in D^{2}T GPs, but fewer than 10 percent of women in non-direct deposit GPs, had been paid MGNREGS wages through individual direct deposit. Online Appendix Figure A5 uses administrative data from one of our partner banks and shows very similar patterns, albeit with higher values of MGNREGS deposits.

The value of MGNREGS deposits is substantial: conditional on getting at least one direct deposit, the MGNREGS administrative data show the average woman in D^{2}T received roughly INR 4,295 ($66 at the 2017 exchange rate of INR 65 per US$) between the baseline and long-run survey. For comparison, annual wage earnings for women in this group was

[13]The "main effects" adjustment includes impacts relative to accounts only ($\beta_j = 0$, $j = 1, 2, 3, 4$) and relative to the control group $\beta_k = \beta_4$, $k = 1, 2, 3$), as well as $\beta_1 = \beta_2$, $\beta_1 = \beta_3$, and $\beta_2 = \beta_3$, For heterogeneous treatment effects tests include $\gamma_i = 0$, $i = 1, 2, 3, 4, 5, 6, 7, 8$, $\gamma_1 = \gamma_7$, $\gamma_1 + \gamma_2 = \gamma_7 + \gamma_8$, $\gamma_1 + \gamma_2 = 0$, $\gamma_3 + \gamma_4 = 0$, $\gamma_5 + \gamma_6 = 0$, $\gamma_7 + \gamma_8 = 0$.

INR 4,865 at the time of the long-run survey. Given the magnitude of these payments, it is plausible that the intervention shifted women's bargaining position in the household.

Figure 2 shows an uptick in individual MGNREGS payment receipt among non-direct deposit intervention arms starting in 2017. This likely reflects the combination of two major government initiatives. First, a few months after implementation of our interventions, the Indian government launched a nationwide, multi-year financial inclusion program, Pradhan Mantri Jan Dhan Yojana (PMJDY).[14] Second, prior to our long-run survey in 2017, the government conducted camps to provide individually linked direct deposit facilities (also known as Aadhar-linked accounts). This policy had an important effect: between intervention launch in 2014 and the short-run survey in 2015, the share of women in our study districts enrolled in individual direct deposit increased modestly, from 11 to 14 percent. However, between the short- and long-run survey, this number more than doubled to 32 percent (online Appendix Figure A1).[15] Thus, our long-run treatment effects reflect the *additional* effect of our interventions beyond these government efforts.

B. Labor Market Engagement

We now ask whether D^2T and associated gains in a woman's financial control altered her and her spouse's labor market engagement. Motivated by our theoretical framework, we pay special attention to impacts on public versus private sector work, and differential effects by a woman's baseline constrained status.

Table 2 assesses impacts on female labor supply. We average three standardized subindices to obtain the aggregate labor supply index (columns 1–3). The *general labor supply* subindex (column 4) includes labor supply measures that are not differentiated by work sector; the *public labor supply* subindex (column 5) only includes MGNREGS work measures; and the *private labor supply* subindex (column 6) only includes measures of private sector work.[16]

Pooling across waves, column 1 shows that D^2T increased female labor supply by 0.11 standard deviation units (significant at the 1 percent level, with a q-value of 0.018 per Figure 1). Columns 4–6 show remarkably similar impacts in the public and private sectors. Online Appendix Table B1 shows these effects reflect a 5 percentage point increase in having worked in the past month and in the past year, an 8 percentage point increase in the likelihood of MGNREGS work per administrative data, as well as a Rs 950 (24 percent) increase in annual private sector earnings. Motivated by our conceptual framework, we interpret the 0.13 standard deviation unit increase in private sector labor supply as

[14]PMJDY began in August 2014. By December 2017 over 300 million bank accounts (27 million in Madhya Pradesh) had been opened (https://data.gov.in/resources/stateut-wise-number-pmjdy-accounts-20122017-ministry-finance, accessed May 28, 2019). Under PMJDY, banks offered low-cost accounts with standard benefits including access to a debit card, accident and life insurance, and an overdraft facility.

[15]As illustrated by Figure 2 the share of women actually receiving direct wage payments is lower, since not all women enrolled work for MGNREGS.

[16]The *general labor supply subindex* includes an indicator for work in past month, earnings in past month, and total months worked over past year. The *public sector subindex* includes (i) MIS-based short-term (past month) and longer-term (past 12 months) work indicators and wages earned over those periods and (ii) survey-based reports of MGNREGS work for the same time periods. The *private sector subindex* includes a private sector work indicator, private sector earnings in past year and a dummy for whether her occupation/main status is a worker. Earnings proxy for intensive margin labor supply, given no substantive shift in market wages (see online Appendix Table A18)

demonstrating that gender norms or other fixed costs to female work constrain (some) women's ability to work.

Consistent with complementarity between direct deposit and training, no other treatment arm significantly impacted the aggregate labor supply index. However, direct deposit (without training) lowered the public sector labor supply subindex by 0.12 standard deviation units (column 5). This reduction is driven by administrative measures of MGNREGS work, not self-reported ones (online Appendix Table B1). One possibility is that the biometrically authenticated accounts opened for treatment women reduced local officials' ability to siphon funds by submitting false work claims in these women's names. If correct, the difference in public labor supply point estimates between D^2T versus direct deposit further highlights the importance of training in helping women effectively leverage the direct deposit facilities.

Columns 2 and 3 show an attenuation in D^2T treatment effects relative to the accounts only group over time (relative to the control group, we observe significant treatment effects for D^2T in both the short run and long run at $p = 0.006$ and $p = 0.043$, respectively, though long-run effects are not significant after FDR adjustments per online Appendix Table C1). Online Appendix Table A6 shows that D^2T impacts attenuate for both the public and private sector subindices, though point estimates on the general work subindex are stable. We further discuss this attenuation of effects below, in the context of heterogeneous treatment effects across constrained and unconstrained women.

Table 3 studies impacts on male labor supply. In the short run, D^2T increased the aggregate index by 0.09 standard deviation units, significant at the 10 percent level, with effects driven by public sector work (column 5). Online Appendix Table A7 shows that D^2T raises male public sector work in the short run, when women also work more in the public sector, and also in the long run, when women do not. As MGNREGS wages are below male private sector wages, this suggests an increased male willingness to accept work at lower wages.

In the presence of gender norm costs, we anticipate that D^2T impacts on female labor supply will be concentrated among constrained women. In Table 4 we examine heterogeneity in male and female labor supply responses by our baseline proxy of whether a woman is constrained.[17] Columns 1–4 consider women's labor supply, pooling short- and long-run survey waves. D^2T has a significantly larger impact on constrained women, increasing their labor supply by 0.21 standard deviation units (column 1, significant at the 1 percent level using standard inference and after FDR correction, see online Appendix Table C3). We reject equality of treatment effects for constrained and unconstrained women for the aggregate labor index, the general subindex, and the private sector subindex. Online Appendix Table A10 and online Appendix Table A11 break down the labor supply indices

[17] As previously discussed, while motivated by the theory our heterogeneity analysis by constrained status was not prespecified. In our pre-analysis plan, we proposed examining heterogeneity in outcomes by above and below median levels of (predicted) empowerment: since we did not collect empowerment data at baseline, we use time-invariant baseline characteristics to predict aggregate empowerment in the control group and use this model to create a predicted empowerment measure. Online Appendix Tables A8 and A9 report heterogeneity in labor supply effects using our prespecified measure. Overall, these results are similar to those obtained when splitting by constraint status: both in the short and long run, women with below median empowerment at baseline increase their labor supply. They also have a larger treatment effect on labor supply than women who are more empowered.

by short and long run for women and men, respectively. Online Appendix Table A10 shows that D^2T increased constrained women's labor supply in both the short run (0.23 standard deviation units) and long run (0.19 standard deviation units), with long-run effects concentrated in private sector work.[18]

In contrast, treatment effects for unconstrained women fade out over time. One potential reason relates to independent government efforts to transition women to MGNREGS direct deposit, discussed in Section IVC. Using administrative data, online Appendix Figure A3 graphs the share of workfare wages paid into individual accounts by quarter.[19] The 2017 government direct deposit enrollment camps are associated with increased direct deposit receipt rates, especially among unconstrained women, shortly before our long-run survey. Unconstrained women, who were more empowered to begin with (see Table 1), may have been better equipped to take advantage of the government's direct deposit campaign.

An alternative is that, in the longer run, the income effect generated by greater bargaining power among unconstrained women discouraged work. This would suggest that unconstrained women in D^2T should be more empowered than peers in accounts only. Indeed, columns 5–8 of Table 4 show that D^2T's effect on male public sector labor supply is qualitatively larger and only statistically significant among spouses of *unconstrained* women (column 7): this follows the prediction of the efficient model, where a negative income effect causes these spouses to work more. These patterns are, however, absent for the overall male labor supply index. We view this as consistent with our qualitative observations that most men work full time in the private sector and rely on MGNREGS, which pays less than the male private sector wage, for "work of last resort."[20]

C. Financial Inclusion and Agency

The domain of empowerment most directly tied to our interventions is financial. We study this in Table 5, which reports impacts on financial activity of women and their husbands, as well as female financial agency. Columns 1–3 report pooled, short-run, and long-run effects on an aggregate index measuring women's account use. This index is based on survey data and includes whether the woman reports owning a bank account at the time of the survey, whether she visited the account in the past six months, and self-reported savings in individual bank accounts.

The short-run control group coefficient (column 2) shows that providing individual bank accounts increases women's account use: control women score over 0.6 standard deviation units lower than accounts only women. However, this difference was roughly halved

[18]To check for misreporting of work type, we examine women's reports of payment method. In both survey waves, less than 2 percent of women reported receiving non-MGNREGS payments into a bank account. Our results are robust to recoding private sector work to zero if it is paid into a bank account. Our qualitative field work found that different recruitment and payment systems meant that villagers clearly distinguish MGNREGS work from other types of casual work.

[19]As we infer direct deposit status when women work, we cannot directly measure the share of all sample women who are signed up for direct deposit in a given quarter.

[20]Another test relates to time trends: policy catch up suggests an upward trend in FLFP, while an income effect a downward trend. However, other changes in the economic environment between the two survey waves, including the 2016 demonetization, makes a causal interpretation of time trends difficult.

between the short- and long-run survey, likely owing to the government's own efforts to bank women through PMJDY and sign them up for MGNREGS deposit.

In light of government policy "catch up," the persistent gains in female account use associated with D^2T are striking: compared to accounts only, women in D^2T score 0.14–0.15 standard deviation units higher on the account use index in both the short and long run; online Appendix Table B5 shows this includes a 6–9 percentage point increase in the probability of having gone to the bank in the past 6 months and an 8 percentage point increase in having an individual account in the long run (despite no initial differences). The estimates suggest a complementarity between direct deposit and training: sending money to a woman's accounts may have little effect if she lacks the capability to access the money on her own; similarly training may not do much if she has no impetus to transact.

To examine whether these reflect meaningful changes in women's financial agency, we consider women's banking knowledge and autonomy. These outcomes were measured for women in the long-run survey. D^2T led to a 0.16 standard deviation units increase in the bank kiosk knowledge index, significant at the 10 percent level (column 4). This index measures whether women have heard of the kiosk and what types of transactions they know they can conduct there. Moreover, column 5 shows that D^2T increased the female banking autonomy index by 0.12 standard deviation units. This index aggregates three types of outcomes: First, whether a woman visits the bank alone or without male supervision and is comfortable doing so. Second, whether she thinks women can visit the bank kiosk without a male relative's supervision. Third, whether she prefers having her wages paid into her own account and whether she prefers her wages are not sent to her husband. Online Appendix Table B5 shows that treatment effects are driven by women's comfort going to the bank alone (an 8 percentage point increase) and conducting transactions independently (a 10 percentage point increase). Online Appendix Table C6 shows that only impacts on the aggregate account use index remain significant after FDR adjustments, with *q*-values of 0.061, 0.201, and 0.041 in the pooled, short-run, and long-run specifications respectively.

Finally, columns 6–8 consider the male account use indices (standardized using the complementary accounts only control mean and standard deviation for women). The accounts only means for husbands show that their account engagement is significantly higher than their wives', especially in the short run. Unlike women, D^2T doesn't change male account use relative to accounts only.[21]

Online Appendix Table A12 breaks the results of the pooled aggregate indices and long-run bank kiosk knowledge and banking autonomy by whether a woman is constrained or unconstrained. As with labor supply we see qualitatively larger impacts for constrained women, though we generally cannot reject the null of equal impacts among the two groups of women.

[21]The large point estimates on some pooled and short-run male treatment effects are because women have limited, and substantially less variable, personal savings compared to men. If we were to instead construct the male index using male standard deviations, point estimates would be 5–10 times smaller.

D. Women's Empowerment

Beyond documented increases in labor supply and banking autonomy, there is scope for D^2T to alter other markers of female agency and empowerment via multiple channels.

First, as discussed in the conceptual framework, D^2T could directly increase a woman's bargaining position within the household by improving her outside option. This channel opens up the possibility of empowerment outcomes improving even when labor supply does not (e.g., among the unconstrained). Here, we anticipate impacts on indicators of both female well-being and of women's preference weights in household decisions. Second, treatment-induced shifts in banking and labor supply, and the increased access to, and control over, resources they bring, could trigger shifts in "downstream" measures of empowerment and agency. For example, women may be more likely to engage in other economic activity (like making household purchases) when they control their earnings. Female mobility could increase as women become more comfortable going to the bank and the job site. An increased capacity to do more could translate into a perceived ability to decide more. Finally, impacts on gender-based violence depend on the net effect of male backlash and female agency on domestic violence.[22]

In Table 6, we consider treatment effects on four domains related to womens economic agency: engagement in making purchases, mobility, self-reported decision-making, and freedom from gender-based violence. For each domain, we construct a subindex of empowerment based on female survey reports. The aggregate empowerment index is the average of the four subindices. Table 6 shows the pooled, short-run and long-run results for the overall summary index, as well as pooled results for its components. Online Appendix Table A13 shows short- and long-run impacts for subindices.

Overall, we find no significant impacts on the aggregate index (columns 1–3). This masks important heterogeneity, however: Figure 3 plots the distribution of the aggregate empowerment index among women in D^2T and accounts only GPs, in the full, the constrained, and the unconstrained samples. While, on average, unconstrained women report higher empowerment than constrained women, D^2T is associated with a significant rightward shift in the empowerment index distribution for constrained women relative to their peers in accounts only GPs; we reject equality of distributions for these two groups at the 1 percent level. Online Appendix Table A14 shows the average D2T treatment effect for constrained women is 0.075 standard deviation units, significant at the 5 percent level using conventional standard errors, with a *q*-value of 0.102 (online Appendix Table C13).

Column 4 of Table 6 shows that, relative to accounts only, D^2T does not increase the average woman's engagement with other markets as captured by the purchase subindex, though we do see a marginally significant difference of 0.06 standard deviation units relative to the control group.[23] This, again, masks heterogeneity by constraint status: online Appendix

[22]The empirical evidence on whether female labor force participation reduces gender-based violence (due to greater economic agency, as in Aizer 2010) or increases it (due to male backlash, as in Luke and Munshi 2011) is unclear.

[23]The purchase subindex captures purchases made by a woman, either at all or (in a separate set of dummy variables) with her own money in the past year.

Table A14 shows that constrained D^{2}T women score 0.18 standard deviation units higher than accounts only women in both the short and long run.

We see similar patterns when studying the mobility subindex, which aggregates dummy variables indicating whether a woman visited a series of common destinations like the local market and health center. Column 5 of Table 6 shows positive, but noisily estimated, gains for D^{2}T women relative to both accounts only and the control group. Looking across components, D^{2}T women are more likely to have visited the childcare center (8 percentage points) and her natal home (3.2 percentage points) (online Appendix Table B10). Again, we see larger, statistically significant effects for constrained women: online Appendix Table A15 shows D^{2}T women score 0.13 standard deviation units higher relative to accounts only.

Finally, we find no significant impacts, both overall and among the constrained, on the decision-making subindex, which aggregates two dummy variables indicating a woman reported having a say in whether she works and how her income and benefits payments are spent, and the freedom from gender-based violence subindex, which aggregates dummy variables measuring a woman's experience of physical, emotional, and sexual violence in the past year.

The measures of empowerment and bargaining power we observe and evaluate are not exhaustive: for instance, women may choose to parlay bargaining power gains into higher levels of transfers from their husbands, which we do not observe. Equally, a woman's reports of her *perceived* decision-making power may be subject to social desirability bias and influenced by prevalent norms. That said, the fact that the observed effects on broad measures of empowerment are also concentrated among the women who are constrained in terms of labor supply suggests that treatment impacts on female agency operate, in part, through women's greater engagement with the labor market (e.g., higher relative earnings and commuting/banking experiences), or, that D^{2}T did more to shift the outside options of constrained women.

E. Gender Norms

The D^{2}T intervention, which occurred in the context of a socially conservative society, increased female labor force participation. We now examine whether social norms around women's work themselves shifted as individuals gain experience with having a working woman in the household, and see more working women in the community.

Measuring Norms.——On norms, we designed three survey modules to capture men and women's beliefs about whether women should work, and the extent of norm costs stemming from own and perceived norms.[24] The first, on *personal beliefs and preferences,* asked individuals whether (i) women should be able to work outside the home, and (ii) they

[24]Our examination of both own norms and the perceptions of community norms is motivated by research in psychology and economics. The psychology literature emphasizes the distinction between own and perceived norms, and how misalignment between the two can lead to equilibria where individuals privately think behavior A is appropriate, but avoid behavior A because they believe others think A is inappropriate (see, e.g., Tankard and Paluck 2016). In Bernhardt et al. (2018), we find that male own and perceived norms, as well as the wife's belief about her husband's preferences, are more predictive of female work than the woman's own preferences.

wanted their sons to marry women who wish to work and their daughters to marry men who permitted work.

The second was a *vignettes module* which elicited individual attitudes toward working women and their husbands, holding household characteristics constant. The vignette featured two hypothetical families belonging to the respondent's caste and living in the respondent's village. The only difference across the two households was that in one case the wife worked for pay, while in the other case she stayed at home. We used pictures to make the households salient to the respondents. Respondents were asked which woman was the better wife, mother, and caretaker. To capture perceptions of community opinions, we asked which woman had more respect in the community. Then we asked which man was the better husband, provider, and who had more community respect (see online Appendix Section E.3 for more detail).

A final *community perceptions module* collected gender-specific measures of perceived norm costs. We asked respondents what fraction of individuals in the community would speak badly of a woman who worked outside the home, and what fraction of respondents would think a man was a bad provider if his wife worked for pay.

We construct five subindices relating to different aspects of norms and, then, aggregate them into two indices, the "own norms index" and the "perceived norms index." All index components are constructed so that higher values reflect fewer costs to female work. To facilitate cross-gender comparisons, we standardize index components relative to women in the accounts only group.

The own norms index includes three subindices. First, the "personal beliefs" subindex which captures whether the respondent thinks women should work and preferences for her child to live in a household with working women. Next, two subindices capture gender-specific norm costs: The "acceptance of working woman" subindex aggregates vignette judgments of whether the working woman is the better wife, the better mother, and the better caretaker. The "acceptance of working woman's husband" subindex aggregates vignette responses regarding which man is the better provider and husband.

The perceived norms index is the average of the two gender-specific perceived acceptance subindices, which include the vignette question on community respect and the "fraction of the community who judges" question.

Treatment Effects.——Table 7 presents results, separately for women and men (panels A and B, respectively). Among women, D^2T liberalized own norms by 0.10 standard deviation units (column 1), significant at the 1 percent level using conventional inference. Figure 1 shows this effect remains significant with $q = 0.046$ after FDR adjustments. Shifts in women's own attitudes indicate a more positive perception about the propriety of women's work, possibly linked to their own choice to work more: online Appendix Table B12 shows that this reflects an 8 percentage point increase in the likelihood a woman would prefer a daughter-in-law who works, a 7 percentage point increase in the likelihood of stating the working woman in the vignette is the better wife, and a 5 percentage point

increase in the likelihood of stating the working woman's spouse is the better husband. Online Appendix Figure A6 shows treatment effects, by gender, in constrained (panel A) and unconstrained (panel B) households. D^2T-induced liberalization of own norms (or, equivalently, the reduction in norm costs) are concentrated among constrained women, i.e., the group that responded to the treatment by increasing labor force attachment (online Appendix Table A16 and online Appendix Figure A6 present the own and perceived norms results for both genders.)

Women's perceived norms also liberalized (by 0.08 standard deviation units, column 5), driven by increases of 4.5 and 7 percentage points in the likelihood the respondent states the working woman and working woman's husband receives more respect in the vignettes (online Appendix Table B13). These effects are consistent with either women learning about more progressive beliefs held by others in the community and/or generalizing from their own liberalizing attitudes regarding women's work. The impact on women's perceived norms remains significant at the 10 percent level after FDR adjustments ($q = 0.092$).

Husbands' own norms were unaffected by the treatments (panel B, column 1). Perceived norms, in contrast, shift, though results are no longer significant at traditional levels after FDR adjustments ($q = 0.194$). Impacts are driven by changes in the husbands' acceptance subindex (column 7): both D^2T and training alone increased male views that husbands with working wives are accepted by others by 0.13 standard deviation units. Online Appendix Table B13 shows this effect is driven by a 0.044 unit (10 percent) increase in a husband's belief about the fraction of the community that does not think the husband of a working woman is a bad provider. These impacts are relevant as men perceive women's work involving more social stigma than women do: in accounts only GPs, the perceived acceptance of husbands index is 0.33 standard deviation units lower among men (relative to women), while the perceived acceptance of wives index is 0.14 standard deviation units lower.

What could cause a husband to update his perceived norms? First, his wife beginning to work may lead him to directly learn that he had overestimated the social sanctions associated with a woman working. Second, seeing more women in his village work as a result of the treatment could lead him to infer that the social costs of work are lower than expected. Although D^2T had a qualitatively larger impact on perceived acceptance of husbands among men in constrained households, the fact that we cannot reject equality of treatment effects between constrained and unconstrained households indicates that social learning may have contributed to a shift in men's perceived norms.

Our norms results raise interesting questions when viewed together with our labor supply results. Specifically, if norms did indeed shift, why did average labor supply effects attenuate? Here we identify two possibilities: online Appendix Table A10 shows that attenuation is entirely driven by unconstrained women. These women, who are less norms-constrained, may work less in the long run due to a bargaining-power-induced income effect. As discussed earlier, another potential driver of attenuation is policy catch-up: specifically, the government-led direct deposit campaign could have been enough to help unconstrained women in accounts only catch up to their D^2T peers. If norms change more slowly than

labor force participation, the current results may reflect that D^2T areas were exposed to greater FLFP for a longer period of time than our accounts only areas.

V. Discussion

As illustrated by Figure 1, D^2T had substantial positive impacts on women's work, including in the private sector, while women's husbands work more in the public sector. We observe larger, longer-lasting effects for constrained women, who are less likely to have worked absent intervention and whose husbands perceive higher social costs to having a wife who works. The persistence of these impacts are particularly striking in light of the Indian government's independent efforts to scale up both financial inclusion and MGNREGS direct deposit for women in the period between our short-run and long-run survey.

These changes translate into significant gains in financial activity and financial agency, though overall impacts on other domains of empowerment, captured by the aggregate empowerment index, are limited. We do, however, see broader empowerment gains for constrained women, especially in terms of mobility and economic engagement. Finally, treated women state more progressive attitudes about women in the labor force, while both genders report lower perceived social costs of female work.

Figure 1 and online Appendix Figure A6 illustrate a consistent story: D^2T has the largest impacts on outcomes across the causal chain, especially for constrained women, who theory predicts should be most affected. The figures also identify which results are robust to accounting for multiple inference. Figure 1 shows that effects on female labor supply and own norms remain significant at the 5 percent level, while effects on account activity and female perceived norms remain significant at the 10 percent level. Online Appendix Figure A6 reports core results splitting by constraint status. Here, we see that treatment effects on constrained women's labor supply, account use, and own norms remain significant at the 5 percent level, while the impact on empowerment is just short of significance at the 10 percent level. In light of these adjustments, we consider our inferences related to perceived norms for men as more speculative. Online Appendix Section C reports the full set of *q*-values for specifications estimating average and heterogeneous treatment effects.

Interpreting our results through the lens of the theory laid out in Section II indicates that D^2T helped women overcome fixed costs associated with work. Given the context, we consider the most likely reason for such fixed costs as related to the costs of violating gender-identity norms linked to women's work.

Below, we discuss several potential alternative channels through which our treatment may have operated and influenced women.

A. Alternative Explanations and Robustness Checks

While norms around women and work are our leading explanation for fixed costs to work, an alternative non-norms fixed cost relates to child care. If women were initially constrained by fixed childcare costs, then we may anticipate larger treatment effects for women with young (especially preschool age) children. Online Appendix Table A17 estimates effects by

whether or not a household had a child under the age of 8 at the time of the short-run survey. Treatment effects are apparent for both subgroups. While point estimates tend to suggest a slightly smaller response among women with young children, in general we cannot reject the null of no difference between the two groups.

Another potential fixed cost relates to learning about work opportunities in the private sector. In the private sector, labor recruiters typically visit households and offer them short-term work opportunities. However, recruiters target both genders, and since most men work, it is unlikely that women's MGNREGS participation increased their access to recruiters.

To rationalize increases across MGNREGS and private sector work, an alternative mechanism needs to impact the return to both forms of work. A natural possibility would be if women's increased participation in MGNREGS changed private sector wages. Online Appendix Table A18 shows that D^2T left these wages unaffected. While confidence intervals on some of these estimates are wide, general equilibrium effects of this sort are a priori unlikely, given that treated women comprised a small share of the population in most GPs (on average, our interventions targeted 28 women per GP, compared to an average female population of 1,625).

A related question is whether our treatment effects are biased by spillover effects, e.g., if MGNREGS funds were directed towards D^2T GPs at the expense of GPs in other treatment arms. Since our study sample is small relative to overall MGNREGS budgets (our study GPs accounted for 0.002 percent of total spending for the state of MP in fiscal year 2016–2017) we do not expect spillover effects to pose a substantive risk.

Another possibility is that D^2T impacted labor supply by easing savings constraints, as in Callen et al. (2019). However, our main comparison does not vary access to financial instruments as accounts only women also received bank accounts (online Appendix Tables A5 and B5). A related possibility is that D^2T reduced wage taxation in the private sector (e.g., if women had these wages deposited directly into their accounts, or made trips to the bank right after working). However, only 2.7 percent of private sector workers in D^2T report having these wages directly deposited, and we find no evidence that women deposit on their own: panels A and B of online Appendix Figure A4 show that non-MGNREGS deposit activity in accounts only is very similar to that in D^2T. It does not appear that treatment effects reflect a sudden surge in women's use of bank accounts for non-MGNREGS transactions.

B. Policy Implications

In recent decades, economic progress in India has translated into better-paying jobs and more attractive work opportunities, with wage growth in rural areas outstripping that in urban areas (Jacoby and Dasgupta 2018). Yet this growth has failed to draw Indian women into the labor market. We argue that social norms around appropriate gender roles play an important role in keeping women out of the labor force, but these norms can be overcome by interventions that increase women's financial control.

Strengthening women's control over MGNREGS wages through D^2T increased women's work both for the program and in the private sector. These changes run counter to the prediction of a basic model of efficient household decision-making, where an increase in bargaining power (precipitated by greater female control over workfare wages) would reduce female labor supply. Allowing for a norms channel rationalizes our main treatment effects and key heterogeneity in effects: treatment effects are largest among the subset of constrained women, who lacked MGNREGS work experience at baseline and had husbands who were significantly opposed to female work.

Our results have multiple policy implications. First, gender targeting can impact women's engagement with workfare programs and the labor market at large. Second, impacts can extend beyond economic fundamentals, reshaping the norms that govern female work. This creates scope for interventions like ours to create further welfare gains by altering the nature of preferences themselves. Third, our long-run results can help inform intervention scale-up discussions. Between our two survey waves, the Indian government began scaling up MGNREGS direct deposit to female-owned accounts across our study area. Different from our intervention, this scale-up did not involve either targeted outreach to eligible women or any systematic account training. It appears that these program features were relevant for the most marginalized women, and an important reason why we find persistent effects on constrained women's labor supply in the long run.

We conclude by highlighting some important open research questions relating to how norms are updated and perceived by community members. While our results make it clear that norms shift with behavior, we cannot say whose behavior (or beliefs) is most influential for changing the beliefs of others. Moreover, we are unable to speak to norms spillover to other members in the community. We see research that examines two-way interactions between social norms and economic activity in communities as a promising avenue for future work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Patrick Agte, Pooja Balasubramanian, Arielle Bernhardt, Vidushi Dhawan, Arti Dhar, Thomas Dreesen, Nikita Kohli, Savannah Noray, Sreelakshmi Papineni, K. C. Sampada, Elena Stacy, Natalie Theys, and Namita Tiwari for research assistance, Sitaram Mukherjee and Maithili Ramachandran for field management, and Eric Dodge and Akshat Goel for help with administrative data. We are grateful to the editor, three anonymous referees, Carlo Cusumano, Pascaline Dupas, Miho Hong, Helena Roy and seminar participants for helpful comments. We are appreciative of intervention implementation support from Dr. Aruna Sharma and Varad Pande. We thank IZA's Growth and Labour Markets in Low Income Countries Programme, the Bill and Melinda Gates Foundation, DFID, Harvard's Center for Population and Development Studies, the J-PAL Governance Initiative, and Innovations for Poverty Action's Financial Services for the Poor Initiative for funding. This project is registered in AEA RCT Registry (AEARCTR-0000115) and was approved by Harvard University and IFMR India IRBs. All views expressed are those of the authors.

REFERENCES

Aizer Anna. 2010. "The Gender Wage Gap and Domestic Violence." American Economic Review 100 (4): 1847–59. [PubMed: 25110354]

Aker Jenny C., Boumnijel Rachid, McClelland Amanda, and Tierney Niall. 2016. “Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger.” Economic Development and Cultural Change 65 (1): 1–37.

Akerlof George A., and Kranton Rachel E.. 2000. “Economics and Identity.” Quarterly Journal of Economics 115 (3): 715–53.

Alesina Alberto, Giuliano Paola, and Nunn Nathan. 2013. “On the Origins of Gender Roles: Women and the Plough.” Quarterly Journal of Economics 128 (2): 469–530.

Almås Ingvild, Armand Alex, Attanasio Orazio, and Carneiro Pedro. 2018. “Measuring and Changing Control: Women’s Empowerment and Targeted Transfers.” Economic Journal 128 (612): F609–39.

Anderson Michael L. 2008. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” Journal of the American Statistical Association 103 (484): 1481–95.

Atkin David. 2009. “Working for the Future: Female Factory Work and Child Health in Mexico.” http://www.econ.ucla.edu/atkin/Women%20in%20Manufacturing_old.pdf (accessed October 14, 2020).

Attanasio Orazio P., and Lechene Valérie. 2014. “Efficient Responses to Targeted Cash Transfers.” Journal of Political Economy 122 (1): 178–222.

Bachas Pierre, Gertler Paul, Higgins Sean, and Seira Enrique. 2020. “How Debit Cards Enable the Poor to Save More.” https://seankhiggins.com/assets/pdf/BachasGertlerHigginsSeira_DebitCards.pdf (accessed October 14, 2020).

Bandiera Oriana, Burgess Robin, Das Narayan, Gulesci Selim, Rasul Imran, and Sulaiman Munshi. 2017. “Labor Markets and Poverty in Village Economies.” Quarterly Journal of Economics 132 (2): 811–70.

Banerjee Abhijit, Duflo Esther, Goldberg Nathanael, Karlan Dean, Osei Robert, William Parienté Jeremy Shapiro, Thuysbaert Bram, and Udry Christopher. 2015. “A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries.” Science 348 (6236): 1260799. [PubMed: 25977558]

Banerjee Abhijit, Duflo Esther, Imbert Clement, Mathew Santhosh, and Pande Rohini. 2020. “E-Governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India.” American Economic Journal: Applied Economics 12 (4): 39–72.

Bedoya Guadalupe, Coville Aidan, Haushofer Johannes, Isaqzadeh Mohammad, and Shapiro Jeremy P.. 2019. “No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation.” NBER Working Paper 25981.

Benjamini Yoav, Krieger Abba M., and Yekutieli Daniel. 2006. “Adaptive Linear Step-Up Procedures That Control the False Discovery Rate.” Biometrika 93 (3): 491–507.

Bernhardt Arielle, Field Erica, Pande Rohini, Rigol Natalia, Schaner Simone, and Troyer-Moore Charity. 2018. “Male Social Status and Women’s Work.” AEA Papers and Proceedings 108: 363–67.

Bertrand Marianne. 2020. “Gender in the Twenty-First Century.” AEA Papers and Proceedings 110: 1–24.

Bertrand Marianne, Kamenica Emir, and Pan Jessica. 2015. “Gender Identity and Relative Income within Households.” Quarterly Journal of Economics 130 (2): 571–614.

Blundell Richard, Chiappori Pierre-André, and Meghir Costas. 2005. “Collective Labor Supply with Children.” Journal of Political Economy 113 (6): 1277–306.

Bobonis Gustavo J. 2009. “Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment.” Journal of Political Economy 117 (3): 453–503.

Bobonis Gustavo J. 2011. “The Impact of Conditional Cash Transfers on Marriage and Divorce.” Economic Development and Cultural Change 59 (2): 281–312. [PubMed: 21174881]

Boudet Ana Maria Muñoz, Petesch Patti, Turk Carolyn, and Thumala Angélica. 2012. “On Norms and Agency: Conversations about Gender Equality with Women and Men in 20 Countries.” World Bank Working Paper 74191.

Bursztyn Leonardo, González Alessandra L., and Yanagizawa-Drott David. 2020. “Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia.” American Economic Review 110 (10): 2997–3029.

Callen Michael, Suresh de Mel Craig McIntosh, and Woodruff Christopher. 2019. “What Are the Headwaters of Formal Savings? Experimental Evidence from Sri Lanka.” Review of Economic Studies 86 (6): 2491–529.

Chiappori Pierre-André. 1992. “Collective Labor Supply and Welfare.” Journal of Political Economy 100 (3): 437–67.

Demographic, and Health Surveys. 2006. “India: Standard DHS, 2005–06.” The DHS Program. https://dhsprogram.com/data/dataset/India_Standard-DHS_2006.cfm?flag=0px.

Fernández Raquel, and Fogli Alessandra. 2009. “Culture: An Empirical Investigation of Beliefs, *Work, and Fertility*.” American Economic Journal: Macroeconomics 1 (1): 146–77.

Fernández Raquel, Fogli Alessandra, and Olivetti Claudia. 2004. “Mothers and Sons: Preference Formation and Female Labor Force Dynamics.” Quarterly Journal of Economics 119 (4): 1249–99.

Field Erica, Pande Rohini, Rigol Natalia, Schaner Simone, and Moore Charity Troyer. 2016. “On Her Account: Can Strengthening Women’s Financial Control Boost Female Labor Supply?” https://economics.mit.edu/files/14005 (accessed October 14, 2020).

Field Erica, Pande Rohini, Rigol Natalia, Schaner Simone, and Charity Troyer Moore. 2020. “Empowering Women through Public Policy: Can Gender-Sensitive Policy Reform to a Major Public Works Program Encourage Female Work and Improve Women’s Wellbeing?” AEA RCTRegistry. October 15. 10.1257/rct.115-5.199999999999999.

Field Erica, Pande Rohini, Rigol Natalia, Schaner Simone, and Charity Troyer Moore. 2021. “Replication Data for: On Her Own Account: How Strengthening Women’s Financial Control Impacts Labor Supply and Gender Norms.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. 10.3886/E132501V1.

Fletcher Erin K., Pande Rohini, and Moore Charity T.. 2018. “Women and Work in India: Descriptive Evidence and a Review of Potential Policies.” India Policy Forum 15 (1): 149–98.

Goldin Claudia. 2006. “The Quiet Revolution That Transformed Women’s Employment, Education, and Family.” American Economic Review 96 (2): 1–21.

Government of India, Ministry of Rural Development. 2013–2017. “MIS Tables.” Management Information System. https://nrega.nic.in/.

Hasan Amer. 2010. “Time Allocation in Rural Households: The Indirect Effects of Conditional Cash Transfer Programs.” World Bank Group Policy Research Working Paper 5256.

Heath Rachel, and Mobarak A. Mushfiq. 2015. “Manufacturing Growth and the Lives of Bangladeshi Women.” Journal of Development Economics 115: 1–15.

ILO. 2020. “ILOSTAT Database.” ILOSTAT 2020. FLFP for women aged 15+. ILO Modelled Estimates https://www.ilo.org/shinyapps/bulkexplorer11/?lang=en&segment=indicator&id=EAP_2EAP_SEX_AGE_NB_A (accessed October 14, 2020).

India Census. 2011. “Sex Ratio in India.” https://censusindia.gov.in/2011census/training/district_pdf_search/district_pdf.aspx (accessed September 25, 2016).

Jacoby Hanan G., and Dasgupta Basab. 2018. “Changing Wage Structure in India in the Post-Reform Era: 1993–2011.” IZA Journal of Development and Migration 8 (1).

Jensen Robert. 2012. “Do Labor Market Opportunities Affect Young Women’s Work and Family Decisions? Experimental Evidence from India.” Quarterly Journal of Economics 127 (2): 753–92.

Kling Jeffrey R., Liebman Jeffrey B., and Katz Lawrence F.. 2007. “Experimental Analysis of Neighborhood Effects.” Econometrica 75 (1): 83–119.

Luke Nancy, and Munshi Kaivan. 2011. “Women as Agents of Change: Female Income and Mobility in India.” Journal of Development Economics 94 (1): 1–17. [PubMed: 24319310]

Lundberg Shelly, and Pollak Robert A.. 1993. “Separate Spheres Bargaining and the Marriage Market.” Journal of Political Economy 101 (6): 988–1010.

Muralidharan Karthik, Niehaus Paul, and Sukhtankar Sandip. 2016. “Building State Capacity: Evidence from Biometric Smartcards in India.” American Economic Review 106 (10): 2895–929.

Oh Suanna. 2020. “Does Identity Affect Labor Supply?” https://drive.google.com/file/d/1OPY49fjA5mVc68FE0Pccj-fboMhXzta-/view (accessed October 25, 2020).

Qian Nancy. 2008. “Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance.” Quarterly Journal of Economics 123 (3): 1251–85.

RBI. 2016. “Handbook of Statistics on Indian Economy 2015–16.” Reserve Bank of India. https://www.rbi.org.in/scripts/AnnualPublications.aspx?head=Handbook+of+Statistics+on+Indian+Economy (accessed September 16, 2020).

Skoufias Emmanuel, Unar Mishel, and Gonzalez de Cossio Teresa. 2013. “The Poverty Impacts of Cash and In-Kind Transfers: Experimental Evidence from Rural Mexico.” Journal of Development Effectiveness 5 (4): 401–29.

Subbarao Kalanidhi, Carlo del Ninno Colin Andrews, and Rodríguez-Alas Claudia. 2013. Public Works as a Safety Net: Design, Evidence, and Implementation. Washington, DC: World Bank Publications.

Tankard Margaret E., and Paluck Elizabeth Levy. 2016. “Norm Perception as a Vehicle for Social Change.” Social Issues and Policy Review 10 (1): 181–211.

UNWOMEN. 2012. “Minister of Rural Development Jairam Ramesh Commits to Making MGNREGA More Women-Friendly.” UNWomen News.

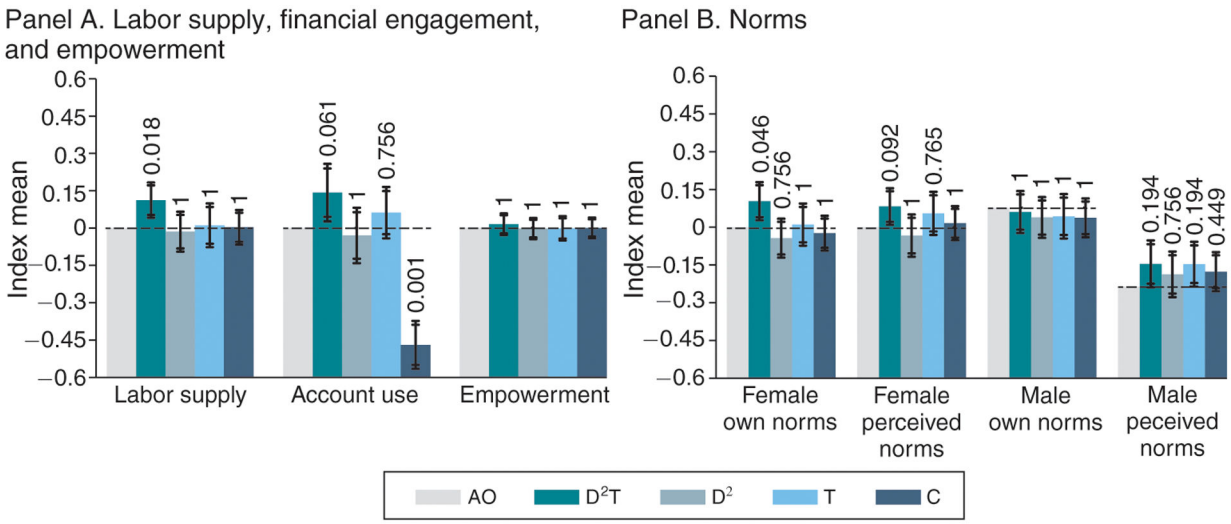

Figure 1.

Treatment Effects on Summary Indices

Notes: All figures used pooled short- and long-run data whenever possible. Light gray bar graphs the mean of each outcome for the accounts only group. The other bars are formed by adding treatment effects (per the specification in equation (2) in Section IIIB) to the accounts only mean. All of the controls included in the regression are listed in Table 2 notes. Missing values for controls are recoded as the mean and accounted for with the inclusion of indicator dummies for missing values. Whiskers display 90 and 95 percent confidence intervals based on robust standard errors clustered at the GP level. Sharpened two-stage *q*-values that control the false discovery rate are displayed above bars. Outcomes are standardized indices; details on index components are available in online Appendix Section E.3. The account use index is standardized relative to the entire female sample, because some index components are always equal to zero in the accounts only group. All other indices are standardized relative to the female mean in the accounts only group. Variables are standardized separately by survey wave; additional details of index construction are available in online Appendix Section E.2.

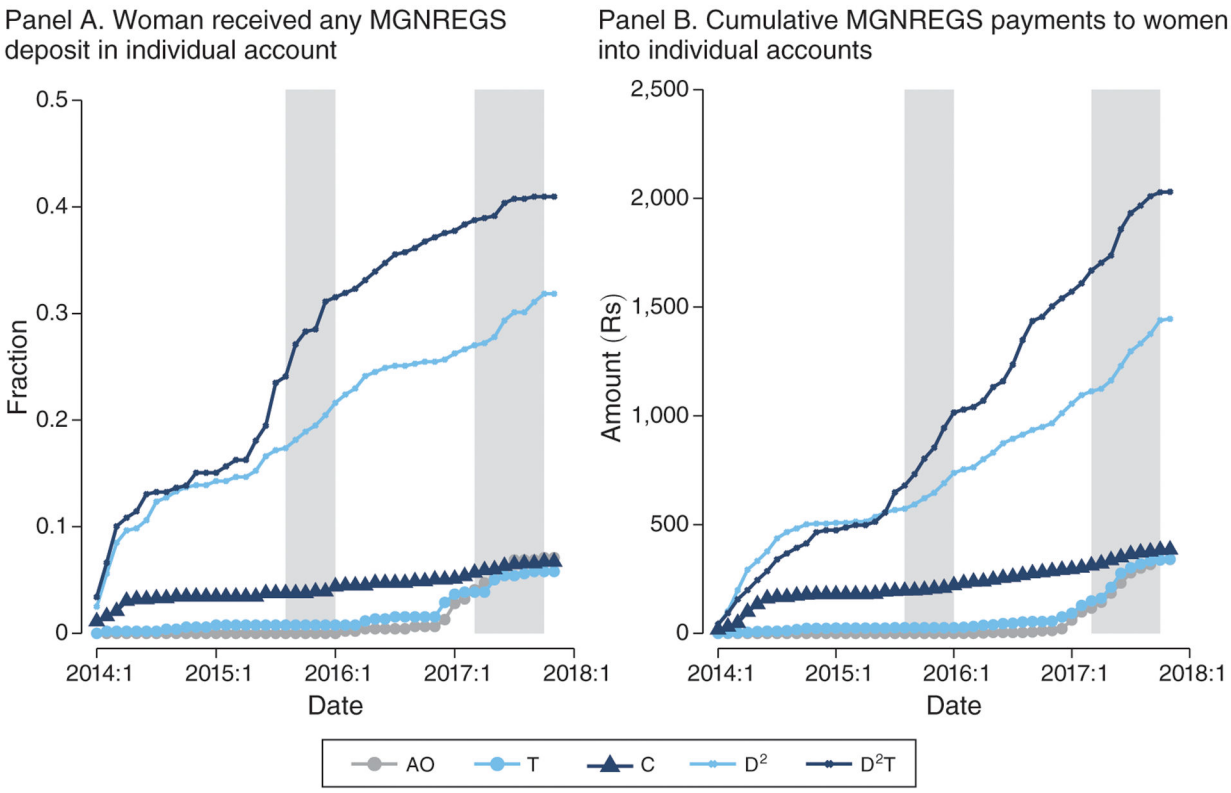

Figure 2.

MGNREGS Administrative Data: MGNREGS Deposits in Individual Accounts over Time

Notes: MGNREGS administrative data. Cumulative MGNREGS deposits are top-coded at the ninety-ninth percentile by month. Shaded bars demarcate the beginning and end of the short-run and long-run surveys. The exchange rate was approximately INR 64 per US$ in 2015 and INR 65 per US$ in 2017. Results exclude 104 women who could no longer be matched to the MGNREGS administrative data.

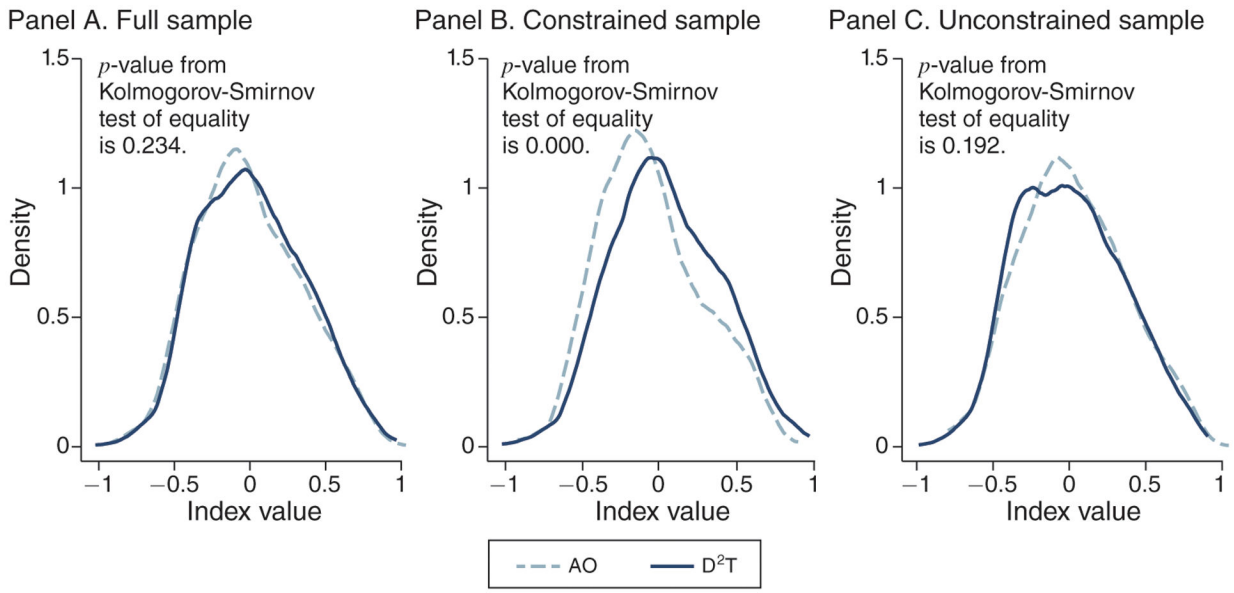

Figure 3.

Distribution of Empowerment by Baseline Constraint Status

Notes: Figure shows kernel density plot of aggregate empowerment index, pooling both the short- and long-run survey data. Outcomes are standardized indices; variables used to construct these indices are available in online Appendix Section E. The aggregate empowerment index is constructed with respect to accounts only females; additional details of index construction are found in online Appendix Section E.2. *Constrained* indicates the household female had not worked for MGNREGS prior to the baseline, while *Unconstrained* indicates the household female had worked for MGNREGS prior to the baseline.

Table 1—

Sample Summary Statistics, Control Group

	Overall mean (1)	Unconstrained mean (2)	Constrained difference (3)	Observations (4)
Panel A. Characteristics of women				
Age[a]	39.598	40.459	−2.626 (0.641)	1,738
Years education[a]	0.686	0.471	0.642 (0.153)	1,683
Age had first child (among women with kids at baseline)[a]	19.109	19.031	0.223 (0.178)	1,631
If worked for pay in past month	0.551	0.587	−0.106 (0.024)	1,654
Earnings last month	807.516	871.999	−191.804 (63.325)	1,630
Private labor subindex	0.003	0.070	−0.197 (0.049)	1,654
Public labor subindex	−0.104	−0.073	−0.076 (0.038)	1,654
Aggregate empowerment index	0.031	0.056	−0.074 (0.020)	1,644
Woman has say in taking employment	0.282	0.305	−0.070 (0.023)	1,651
Believes women can work	0.744	0.755	−0.035 (0.033)	1,650
Frac. community who will think poorly of working woman	0.387	0.378	0.027 (0.018)	1,648
Panel B. Characteristics of husbands				
Age[a]	44.238	44.962	−2.148 (0.792)	1,694
Years education[a]	3.879	3.266	1.732 (0.260)	1,688
If worked for pay in past month	0.682	0.693	−0.041 (0.027)	1,521
Earnings last month	1,473.888	1,438.257	71.473 (139.363)	1,503
Private labor subindex	0.589	0.564	0.066 (0.028)	1,651
Public labor subindex	0.127	0.173	−0.114 (0.058)	1,651
Believes women can work	0.656	0.668	−0.046 (0.033)	1,520
Frac. community who will think poorly of husband	0.564	0.542	0.065 (0.018)	1,519
Panel C. Household characteristics				
Scheduled caste/scheduled tribe[a]	0.441	0.458	−0.054 (0.054)	1,614
Household income per capita last month (male report)	980.419	1,059.449	−231.334 (57.005)	1,518
DHS work index[b]	0.014	0.028	−0.043 (0.018)	1,583

Notes: Robust standard errors clustered at the GP level in parentheses. Sample limited to control group.

[a]indicates that outcomes are from short-run survey; otherwise outcomes are from long-run survey.

Constrained indicates the household female had not worked for MGNREGS prior to the baseline, while *Unconstrained* indicates the household female had worked for MGNREGS prior to the baseline. The mean of the constrained indicator for this sample is 0.337. The first two columns show the means of the outcome variable (leftmost column) for the full control sample (column 1) and for the unconstrained sample (column 2). The third column shows the regression coefficient of the outcome variable on an indicator variable for being constrained.

[b]Indicates index was constructed using the Indian Demographic and Health Survey V (2005–2006) and merged onto our sample at the subcaste level, see online Appendix Section E.1 for more details.

Online Appendix Section E describes variable construction. Variables measured in INR topcoded at the ninety-ninth percentile. The exchange rate was approximately INR 64 per US$ in 2015 and INR 65 per US$ in 2017.

Table 2—

Impact of Treatments on Women's Labor Supply

	Aggregate labor supply index			General labor supply subindex	Public labor supply subindex	Private labor supply subindex
	Pooled (1)	Short-run (2)	Long-run (3)	Pooled (4)	Pooled (5)	Pooled (6)
β_1: Direct deposit and training (D^2T)	0.111 (0.036)	0.162 (0.040)	0.059 (0.049)	0.098 (0.053)	0.107 (0.064)	0.128 (0.048)
β_2: Direct deposit only (D^2)	−0.017 (0.040)	0.011 (0.042)	−0.048 (0.052)	0.016 (0.058)	−0.119 (0.058)	0.051 (0.046)
β_3: Training only (T)	0.013 (0.044)	0.021 (0.049)	0.002 (0.053)	0.004 (0.050)	−0.003 (0.085)	0.038 (0.042)
β_4: Control (C)	0.008 (0.035)	0.048 (0.038)	−0.024 (0.045)	−0.005 (0.046)	−0.002 (0.061)	0.030 (0.041)
Accounts only mean	−0.000	0.000	−0.000	0.000	0.000	0.000
Observations	8,297	4,179	4,118	8,297	8,297	8,297
p-values from F-tests						
$\beta_1 = \beta_4$	0.005	0.006	0.043	0.029	0.067	0.019
$\beta_2 = \beta_4$	0.517	0.358	0.595	0.678	0.029	0.560
$\beta_3 = \beta_4$	0.894	0.569	0.583	0.830	0.995	0.820
β_1: Short-run = long-run	0.059			0.893	0.167	0.014
β_2: Short-run = long-run	0.245			0.233	0.776	0.061
β_3: Short-run = long-run	0.722			0.816	0.761	0.923
β_4: Short-run = long-run	0.115			0.708	0.134	0.045

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column and as specified in equation (2) in Section IIIB. *Pooled* columns include outcomes from both the shortand long-run surveys. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. All regressions include strata, district, and wave-specific survey month fixed effects. GP (locality)-level controls include number of new MGNREGS work projects over the two years prior to baseline ratio of MGNREGS workers in two years prior to the baseline to 2011 census GP population proportion of in-sample MGNREGS workers in administrative data that did not self-report having worked for MGNREGS at baseline GP sex ratio, calculated from 2011 census data proportion of GP population that is scheduled caste proportion of GP population that is scheduled tribe sarpanch (elected GP leader) caste and sarpanch gender. Individual level controls include whether the respondent is part of a scheduled caste, scheduled tribe, age, household size, number of children over age three, whether the respondent worked for MGNREGS before baseline, age difference between husband and wife, education difference between husband and wife, and distance from nearest banking kiosk. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage q-values that correct the p-values of the aggregate indices in this table for the false discovery rate (FDR).

Table 3—

Impact of Treatments on Men's Labor Supply

	Aggregate labor supply index			General labor supply subindex	Public labor supply subindex	Private labor supply subindex
	Pooled (1)	Short-run (2)	Long-run (3)	Pooled (4)	Pooled (5)	Pooled (6)
β_1: Direct deposit and training (D^2T)	0.034 (0.040)	0.094 (0.051)	0.000 (0.045)	−0.053 (0.063)	0.172 (0.076)	−0.018 (0.038)
β_2: Direct deposit only (D^2)	−0.003 (0.047)	0.036 (0.062)	−0.017 (0.049)	0.074 (0.075)	−0.132 (0.069)	0.049 (0.048)
β_3: Training only (T)	0.031 (0.042)	0.065 (0.055)	−0.001 (0.049)	0.073 (0.079)	−0.041 (0.083)	0.060 (0.047)
β_4: Control (C)	−0.005 (0.039)	0.033 (0.055)	−0.032 (0.042)	−0.031 (0.063)	−0.015 (0.070)	0.030 (0.040)
Accounts only mean	0.509	0.553	0.466	0.647	0.174	0.707
Observations	8,065	3,957	4,108	8,065	8,065	8,065
p-values from F-tests						
$\beta_1 = \beta_4$	0.363	0.266	0.455	0.710	0.013	0.207
$\beta_2 = \beta_4$	0.962	0.953	0.715	0.135	0.078	0.643
$\beta_3 = \beta_4$	0.399	0.551	0.519	0.161	0.735	0.502
β_1: Short-run = long-run	0.088			0.064	0.998	0.026
β_2: Short-run = long-run	0.351			0.273	0.933	0.257
β_3: Short-run = long-run	0.258			0.451	0.566	0.286
β_4: Short-run = long-run	0.243			0.432	0.474	0.300

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column and as specified in equation (2) in Section IIIB. *Pooled* columns include outcomes from both the short- and long-run surveys. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. All regressions include strata, district, and wave-specific survey month fixed effects. Additional controls included are listed in Table 2 notes. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage q-values that correct the p-values of the aggregate indices in this table for the false discovery rate (FDR).

Table 4—

Heterogeneous Impact of Treatments on Labor Supply: Pooling Short-Run and Long-Run

	Women's labor supply				Men's labor supply			
		Aggregate index components				Aggregate index components		
	Aggregate labor supply index (1)	General labor supply subindex (2)	Public labor supply subindex (3)	Private labor supply subindex (4)	Aggregate labor supply index (5)	General labor supply subindex (6)	Public labor supply subindex (7)	Private labor supply subindex (8)
γ_1: Direct deposit and training (D^2T)	0.208 (0.044)	0.202 (0.058)	0.135 (0.090)	0.286 (0.070)	0.013 (0.058)	−0.035 (0.100)	0.124 (0.111)	−0.051 (0.058)
γ_2: Direct deposit and training (D^2T) × unconstrained	−0.147 (0.054)	−0.148 (0.066)	−0.044 (0.077)	−0.251 (0.079)	0.039 (0.067)	−0.017 (0.116)	0.076 (0.095)	0.059 (0.066)
γ_3: Direct deposit only (D^2)	0.033 (0.048)	0.053 (0.064)	−0.088 (0.069)	0.135 (0.068)	0.083 (0.068)	0.244 (0.124)	−0.089 (0.084)	0.096 (0.075)
γ_4: Direct deposit only (D^2) × unconstrained	−0.076 (0.052)	−0.050 (0.064)	−0.050 (0.063)	−0.129 (0.076)	−0.134 (0.074)	−0.258 (0.136)	−0.075 (0.076)	−0.070 (0.087)
γ_5: Training only (T)	0.076 (0.053)	0.092 (0.057)	−0.008 (0.086)	0.144 (0.070)	0.073 (0.053)	0.174 (0.109)	−0.067 (0.099)	0.112 (0.066)
γ_6: Training only (T) × unconstrained	−0.092 (0.054)	−0.128 (0.064)	0.014 (0.081)	−0.161 (0.080)	−0.059 (0.068)	−0.142 (0.128)	0.045 (0.088)	−0.079 (0.071)
γ_7: Control	0.098 (0.042)	0.084 (0.053)	0.055 (0.073)	0.155 (0.062)	0.009 (0.055)	0.034 (0.099)	−0.016 (0.088)	0.010 (0.060)
γ_8: Control × unconstrained	−0.138 (0.041)	−0.132 (0.052)	−0.090 (0.062)	−0.193 (0.066)	−0.023 (0.054)	−0.096 (0.107)	0.001 (0.067)	0.027 (0.058)
γ_9 : Unconstrained	0.224 (0.035)	0.248 (0.045)	0.095 (0.042)	0.328 (0.060)	0.078 (0.044)	0.146 (0.092)	0.035 (0.048)	0.053 (0.048)
p-values from *F*-tests								
$\gamma_1 = \gamma_7$	0.007	0.012	0.312	0.019	0.953	0.451	0.195	0.275
$\gamma_1 + \gamma_2 = \gamma_7 + \gamma_8$	0.026	0.086	0.044	0.130	0.116	0.866	0.003	0.495
$\gamma_1 + \gamma_2 = 0$	0.192	0.398	0.158	0.532	0.256	0.466	0.006	0.852
$\gamma_3 + \gamma_4 = 0$	0.371	0.969	0.035	0.915	0.336	0.856	0.037	0.657
$\gamma_5 + \gamma_6 = 0$	0.760	0.558	0.952	0.729	0.791	0.732	0.816	0.524
$\gamma_7 + \gamma_8 = 0$	0.302	0.361	0.596	0.402	0.747	0.363	0.829	0.380
Accounts only mean: constrained	−0.162	−0.183	−0.075	−0.228	0.517	0.654	0.159	0.737
Observations	8,297	8,297	8,297	8,297	8,065	8,065	8,065	8,065

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column. Regression is as specified in equation (2) in Section IIIB, with the addition of interactions of treatment dummies with an indicator that the woman was unconstrained, meaning she had worked for MGNREGS prior to the baseline. All columns include pooled outcomes from both the short- and long-run surveys. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. All regressions include strata, district, and wave-specific survey month fixed effects. Additional controls included are listed in Table 2 notes. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage *q*-values that correct the *p*-values of the aggregate indices in this table for the false discovery rate (FDR).

Table 5—

Impact of Treatments on Financial Inclusion and Agency

	Female reports					Male reports		
	Aggregate account use index			Bank kiosk knowledge index	Banking autonomy index	Aggregate account use index		
	Pooled (1)	Short-run (2)	Long-run (3)	Long-run (4)	Long-run (5)	Pooled (6)	Short-run (7)	Long-run (8)
β_1: Direct deposit and training (D^2T)	0.149 (0.059)	0.144 (0.074)	0.147 (0.054)	0.162 (0.091)	0.124 (0.058)	0.266 (0.210)	0.477 (0.384)	0.043 (0.088)
β_2: Direct deposit only (D^2)	−0.024 (0.056)	−0.058 (0.075)	−0.005 (0.053)	−0.066 (0.091)	−0.035 (0.057)	0.019 (0.192)	0.154 (0.352)	−0.043 (0.099)
β_3: Training only (T)	0.064 (0.052)	0.103 (0.065)	0.013 (0.052)	−0.075 (0.089)	0.018 (0.059)	0.321 (0.175)	0.514 (0.325)	0.049 (0.091)
β_4: Control (C)	−0.467 (0.049)	−0.644 (0.061)	−0.303 (0.045)	−0.515 (0.076)	−0.226 (0.050)	0.102 (0.160)	0.210 (0.298)	−0.103 (0.077)
Accounts only mean	−0.000	−0.000	−0.000	0.000	−0.000	1.110	1.682	0.560
Observations	8,297	4,179	4,118	4,118	4,118	8,065	3,957	4,108
p-values from F-tests								
$\beta_1 = \beta_4$	0.000	0.000	0.000	0.000	0.000	0.374	0.426	0.041
$\beta_2 = \beta_4$	0.000	0.000	0.000	0.000	0.000	0.616	0.846	0.448
$\beta_3 = \beta_4$	0.000	0.000	0.000	0.000	0.000	0.139	0.236	0.054
β_1: Short-run = long-run	0.964					0.236		
β_2: Short-run = long-run	0.444					0.564		
β_3: Short-run = long-run	0.131					0.154		
β_4: Short-run = long-run	0.000					0.285		

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column and as specified in equation (2) in Section IIIB. *Pooled* columns include outcomes from both the short- and long-run surveys. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. *Aggregate account use indices* in columns 1–3 and 6–8 are standardized relative to the entire female sample because some index components are always equal to zero in the accounts only group. All regressions include strata, district, and wave-specific survey month fixed effects. Additional controls included are listed in Table 2 notes. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage q-values that correct the p-values of the aggregate indices in this table for the false discovery rate (FDR).

Table 6—

Impact of Treatments on Other Empowerment Dimensions

	Aggregate empowerment index			Aggregate index components			
				Purchase subindex	Mobility subindex	Decision-making subindex	Freedom from gender-based violence subindex
	Pooled (1)	Short-run (2)	Long-run (3)	Pooled (4)	Pooled (5)	Pooled (6)	Pooled (7)
β_1: Direct deposit and training (D^2T)	0.015 (0.022)	0.004 (0.026)	0.023 (0.030)	0.025 (0.048)	0.053 (0.034)	−0.021 (0.041)	0.007 (0.030)
β_2: Direct deposit only (D^2)	−0.004 (0.021)	−0.013 (0.029)	0.000 (0.023)	−0.046 (0.044)	0.003 (0.036)	0.028 (0.042)	0.004 (0.031)
β_3: Training only (T)	0.001 (0.025)	−0.038 (0.029)	0.036 (0.031)	−0.024 (0.047)	0.038 (0.035)	0.003 (0.045)	−0.007 (0.034)
β_4: Control (C)	−0.001 (0.020)	−0.011 (0.026)	0.011 (0.024)	−0.033 (0.041)	0.009 (0.030)	−0.008 (0.040)	0.031 (0.029)
Accounts only mean	0.001	0.000	0.002	0.000	−0.000	−0.000	−0.000
Observations	8,276	4,179	4,097	8,276	8,297	8,297	8,297
p-values from F-tests							
$\beta_1 = \beta_4$	0.377	0.438	0.645	0.096	0.123	0.708	0.320
$\beta_2 = \beta_4$	0.865	0.935	0.521	0.705	0.792	0.249	0.229
$\beta_3 = \beta_4$	0.912	0.237	0.368	0.814	0.270	0.750	0.180
β_1: Short-run = long-run	0.560			0.258	0.979	0.141	0.090
β_2: Short-run = long-run	0.646			0.602	0.142	0.023	0.719
β_3: Short-run = long-run	0.025			0.092	0.534	0.045	0.061
β_4: Short-run = long-run	0.387			0.848	0.643	0.807	0.097

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column and as specified in equation (2) in Section IIIB. *Pooled* columns include outcomes from both the short- and long-run surveys. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. All regressions include strata, district, and wave-specific survey month fixed effects. Additional controls included are listed in Table 2 notes. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage q-values that correct the p-values of the aggregate indices in this table for the false discovery rate (FDR).

Table 7—

Impact of Treatments on Norms

		Aggregate index components				Aggregate index components	
	Aggregate own norms index (1)	Personal beliefs subindex (2)	Working women acceptance subindex (3)	Husband acceptance subindex (4)	Aggregate perceived norms index (5)	Perceived working women acceptance subindex (6)	Perceived husbands acceptance subindex (7)
Panel A. Female reports							
β_1 Direct deposit and training (D^2T)	0.102 (0.036)	0.114 (0.040)	0.087 (0.058)	0.106 (0.052)	0.078 (0.037)	0.078 (0.041)	0.078 (0.043)
β_2: Direct deposit only (D^2)	−0.032 (0.037)	0.015 (0.048)	−0.039 (0.058)	−0.071 (0.049)	−0.024 (0.040)	−0.042 (0.046)	−0.006 (0.043)
β_3: Training only (T)	0.016 (0.042)	−0.003 (0.041)	0.021 (0.054)	0.029 (0.064)	0.046 (0.040)	0.061 (0.041)	0.032 (0.047)
β_4: Control (C)	−0.015 (0.035)	−0.012 (0.036)	−0.032 (0.051)	−0.001 (0.049)	0.020 (0.037)	−0.008 (0.042)	0.048 (0.039)
Accounts only mean	−0.000	0.000	0.000	−0.000	−0.000	−0.000	0.000
Observations	4,118	4,118	4,118	4,118	4,116	4,116	4,116
p-values from F-tests							
$\beta_1 = \beta_4$	0.000	0.000	0.006	0.011	0.061	0.013	0.415
$\beta_2 = \beta_4$	0.566	0.502	0.874	0.084	0.212	0.391	0.174
$\beta_3 = \beta_4$	0.392	0.808	0.157	0.611	0.511	0.094	0.724
Panel B. Male reports							
β_1: Direct deposit and training (D^2T)	−0.023 (0.042)	−0.070 (0.056)	0.034 (0.051)	−0.032 (0.057)	0.086 (0.045)	0.044 (0.054)	0.127 (0.054)
β_2: Direct deposit only (D^2)	−0.023 (0.038)	−0.036 (0.061)	0.018 (0.049)	−0.050 (0.047)	0.062 (0.045)	0.042 (0.057)	0.082 (0.051)
β_3: Training only (T)	−0.033 (0.043)	−0.026 (0.063)	−0.005 (0.050)	−0.070 (0.057)	0.083 (0.044)	0.046 (0.052)	0.121 (0.054)
β_4: Control (C)	−0.033 (0.037)	−0.049 (0.054)	0.009 (0.045)	−0.059 (0.049)	0.068 (0.038)	0.054 (0.047)	0.082 (0.046)
Accounts only mean	0.077	0.180	0.001	0.049	−0.236	−0.138	−0.334
Observations	3,814	3,814	3,814	3,814	3,813	3,813	3,813
p-values from F-tests							
$\beta_1 = \beta_4$	0.769	0.670	0.502	0.579	0.649	0.826	0.300
$\beta_2 = \beta_4$	0.685	0.763	0.797	0.780	0.856	0.780	0.989
$\beta_3 = \beta_4$	0.991	0.634	0.681	0.792	0.620	0.831	0.298
β_1: Male = female	0.025	0.002	0.540	0.088	0.886	0.585	0.473
β_2: Male = female	0.853	0.426	0.450	0.745	0.128	0.188	0.182
β_3: Male = female	0.380	0.738	0.739	0.233	0.455	0.763	0.192
β_4: Male = female	0.710	0.520	0.561	0.370	0.252	0.187	0.536

Notes: Column headers list outcome variables of regressions including the treatment coefficients in the leftmost column and as specified in equation (2) in Section IIIB. All columns show long-run results. Outcomes are indices standardized relative to the female accounts only group separately by survey wave. Details of index construction are described in online Appendix Section E.2 and definitions of variables used to construct the indices are available in online Appendix Section E.3. Results on individual index components available in online Appendix Section B. All

regressions include strata, district, and wave-specific survey month fixed effects. Additional controls included are listed in Table 2 notes. Missing values for controls are recoded as the mean and regressions include an indicator dummy variable for variable-specific missing values. Robust standard errors clustered at the GP level in parentheses. See online Appendix Section C to view the sharpened two-stage *q*-values that correct the *p*-values of the aggregate indices in this table for the false discovery rate (FDR).