G²LM|LIC
Gender, Growth, and Labour Markets in Low Income Countries Programme

G²LM|LIC Working Paper No. 113 | August 2026

Income Growth and Women's Work:

Long-Run Evidence from India

Angelica Bozzi (Imperial College London)
Sam Asher (Imperial College London)
Rebecca Cai (MIT)

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid from the British people

G²LM|LIC

G²LM|LIC Working Paper No. 113 | August 2026

Income Growth and Women's Work:

Long-Run Evidence from India

Angelica Bozzi (Imperial College London)
Rebecca Cai (MIT)
Sam Asher (Imperial College London)

G²LM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-510

Email: glm-contact@iza.org

**G²LM|LIC Working Paper No. 113 | ** August 2026

ABSTRACT

Income Growth and Women's Work:*
Long-Run Evidence from India

This paper asks whether rising rural incomes are driving some of India's falling female labor force participation. Identification comes from India's vast network of gravity-driven irrigation canals, which generate exogenous variation in agricultural productivity by distributing water only to nearby lower-elevation villages. We find that canal treatment reduces the employment share in agriculture for both men and women, but labor reallocation differs by gender: men primarily shift toward non-farm employment while women primarily shift toward unpaid household production. These results show that economic growth may reinforce gender specialization when suitable nonfarm opportunities for women remain limited.

JEL Classification:
D13, J16, J22, O13, Q15

Keywords:
Female labor force participation; agricultural productivity; gendered labor reallocation; household production; rural development

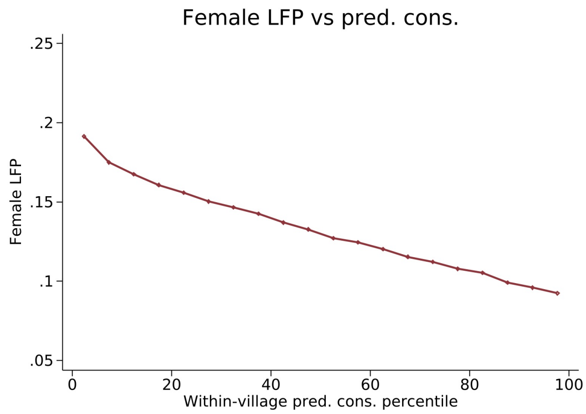
Corresponding author:
Angelica Bozzi (Imperial College London)
Email: a.bozzi22@imperial.ac.uk

* We appreciate feedback from seminar participants at STEG and UCPH workshop, Imperial, Harvard, MIT and at the 2026 World Inequality Conference. Financial assistance from the UK Foreign, Common-wealth & Development Office (FCDO) and the Institute of Labor Economics (IZA) within the G2LM|LIC Programme, STEG, the World Bank, and the European Union (ERC Consolidator Grant 101170766) is grate-fully acknowledged. . Views and opinions expressed are those of the authors only and do not necessarily reflect those of the World Bank, the European Union, the European Research Council, FCDO or IZA. None of these institutions can be held responsible for them.

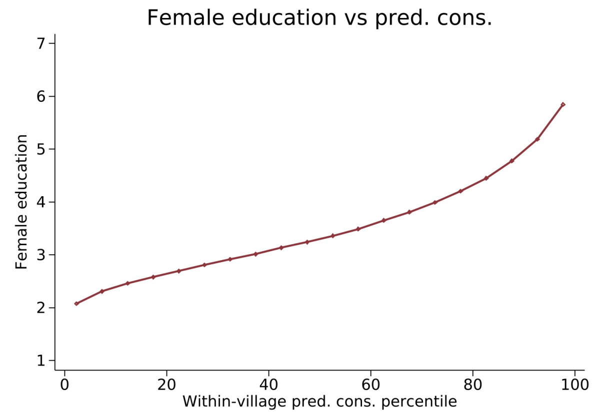
1 Introduction

In many developing countries, periods of rising income and schooling have coincided with stagnant or declining measured female labor force participation (World Bank, 2012). India is the most prominent recent example. Rural consumption increased, educational attainment rose, and agricultural productivity improved in many regions, yet female labor force participation remained low and in many datasets declined over time (Afridi *et al.*, 2018; Mehrotra & Parida, 2017). This combination has motivated a large literature on whether development eventually lifts women into market work or instead initially deepens specialization between market and home production (Mammen & Paxson, 2000; Goldin, 1994; Ngai & Petrongolo, 2017; Jayachandran, 2021). One possibility is that rising prosperity allows households to move toward a single-earner model, with women more likely than men to withdraw from market work because of social norms, barriers to entering new sectors, safety and mobility constraints, or a comparative advantage in household production. In the Indian context, this interpretation is consistent with widely documented norms that attach status to women's withdrawal from work outside the home when male earnings are sufficient, and with survey evidence that many adults disapprove of married women working outside the home when husbands earn well (Coffey *et al.*, 2018; Eswaran *et al.*, 2013; Jayachandran, 2021). But the central empirical challenge is that places with higher income are also places with different labor demand, schooling systems, and historical gender norms. As a result, it is difficult to know whether rising prosperity causally widens gender gaps or whether richer places simply look different for myriad other reasons.

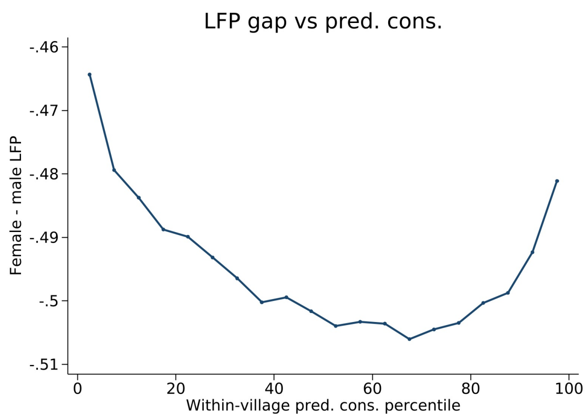
Figure 1 illustrates the motivating pattern using a 10% random sample of villages (N $\approx$ 55 million people) in the 2012 Socio-Economic and Caste Census. Women whose households are relatively wealthy within their village are monotonically less likely to participate in the labor force (Panel A) and more educated (Panel B), but increasingly less educated than their husbands and male family members (Panel D). Panel C plots the gender gap in labor force participation. The gap is narrower among the poorest and richest households and widest in the middle of the income distribution, echoing the much-discussed "U-shape" in female employment and development (Goldin, 1994). These correlations may reflect greater household specialization as resources rise. But they are also consistent with alternative explanations: richer households may differ in family structure, labor demand, or unobserved preferences. We make progress on this question by examining how a large and persistent income shock causally affects gender gaps in work and education.



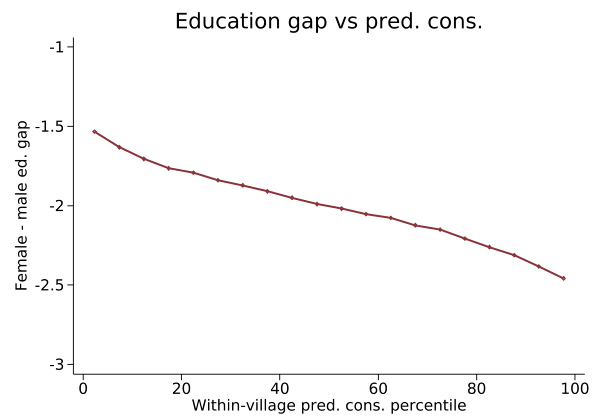
(a) Female labor force participation



(b) Female education



(c) Gender gap (F-M) in labor force participation



(d) Gender gap (F-M) in educational attainment

Figure 1: Descriptive patterns of female labor and education against household within-village wealth status

Notes: These descriptive graphs motivate the main empirical question. The source dataset is a 10% random sample of villages in the 2012 Socio-Economic and Caste Census (SECC) microdata. This graph represents a 10% random sample of the entire Indian rural population, instead of the balanced analysis sample used for the canals specification. In every subplot, each dot along the x-axis represents a "bin" of household consumption value (a common proxy for wealth in developing countries), normalized to percentile within village. In (a), the y-axis is female labor force participation rate: the share of women aged 18+ in the consumption percentile bin who work in agriculture, services, or manufacturing. In (b), the y-axis is female education attainment (years). In (c), the y-axis is the female-minus-male gender gap in LFP. In (d), the y-axis is the within-household female-minus-male gap in adult education attainment.

Our empirical setting uses the expansion of India's national irrigation canal network as a rural income shock. Canal irrigation was one of the largest agricultural infrastructure investments of the twentieth century in rural India: the network comprises more than 300,000 km of channels reaching over 130,000 villages (Asher *et al.*, 2022). Because the canals are gravity-fed, only villages below a nearby canal can receive water. Villages just uphill are nearby and often similar in geography and institutions, but cannot be irrigated. Following our companion paper, we exploit this engineering threshold in a regression discontinuity design comparing villages slightly above and below canals (Asher *et al.*, 2022). Most canals in our data were completed between 1920 and 1980, while outcomes are measured in 2011–2012. We therefore observe long-run re-allocations rather than short-run adjustment.

Previous work using the same design finds that canal exposure increased agricultural productivity, cultivated area, intensive dry-season cultivation, and household consumption (Asher *et al.*, 2022). These gains were not evenly distributed. Because canals raise the value of land, the largest persistent benefits accrued to landowners, while landless households received much smaller direct gains. Importantly, canal exposure is not found to have induced a broad sectoral shift toward manufacturing or services. The setting therefore captures a long-run income shock generated by a land-augmenting productivity shock without broader local structural transformation.

We find that canals widen the gender gap in labor force participation. Female labor force participation falls by 1.7 percentage points, about 5% of its control-group mean of 26%, compared with 0.5 percentage points, or less than 1% of the control mean, for men. Agricultural work declines similarly for women and men, by 1.7 and 1.5 percentage points. Their subsequent reallocation differs markedly: men move into paid nonfarm work (+1.0 percentage point) and studying (+0.3), while women move primarily into unpaid household production (+1.0) and studying (+0.2). To put this magnitude into context, the female decline accounts for more than one-tenth of the national decline in rural female labor force participation over the two decades preceding our outcome data.[1]

Canal exposure also raises education for both genders, but more for men: adult women's schooling rises by about 0.07 years, or 2.5% of the control mean, while men's rises by 0.19 years, or 3.8%. The results do not indicate that women are excluded from educational gains. Rather, schooling rises for both genders while still favoring men more strongly, so labor and education gaps widen together. These patterns are strongest among landown-

[1]National Sample Survey estimates indicate that labor force participation among working-age rural women fell 14 percentage points between 1987 and 2011 (Afridi *et al.*, 2018).

ers, who experienced the largest consumption gains. Female labor force participation falls more sharply, women's shift into household production is larger, and education gaps widen more than among the landless. This heterogeneity links gendered reallocation to the intensity of the productivity gains: the group that captures the largest gains from higher land productivity is also the group in which gender gaps widen the most.

The canal setting isolates a type of development shock that is common in agrarian economies but difficult to study causally: a persistent increase in the productivity of an immobile asset. When land becomes more productive, households may become richer without facing the changes in local industrial structure associated with urbanization or export growth. The shock raises agricultural income and loosens household resource constraints, but it does not automatically expand nearby paid jobs, especially for women. The gender response therefore reveals how income gains interact with pre-existing barriers in local labor markets rather than how women respond to a generalized nonfarm boom.

We claim the evidence is consistent with a long-run income shock interacting with gender-specific barriers to market work. Descriptive evidence from the India Human Development Survey shows that nonagricultural sectors are more male dominated and exhibit larger gender gaps in commuting than agriculture. These patterns are consistent with women facing greater barriers in moving from family farm work into paid nonfarm employment. Such barriers may operate through social norms, mobility constraints, safety concerns, or employer discrimination, though we cannot distinguish among these channels.

The paper contributes to three literatures. First, it provides long-run quasi-experimental evidence on how an agricultural income shock reshapes gender gaps in labor and education. Studies from India and China show that income growth can worsen women's outcomes, both absolutely and relative to men's, but focus primarily on human-capital investments rather than labor-market outcomes (Bhalotra *et al.*, 2019; Dreze & Sen, 1996; Almond *et al.*, 2019). Work connecting rising incomes to declining female labor force participation typically relies on cross-sectional decompositions, requiring strong assumptions that determinants such as education have stable effects over time (Afridi *et al.*, 2018; Mehrotra & Parida, 2017; Kapsos *et al.*, 2014). Our design instead exploits plausibly exogenous, persistent differences in agricultural productivity and identifies their long-run effects on labor allocation and education. It thereby strengthens the causal interpretation of the relationships documented in the cross-sectional literature.

Second, the paper contributes to research on structural transformation and the U-shaped relationship between development and female labor force participation (Goldin, 1994;

Ngai & Petrongolo, 2017). In the U-shape framework, female participation initially declines as agrarian incomes rise and employment shifts toward "brawny," male-dominated manufacturing sectors. It later recovers as women's home-based care work becomes marketized and female-dominated service employment expands. Recent research documents this upturn across developed economies (Olivetti & Petrongolo, 2014, 2016; Ngai & Petrongolo, 2017; Petrongolo & Ronchi, 2020). Evidence on the initial downturn is scarcer because it occurred centuries ago in wealthier countries and more recently, but with weaker documentation, in poorer ones (Ngai *et al.*, 2024). Existing methods are data-demanding, requiring granular household surveys reaching back to the agrarian phase of development. In developing countries, this type of data often does not exist. By identifying villages persistently assigned to different agricultural development paths and measuring outcomes decades later, canal exposure provides a rare setting for studying long-run allocations on the downturn side of the U-shape.

More broadly, the paper helps interpret measured female labor force participation by showing that movement between counted and uncounted work categories can be central to the response to development. Standard labor-force concepts count unpaid family work in market production as employment but exclude unpaid own-use household production (International Labour Office, 2013; Gaddis *et al.*, 2020). In rural India, many women perform agricultural work on family farms that is visible as labor force participation. A woman moving from a family farm into cooking, cleaning, childcare, or other domestic tasks therefore leaves measured employment even if her total work hours do not fall.

This distinction has been emphasized in statistical and time-use research, but is rarely brought to long-run causal evidence on development shocks (Burda *et al.*, 2013; Bridgman *et al.*, 2018; Dinkelman & Ngai, 2022). Our setting therefore clarifies one channel through which development can lower measured female labor force participation without implying that women stop working altogether.

Because canal irrigation generated sustained growth in agricultural productivity and income, the estimates capture long-run allocations after decades of household and labor-market adjustment. This differs from studies of temporary fluctuations in rainfall, commodity prices, or demand, which may not reveal how households allocate resources when higher productivity persists. By the time outcomes are observed, households, employers, and villages had decades to adjust schooling, fertility, migration, and labor allocation. The estimates therefore describe allocations following sustained agricultural productivity growth rather than short-run responses during canal construction or to a one-season harvest shock.

The rest of the paper proceeds as follows. Section 2 describes the canal setting, data, and empirical design. Section 3 presents the main results, suggestive mechanism evidence, and limits of our data. Section 4 concludes.

2 Setting, Data, and Empirical Design

2.1 Canals as a long-run land-augmenting income shock

Canals have been central to Indian agriculture for more than a century, but the modern national network expanded rapidly after independence in 1947. The systems in our data are designed for gravity-fed irrigation: water can flow only to lower-elevation terrain. This feature creates a sharp engineering threshold for causal inference. Villages below a nearby canal can receive water and experience a persistent improvement in irrigation access. Villages just above the same canal are nearby and share similar geography, but are not directly irrigated. Because treatment varies at a fine spatial scale, the design compares villages exposed to the same regional institutions and markets.

The canal is a land-augmenting productivity shock that generates a long-run income shock. It makes it possible to cultivate more land more intensively, especially in the dry season, and enables a second crop cycle in previously less productive places. Prior work finds substantial increases in cultivated area, irrigated acreage, and agricultural output below canals, alongside increases in household consumption and population density (Asher *et al.*, 2022).[2] At the same time, that work finds little evidence of broad local structural transition toward nonagricultural sectors. This combination is important for interpretation: canals raise agricultural prosperity persistently but do not mechanically create a large local market for paid work outside agriculture.

2.2 Outcome data and work categories

Our main outcomes come from the rural Socio-Economic and Caste Census (SECC), which records the economic status and primary activity of nearly the entire rural population. We study adults aged 20–64 for labor outcomes and adults aged 18 and above for education. The SECC is valuable because it includes free-text descriptions of adults' primary activity,

[2]Our companion paper Asher *et al.* (2022) provides contextual details, validates the RDD strategy, and reports results unrelated to gender.

which can be translated and grouped into standardized two-digit occupation categories.[3] We first classify whether working-age adults are working[4]. We then distinguish agriculture and paid nonfarm work, including manufacturing, services, and generic responses such as "laborer." Finally, we isolate two important activities outside the labor force: household production and attending higher education.

This classification is conceptually central. Standard labor statistics count unpaid family work contributing to market production as employment but exclude unpaid production for one's own household (International Labour Office, 2013; Gaddis *et al.*, 2020). In our setting, this separates unpaid family-farm work from unpaid household production. A woman who stops working on the family farm and spends more time on domestic tasks has moved from counted employment into uncounted work. The distinction matters for interpretation particularly in developing economies like India, where many women perform flexible, home-adjacent farm work. Throughout the paper, "household production" refers to nonmarket domestic activities, including housework, care work, and unpaid production for personal consumption, that standard measures exclude from labor force participation.

For education, we use years of education for adult women and men, as well as female-minus-male adult education gaps at the village and household levels. The household gap is useful because it captures within-household gender inequality directly rather than only the composition of adults in a village.

2.3 Empirical design

Our core regression discontinuity specification follows Asher *et al.* (2022). $BelowCanal_v$ indicates whether village v lies below the nearest canal and can receive gravity-fed irrigation; settlements just above the canal form the control group. Following Imbens & Lemieux (2008) and Gelman & Imbens (2019), we estimate:

$$\begin{aligned} y_{v,s} = {} & \beta_0 + \beta_1(BelowCanal_{v,s}) + \beta_2 RelElev_{v,s} + \\ & + \beta_3(RelElev_{v,s} \times BelowCanal_{v,s}) + \beta_4 X_{v,s} + v_s + \epsilon_{v,s} \end{aligned} \tag{1}$$

[3]Classification of local-language strings using India's National Classification of Occupations (2004) is in Online Appendix of Asher & Novosad (2020).

[4]Primary activity is agriculture or a non-farm job that is typically paid.

Here, $y_{v,s}$ is the outcome in settlement i within subdistrict s, and $RelElev_{v,s}$ is the running variable, measuring the elevation difference between the settlement and the canal. $X_{v,s}$ includes terrain ruggedness, rainfall, distances to the nearest river and coast, and potential rice and wheat yields based on geophysical fundamentals.[5] We include subdistrict fixed effects and cluster standard errors at the subdistrict level. The coefficient β_1 captures the difference in outcomes between villages that receive canal irrigation and those that do not.

Tables 2 and A2 of Asher *et al.* (2022) reports the corresponding balance evidence for the estimation sample. The balance tables examine both time-invariant geophysical characteristics and pre-treatment village characteristics measured in the 1951 Population Census. The latter include log population density, average household size, literacy rate, the male share of the population, and the share of households owning land. Across these covariates, the estimated discontinuities around the canal threshold are small and do not display a systematic pattern, which is consistent with villages just above and below the threshold being comparable prior to irrigation access.

Following Asher *et al.* (2022), we include settlements within 10 km and 50 vertical meters of the nearest canal and exclude villages exactly at canal elevation, where treatment status is uncertain. The education sample contains 57,808 villages. Due to inconsistencies in how some districts recorded primary activity in the SECC, some districts in our data report implausibly low labor force participation relative to the same-year National Sample Survey. We exclude these data-quality outliers, defined as districts with adult female participation below 2% or adult male participation below 40%, leaving 37,774 villages for the labor analysis.

Outcomes are observed roughly thirty years after the median canal's completion, so the estimates capture long-run allocations rather than short-run adjustment. By then, irrigation gains had been absorbed into land values, schooling, migration, and household labor allocation.

[5]Potential yields come from the UN Food and Agricultural Organization Global Agro-Ecological Zones data.

3 Results

3.1 Main labor results

Table 1 presents the main labor results. Panel A shows that canal exposure reduces overall labor force participation for both genders, but the effect is much larger for women. Female labor force participation falls by 1.66 percentage points ($p < 0.01$), compared with 0.51 percentage points for men ($p < 0.05$). Relative to the control-group means, the female decline is 5 percent while the male decline is less than 1 percent ($p < 0.05$). The female-minus-male labor gap becomes significantly more negative. Thus, the shock lowers employment but the decline is strongly gendered.

Table 1 Panel B shows why the aggregate result widens the gender gap. Agricultural work falls sharply for both women and men. The female agricultural employment rate falls by 1.7 percentage points and the male agricultural employment rate by 1.5 percentage points (Cols 1-2, $p < 0.01$). The initial response is therefore similar across gender. The divergence emerges once workers leave the farm. Despite working less in agriculture, women do not shift into non-farm paid jobs (Col 3). Instead, women's household production participation rises by 1.0 percentage point ($p < 0.10$), while studying rises by 0.2 percentage points ($p < 0.05$) (Cols 5 and 7). For men, by contrast, participation in paid non-agricultural work (services, manufacturing, and generic "laborer" jobs) rises by 1 percentage point ($p < 0.10$), and studying rises by 0.3 percentage points ($p < 0.05$) (Cols 4 and 8). The reallocation after agriculture is therefore strongly asymmetric.

In Appendix Table A1, we analyze the original "primary activity" strings for agricultural work, to show that the exit from agriculture is driven by working less on one's own family plot, rather than as hired field laborers. Importantly, own-plot work is not typically waged, but rather an unpaid family enterprise. This pattern matters for how we interpret measured female labor force participation. The women leaving agriculture are not moving one-for-one from paid to unpaid work. In absence of the canal treatment, these women would primarily work without pay on their own family's field, but instead they work in household production. In a standard labor-force measure, that shift lowers employment. In a time-allocation sense, it is a reallocation from counted work to uncounted work. The male response, in contrast, is more clearly a move across counted activities: away from agriculture and toward waged nonfarm work and higher education. The combination produces a widening gender gap even though the initial decline in agricultural work is broadly shared.

Table 1: Post-canal labor market outcomes and sectoral reallocation

Panel A. Overall labor force participation and gender gap

	(1) Female LFP	(2) Male LFP	(3) Gender Gap
Below Canal	-0.0166***	-0.0054**	-0.0112**
	(0.0057)	(0.0021)	(0.0053)
Geographic controls	✓	✓	✓
Subdistrict FE	✓	✓	✓
Mean in control group	0.305	0.867	-0.562
R-squared	0.525	0.387	0.541
Obs	37774	37774	37774

Panel B. Sectoral reallocation

	LFP				Not LFP			
	Agriculture		Nonfarm paid work		Housework		Studying	
	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male	(7) Female	(8) Male
Below Canal	-0.017***	-0.015**	0.000	0.010*	0.010*	0.000	0.002**	0.003**
	(0.005)	(0.006)	(0.004)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	0.153	0.464	0.189	0.410	0.552	0.008	0.029	0.069
R^2	0.37	0.40	0.40	0.38	0.54	0.10	0.41	0.39
Obs	37,774	37,774	37,774	37,774	37,774	37,774	37,774	37,774

Notes: Panel A reports the main regression discontinuity estimates for overall labor force participation using the SECC's variable for person's "primary activity". The sample is restricted to adults aged 20–64 and (to drop areas with low data quality) to districts where recorded female labor force participation is at least 2 percent and male labor force participation is at least 40 percent. In Panel A Cols 1-2, LFP equals 1 if a person works in agriculture, services, or manufacturing. In Panel A Col 3, the outcome is the (F - M) gender gap in LFP. Negative values imply women work at lower rates than men. Panel B reports the corresponding activity-level estimates. In Cols 1-2, the outcome is adult population share working in agriculture (own or others' plots). In Cols 3-4, "Nonfarm paid work" is employment in services or manufacturing. In Cols 5-6, "Household production" includes housework and unpaid positions in household enterprises. The terms "housewife" and "household chores" predominate this category. In Cols 7-8, "Studying" records the person is a student.

The reallocation margins also help distinguish the results from a purely technological decline in farm labor demand. The widening gender gap does not arise from different agricultural exits, but from what happens afterward. Men move toward paid nonfarm work and studying, whereas women move toward household production. This pattern is more consistent with differential opportunities or constraints outside agriculture than with a gender-specific reduction in agricultural labor demand.

3.2 Education results

Table 2 reports the main education effects. Canal exposure raises education for both women and men, but the increase is substantially larger for men. Adult women's years of education rise by 0.07 years ($p < 0.10$), while men's rise by 0.18 years ($p < 0.01$). Both education-gap measures become more negative by about 0.11 years ($p < 0.01$). Thus, the shock raises human capital investment but does not do so in a gender-neutral way.

Table 2: Exhibit 3. Education outcomes

	(1)	(2)	(3)	(4)
	Years of Education		Gaps in Education	
	Females	Males	Village	Household
Below Canal	0.0734*	0.1837***	-0.1109***	-0.1101***
	(0.0375)	(0.0457)	(0.0226)	(0.0227)
Geographic controls	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓
Mean in control group	2.683	4.663	-1.976	-1.995
R-squared	0.560	0.456	0.460	0.461
Obs	57808	57808	57806	57773

Notes: The table reports the main RDD estimates for years of education and gender gaps in education. Columns 1 and 2 report average years of education for women and men aged 18 and above. Columns 3 and 4 report the female-minus-male gap at the within-village and within-household levels.

This result complements rather than offsets the labor findings. A common view in the development literature is that rising prosperity increases schooling and may eventually expand women's market opportunities (Goldin, 1994; Ngai & Petrongolo, 2017; Jayachandran, 2021). Our estimates clearly support the first margin: education rises. The expansion of women's work, however, does not follow in our setting. Women are more educated in treated villages, yet women also move out of measured market work. Men obtain even larger schooling gains and are more likely to move toward paid nonfarm work. The same shock therefore widens labor gaps and education gaps together.

In our sample, families are predominantly patrilocal, meaning that women migrate from their birth villages to their husbands' homes after marriage. The village-level estimate of canal effects on women therefore bundles two distinct channels. The education of a married woman reflects the bridal choice of men raised in the village. The education of a never-married woman, who almost certainly lives with her parents, reflects her own parents' human capital investment decision. To separate these two channels, we re-estimate a person-level analog of Equation 1, fully interacted with an indicator for marital status (Appendix Section A4). Because marriage is nearly universal after age 30, we restrict our sample to adults aged 18-29. We find that both married and never-married women are significantly more educated in canal villages, and the effect of education does not vary significantly by marital status (Appendix Figure A5). The results are therefore consistent with a combination of more-educated women marrying into treated villages and educational gains among young women raised there, although the data do not allow us to quantify the relative importance of the two channels.

3.3 Heterogeneity by landownership

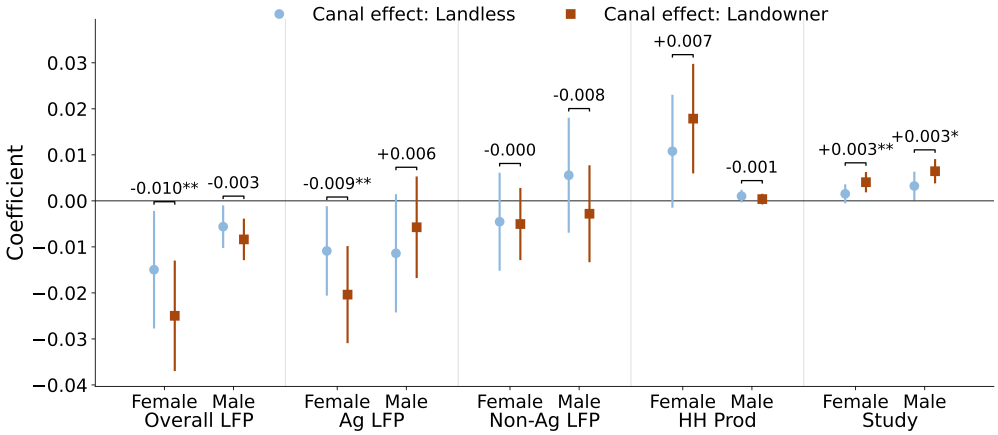
Figure 2 examines whether gendered reallocation is more pronounced among landowning households, which prior work shows received the largest consumption gains from canal access (Asher *et al.*, 2022). Stronger effects among landowners would strengthen the interpretation that income gains contribute to gendered reallocation.

To test whether canal effects differ by landownership, we re-estimate Equation 1, fully interacted with a landownership indicator:

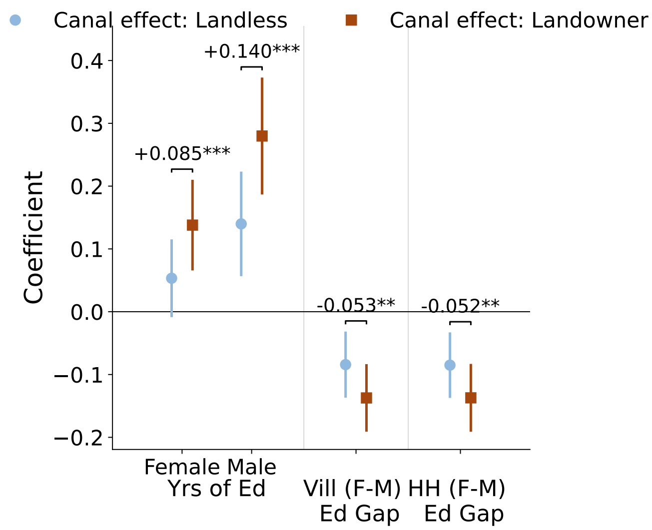
$$\begin{aligned} y_{v,s,LandOwn} = \beta_0 + \beta_1 BelowCanal_v + \beta_2 LandOwn + \beta_3(BelowCanal_v \times LandOwn) \\ + \beta_4 W_{v,s} + \beta_5(LandOwn \times W_{v,s}) + \upsilon_{s,LandOwn} + \epsilon_{v,s,LandOwn}. \end{aligned} \tag{2}$$

where $LandOwn_v$ indicates landownership, $W_{v,s}$ contains the running variable $RelElev_{v,s}$ and geographic controls. Subdistrict fixed effects are also interacted with landownership. Figure 2 plots canal impacts on landless ($\hat{\beta}_1$) and landowners ($\hat{\beta}_1 + \hat{\beta}_3$), and reports the difference $\hat{\beta}_3$, with stars indicating significance. Full regression results are in Appendix Tables A2 and A3.

Among landowners, women's labor force participation falls significantly more sharply than for the landless, and household production rises especially strongly. Men's response



(a) Labor market outcomes by landownership



(b) Education outcomes by landownership

Figure 2: Heterogeneity by landownership

Notes: The coefficient plot displays the canal treatment effect interacted with household's land ownership status. Blue circles plot the canal treatment effects and 95% CI on the baseline group (landless), corresponding to β_1 in Equation 2. Brackets above each cluster display β_3, the coefficient on $(BelowCanal_v \times LandOwn)$ capturing the additional change in the outcome in landowning, compared to landless, households. Stars on brackets indicate significance of β_3 (* $p < 0.10$, ** 0.05, *** 0.01). Red squares plot the treatment effect and 95% CI on landowners, where standard error of the treatment effect on landowners ($\hat{\beta}_1 + \hat{\beta}_3$) is the square root of:

$$\widehat{Var}(\hat{\beta}_1 + \hat{\beta}_3) = \widehat{Var}(\hat{\beta}_1) + \widehat{Var}(\hat{\beta}_3) + 2\,\widehat{Cov}(\hat{\beta}_1, \hat{\beta}_3)$$

.

is smaller and split between paid services and studying. Among the landless, the pattern is similar but attenuated: women leave agriculture for household production, but significantly less than landowning women. Gendered reallocation is therefore strongest among households receiving the largest consumption gains.

Education gains and the widening of education gaps are also significantly larger among landowners ($p < 0.05$). Together, these estimates are consistent with an income channel. Because canals may also alter landowners' demand for family labor, we interpret this evidence as supportive rather than conclusive.

3.4 Robustness checks

Our main findings are robust to several checks reported in the Appendix. First, RD scatter plots show that the results are not driven by functional-form choices (Appendix Section A1). Second, to address the SECC's reliance on self-reported primary activity, we replicate the result in India's 2011 Population Census where the same pattern emerges: FLFP falls more than MLFP, driven by agricultural labor decline (Appendix Table A6). Third, to distinguish from a direct labor-demand explanation, we consider if canals directly displace agricultural jobs primarily done by women. In fact, men–not women–overwhelmingly perform irrigation tasks replaced by canals (Appendix Figure A7). If canals reduced agricultural LFP through labor displacement, then men would leave agriculture at higher rates: the opposite of what we see.

Fourth, we test whether canals indirectly affected control villages through integrated labor markets, a key concern with our empirical strategy. If so, control villages near canals should differ from those farther away; in standard spillovers analysis we find no such differences (Appendix Tables A8–A10). Finally, we find no evidence that selective in-migration altered demographic composition in ways that explain our results. Treated and control villages are similar in age, marriage, and caste (Appendix Figure A8; Tables A11 and A12). Person-level estimates for 25 million adults remain nearly identical after controlling for these characteristics, making composition unlikely to explain the village-level results (Appendix Tables A13 and A14).

3.5 Interpretation and suggestive mechanism evidence

Taking stock, women and men both leave agricultural work following a long-run productivity increase, but men reallocate toward paid nonfarm work and higher education while women shift strongly toward unpaid household production. Education increases in canal villages, but more strongly for men. Both labor and education effects are concentrated among landowners. These results raise two questions: why education gains favor men, and why women and men reallocate differently after leaving agriculture.

In this section, we review the mechanism questions raised by these results. We argue the main results are most consistent with i) conservative gender norms against women working unless needed for subsistence, and ii) labor market structures (norms, skill requirements, mobility restrictions) that ration suitable jobs for rural women outside agriculture. Our interpretation of the joint labor and education response is that the canal treatment relaxes resource constraints without relaxing all market constraints. Higher capital income from greater land productivity allows households to invest more in education and to reduce reliance on family labor in agriculture. If the paid jobs outside agriculture are more available, more acceptable, or higher-paid for men than for women, then the same increase in resources can reinforce rather than reduce gender inequality in paid work. Families can invest in both daughters and sons while still favoring sons more strongly at the margin and reallocating women away from market work.

Why do parents increase daughters' education investments and find more educated brides for their sons if they don't anticipate greater labor market returns for adult women? While this contradicts standard labor economics theories on human capital investment, for many rural South Asian families, the dominant motive behind daughters' education is marriage market returns (i.e. matching to a higher-quality groom), not labor market returns (Buchmann *et al.*, 2023; Andrew & Adams, 2025).

Most centrally, why do men and women exit agriculture symmetrically, but men shift to paid jobs outside the home while women shift to unpaid work inside the home? We first exclude that the increase in women's household production is driven by changes in household composition. Canal-induced income gains could raise fertility, thereby increasing childcare needs, or alter marriage rates and women's domestic responsibilities. However, we find no detectable effects on either average fertility or marriage rates (Appendix Table A7).

We then consider two explanations for the gendered reallocation of work within households. First, gender norms may make it more acceptable for women, rather than men, to withdraw from visible market work when household resources rise, especially where women's non-employment outside the home conveys status and women remain responsible for most care and housework (Coffey *et al.*, 2018; Eswaran *et al.*, 2013; Jayachandran, 2021; Fletcher *et al.*, 2017; Deshpande & Singh, 2024). Such male-breadwinner norms are widely documented in India, although our village-level data do not directly measure them.

The second explanation is that the jobs available outside agriculture may be less accessible to women. The literature on women's work in rural India points to three related explanations. Non-farm gender wage gaps are larger outside of agriculture (Pattayat *et al.*, 2023). Women's non-farm labor demand may be rationed, due to occupational segregation from strong gender norms about "suitable work" for women (Chatterjee *et al.*, 2015; Chowdhury *et al.*, 2018). Third, agricultural labor–especially on one's own family farm–may be easier to combine with domestic responsibilities, whereas non-farm jobs require longer commutes and have less flexible schedules. Constraints to women's labor mobility are well-documented, especially when women have to leave their own village to access work opportunities (Cheema *et al.*, 2024). Combining these reasons, the supply of "suitable" non-farm jobs may be insufficient to meet women's demand, resulting in tens of millions of women who are willing to work but have difficulty matching (Fletcher *et al.*, 2017).

Our main analysis utilizes census data on the entire population and exploits variation across nearby villages. But in India, there is no available data on wages, commute distance, and job flexibility at this granularity. To provide suggestive evidence supporting the three claims above, we turn to the 2011-2012 India Human Development Survey (IHDS), which is population-representative at the district-level (N = 640). In Figure 3 we provide descriptive evidence consistent with non-farm jobs being rationed for women, using the IHDS rural data from the nine states that dominate the canals analysis sample.

Figure 3 plots men and women's respective share of total employment in agriculture, manufacturing, and services, using the same occupation classification as the main analysis. Agriculture is far more gender-symmetric, with a roughly 50-50 split of male and female workers. In contrast, women make up only 25.6% of construction/manufacturing workers, and 29% of service workers, despite evidence that women have a comparative advantage in services in developed and urban settings.

In the lower panel of Figure 3, we show that commuting patterns also differ by sector and gender. Agriculture has short and similar commutes for women and men. Importantly, these agricultural commute times are biased upwards, because the IHDS only asks commute times to paid agricultural wage labor (i.e. hired hands). Family farms, which underlie the canal-driven exit from agriculture, are presumably even closer to home. Construction, manufacturing, and service sectors require longer commutes for men (roughly 40 minutes), while women's commutes remain shorter (less than 30). We do not interpret this as a clean mobility estimate. But the pattern is consistent with evidence that many non-farm jobs are farther away, outside the village, or otherwise harder for women to access. In non-agricultural sectors, women only take up jobs closer to home, which are comparatively rationed.

This evidence does not distinguish whether the constraint is social acceptability, safety, employer demand, or commuting costs. It does show that the sectors into which workers would need to move after leaving agriculture are substantially more male dominated, and that existing male workers in those sectors travel farther than agricultural workers. The pattern is consistent with a setting in which nonfarm work expands the feasible choice set for men more than for women.

The IHDS evidence therefore speaks most directly to the second channel. It does not rule out norms, but it highlights a complementary constraint: if nonfarm jobs are farther from home, less flexible, more male dominated, or rationed through male networks, the same agricultural income shock can move men into paid work while moving women into household production. In such a setting, a household that can afford to release labor from farming may still find it much easier to reallocate men rather than women into paid work.

4 Discussion

Across the twentieth century, India constructed an extensive irrigation network comprising more than 300,000 kilometers of channels and reaching over 130,000 villages. Comparing villages that received gravity-fed irrigation with nearby villages just above canals reveals a distinct gendered adjustment. Agricultural participation falls by 1.7 percentage points for women and 1.5 percentage points for men, but their subsequent allocations differ: men move toward paid nonfarm work and studying, while women move primarily toward household production. These differences are robust to checks for spillover and differential migration, and remain similar in person-level specifications.

Education rises alongside this reallocation, but more for men. Because residence is predominantly patrilocal, the education of women observed in treated villages combines two margins: the education of young women raised there and that of women marrying into the village. Among adults aged 18–29, education rises for both married and never-married

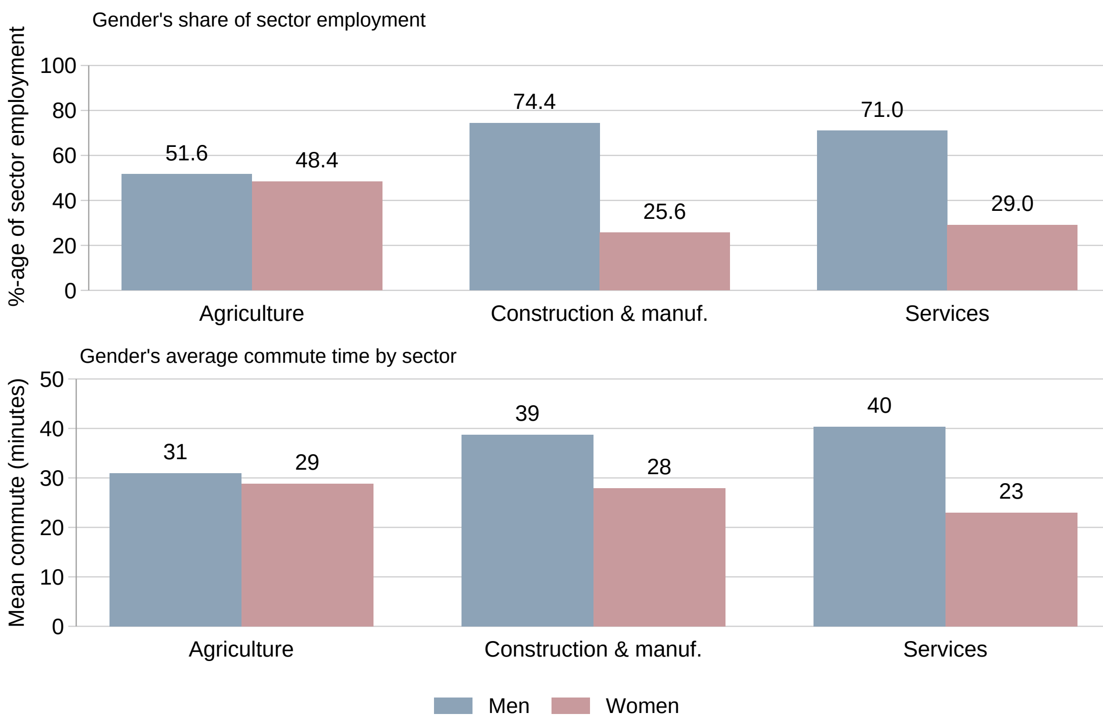


Figure 3: Descriptive mechanism evidence from sectoral segregation and commute times

Notes: This descriptive exhibit uses data from the population-representative rural 2012 India Human Development Survey (IHDS) to compare men and women's employment share and average commute time by sector. We restrict the IHDS sample to the 9 states that dominate our canals analysis. In the upper graph, the y-axis plots the gender's share of total employment in this sector (male and female bars total to 100%). In the lower graph, the y-axis plots the mean commute time (from home to work) by gender and sector for existing workers. The figure is intended as mechanism context rather than a causal test: it illustrates that compared to agriculture, non-agricultural paid work is more male dominated, and has larger gender gaps in commute time. Importantly, "Agriculture" in these graphs is wage/salaried labor, because the IHDS does not ask commute distance to household farm.

women, with no significant difference between the estimates. The results are therefore consistent with both educational gains among women raised in canal villages and more-educated women marrying into them. Labor and education effects are strongest among landowners, who received the largest consumption gains, supporting the income-channel interpretation. Moreover, men predominantly perform the irrigation tasks most directly affected by canals, making technological displacement of female-intensive farm tasks unlikely to be the main explanation.

The findings provide one possible margin underlying the downturn of the U-shaped relationship between development and female labor force participation (Goldin, 1994; Ngai & Petrongolo, 2017; Dinkelman & Ngai, 2022). Higher agricultural productivity reduces reliance on family farm labor, but the activities available afterward are not equally accessible to women and men. Unpaid work on a family farm is counted as employment, whereas household production generally is not. Women's movement between these categories therefore lowers measured labor force participation, while men are more likely to transition into paid work. The evidence is consistent with income gains interacting with gender norms and barriers to women's access to nonfarm employment. The IHDS evidence shows that nonagricultural sectors are substantially more male dominated and that male workers in these sectors travel farther than agricultural workers. These patterns do not isolate a single mechanism, but they illustrate why leaving agriculture may expand men's paid employment opportunities more readily than women's. Because the outcomes are observed approximately thirty years after the median canal's completion, the estimates describe allocations long after construction; they do not trace the adjustment process over time.

The findings should not be interpreted as implying that agricultural productivity growth is undesirable. Canal irrigation raised productivity and consumption, and education increased for both women and men. At the same time, barriers to women's access to paid nonfarm work mean that these aggregate gains coexist with a gendered reallocation of labor whose welfare consequences are ambiguous. Greater concentration of women's activity within the home may limit mobility, social interaction, access to higher-paid work, and economic autonomy. Recent descriptive evidence from India documents high levels of female seclusion and shows that women are disproportionately concentrated in lower-paid jobs compatible with housework (Andrew & Smurra, 2024). At the same time, leaving agricultural work may reduce physically demanding labor and permit greater investment into childcare and other forms of household production. Because our data do not measure total working hours, women's wellbeing or autonomy, or child outcomes such as infant mortality or nutrition, we cannot determine whether this reallocation increases or reduces welfare. The narrower implication is that productivity gains need not generate convergence in economic opportunities when women face greater barriers to the activities that expand outside agriculture. Complementary improvements in mobility, childcare, flexible work arrangements, and access to nearby paid employment may therefore shape whether women can translate rising household resources into market opportunities.

5 Conclusion

This paper studies the long-run gender consequences of a land-augmenting productivity shock operating through a persistent income channel. Using the elevation threshold around India's gravity-fed irrigation canals, we show that canal exposure leads women and men to reduce agricultural work. Their subsequent reallocation differs sharply: men move toward paid nonfarm work and studying, while women move primarily toward household production. Education rises for both genders but more for men, widening education gaps as well. These effects are strongest among landowners, who experienced the largest gains from higher land productivity.

These findings help interpret rural development episodes in which female labor force participation declines while schooling and consumption improve. In such settings, lower female employment need not be inconsistent with gains in living standards and education. It may instead reflect women moving from unpaid family agriculture, which is counted as employment, into unpaid household production, while men transition into paid work outside agriculture. The decline in measured participation is real, but reflects a particular reorganization of work within households rather than an absence of effort. Distinguishing among these margins is essential for determining whether low productivity, low schooling, weak labor demand, or barriers to women's access to paid work constrain market opportunities.

The relationship between development and gender equality therefore depends not only on whether incomes rise, but also on which forms of work expand and which remain closed. Where rural prosperity reduces the need for family farm labor without creating comparable market opportunities for women, gender specialization can deepen rather than diminish. Income growth alone need not guarantee convergence in gender outcomes.

References

AFRIDI, FARZANA, DINKELMAN, TARYN, & MAHAJAN, KANIKA. 2018. Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades. *Journal of Population Economics*, **31**(July), 783–818.

ALMOND, DOUGLAS, LI, HONGBIN, & ZHANG, SHUANG. 2019. Land reform and sex selection in China. *Journal of Political Economy*, **127**(2), 560–585.

ANDREW, ALISON, & ADAMS, ABI. 2025. Revealed beliefs and the marriage market return to education. *The Quarterly Journal of Economics*, **140**(3), 2107–2162.

ANDREW, ALISON, & SMURRA, ANDREA. 2024. *Seclusion and Women's Time: Descriptive Evidence from India*. IFS Working Paper 24/39. Institute for Fiscal Studies.

ASHER, SAM, & NOVOSAD, PAUL. 2020. Rural roads and local economic development. *American economic review*, **110**(3), 797–823. ISBN: 0002-8282 Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

ASHER, SAM, CAMPION, ALISON, GOLLIN, DOUGLAS, & NOVOSAD, PAUL. 2022. *The Long-Run Development Impacts of Agricultural Productivity Gains: Evidence from Irrigation Canals in India*. CEPR Discussion Paper 17414. Centre for Economic Policy Research. Revise and resubmit, *Review of Economic Studies*.

BHALOTRA, SONIA, CHAKRAVARTY, ABHISHEK, MOOKHERJEE, DILIP, & PINO, FRANCISCO J. 2019. Property rights and gender bias: Evidence from land reform in West Bengal. *American Economic Journal: Applied Economics*, **11**(2), 205–237.

BRIDGMAN, BENJAMIN, DUERNECKER, GEORG, & HERRENDORF, BERTHOLD. 2018. Structural Transformation, Marketization, and Household Production Around the World. *Journal of Development Economics*, **133**, 102–126.

BUCHMANN, NINA, FIELD, ERICA, & GLENNERSTER, RACHEL. 2023. *The marriage market value of bride age and education: A vignette approach to decomposing marriage transfers in rural bangladesh*.

BURDA, MICHAEL C., HAMERMESH, DANIEL S., & WEIL, PHILIPPE. 2013. Total Work and Gender: Facts and Possible Explanations. *Journal of Population Economics*, **26**(1), 239–261.

CHATTERJEE, URMILA, MURGAI, RINKU, & RAMA, MARTIN. 2015. *Job Opportunities along the Rural-Urban Gradation and Female Labor Force Participation in India*. Policy Research Working Paper 7412. World Bank, Washington, DC.

CHEEMA, ALI, KHWAJA, ASIM I, NASEER, FAROOQ, & SHAPIRO, JACOB N. 2024. Glass walls: Experimental evidence on access constraints faced by women. *Journal of Political Economy*, **Forthcoming**.

CHOWDHURY, AFRA RAHMAN, AREIAS, ANA CAROLINA, IMAIZUMI, SAORI, NOMURA, SHINSAKU, & YAMAUCHI, FUTOSHI. 2018. *Reflections of Employers' Gender Preferences in Job Ads in India: An Analysis of Online Job Portal Data*. Policy Research Working Paper 8379. World Bank, Washington, DC.

COFFEY, DIANE, HATHI, PAYAL, KHURANA, NIDHI, & THORAT, AMIT. 2018. Explicit Prejudice: Evidence from a New Survey. *Economic and Political Weekly*, **53**(1), 46–54.

DESHPANDE, ASHWINI, & SINGH, JITENDRA. 2024 (Sept.). *The Demand-Side Story: Structural Change and the Decline in Female Labour Force Participation in India*. IZA Discussion Paper 17368. IZA Institute of Labor Economics.

DINKELMAN, TARYN, & NGAI, L. RACHEL. 2022. Time Use and Gender in Africa in Times of Structural Transformation. *Journal of Economic Perspectives*, **36**(1), 57–80.

DREZE, JEAN, & SEN, AMARTYA. 1996. *India economic development and social opportunity*. Oxford University Press.

ESWARAN, MUKESH, RAMASWAMI, BHARAT, & WADHWA, WILIMA. 2013. Status, Caste, and the Time Allocation of Women in Rural India. *Economic Development and Cultural Change*, **61**(2), 311–333.

FLETCHER, ERIN K., PANDE, ROHINI, & TROYER MOORE, CHARITY. 2017 (Dec.). *Women and Work in India: Descriptive Evidence and a Review of Potential Policies*. CID Faculty Working Paper 339. Center for International Development at Harvard University.

GADDIS, ISIS, OSENI, GBEMISOLA, PALACIOS-LÓPEZ, AMPARO, & PIETERS, JANNEKE. 2020. *Who Is Employed? Evidence from Sub-Saharan Africa on Redefining Employment*. Tech. rept. 9370. World Bank.

GELMAN, ANDREW, & IMBENS, GUIDO. 2019. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, **37**(3), 447–456.

GOLDIN, CLAUDIA. 1994. *The U-shaped female labor force function in economic development and economic history*.

IMBENS, GUIDO W, & LEMIEUX, THOMAS. 2008. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, **142**(2), 615–635.

INTERNATIONAL LABOUR OFFICE. 2013. *Resolution Concerning Statistics of Work, Employment and Labour Underutilization*. Adopted by the 19th International Conference of Labour Statisticians, Geneva.

JAYACHANDRAN, SEEMA. 2021. Social Norms as a Barrier to Women's Employment in Developing Countries. *IMF Economic Review*, **69**(3), 576–595.

KAPSOS, STEVEN, SILBERMANN, ANDREA, & BOURMPOULA, EVANGELIA. 2014. *Why is female labour force participation declining so sharply in India?* ILO Geneva.

MAMMEN, KRISTIN, & PAXSON, CHRISTINA. 2000. Women's Work and Economic Development. *Journal of Economic Perspectives*, **14**(4), 141–164.

MEHROTRA, SANTOSH, & PARIDA, JAJATI K. 2017. Why is the labour force participation of women declining in India? *World Development*, **98**, 360–380.

NGAI, L RACHEL, & PETRONGOLO, BARBARA. 2017. Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, **9**(4), 1–44.

NGAI, L RACHEL, OLIVETTI, CLAUDIA, & PETRONGOLO, BARBARA. 2024. *Gendered change: 150 years of transformation in US hours*. Tech. rept. National Bureau of Economic Research.

OLIVETTI, CLAUDIA, & PETRONGOLO, BARBARA. 2014. Gender gaps across countries and skills: Demand, supply and the industry structure. *Review of Economic Dynamics*, **17**(4), 842–859.

OLIVETTI, CLAUDIA, & PETRONGOLO, BARBARA. 2016. The evolution of gender gaps in industrialized countries. *Annual review of Economics*, **8**(1), 405–434.

PATTAYAT, SHIBA SHANKAR, PARIDA, JAJATI KESHARI, & PALTASINGH, KIRTTI RANJAN. 2023. Gender Wage Gap among Rural Non-farm Sector Employees in India: Evidence from Nationally Representative Survey. *Review of Development and Change*, **28**(1), 22–44.

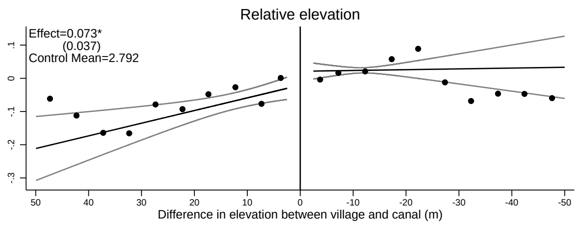
PETRONGOLO, BARBARA, & RONCHI, MADDALENA. 2020. Gender gaps and the structure of local labor markets. *Labour Economics*, **64**, 101819.

WORLD BANK. 2012. *World Development Report 2012: Gender Equality and Development*. Washington, DC: World Bank.

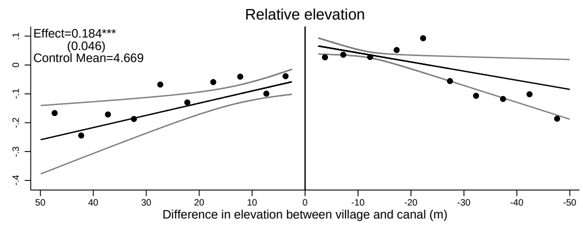
Supplementary Appendix for "Income Growth and Women's Work: Long-Run Evidence from India"

Angelica Bozzi   Rebecca Cai   Sam Asher

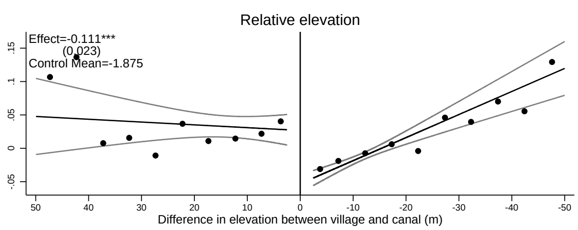
A1 RD graphs for main results



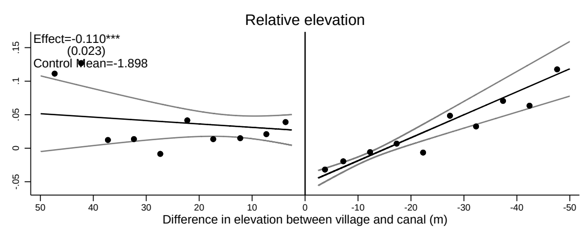
(a) Female years of education



(b) Male years of education



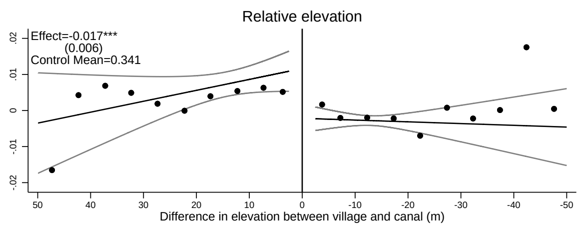
(c) Gender gap in education



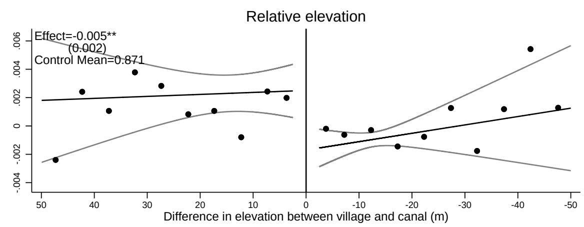
(d) Within-household education gap

Figure A1: Educational outcomes

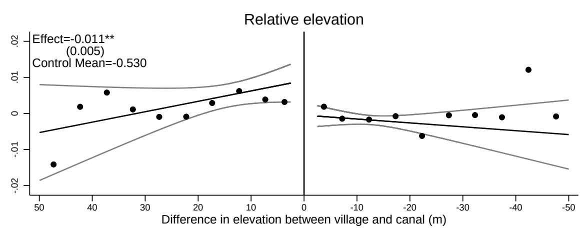
Notes: The regression-discontinuity plots visualize the canal treatment effect. The running variable (x-axis) is the village's elevation relative to the canal ($RelElev$), where villages below canal in elevation (to the right of the vertical line) can receive irrigation. The dots plot the coefficient on each $RelElev$ bin, black lines plot the fitted line through coefficients, and grey lines represent 95% confidence intervals. * 0.1 ** 0.05 *** 0.01. Village and canal elevation are both noisily measured, so we drop a "donut hole" of villages within 2.5 meters elevation of the canal, leading to the small gap around the $RelElev = 0$ line. The outcomes are: (a) years of education for adult women, (b) years of education for adult men, (c) the (F - M) gender gap in average education years among adults at the village-level, and (d) the analogous education gender gap, but within household.



(a) Female labor force participation



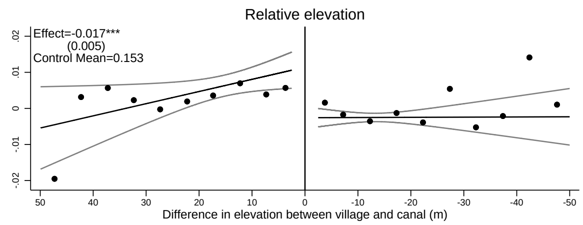
(b) Male labor force participation



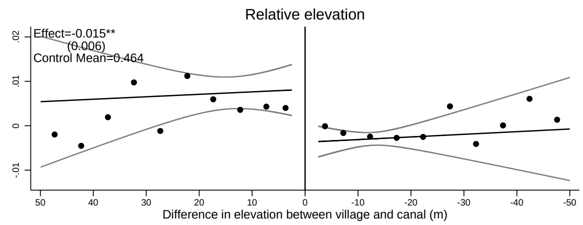
(c) Gender gap in labor force participation

Figure A2: General labor market outcomes

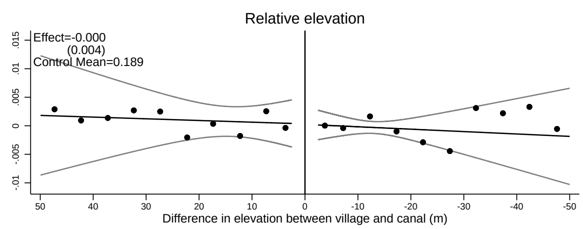
Notes: The regression-discontinuity plots visualize the canal treatment effect. The running variable (x-axis) is the village's elevation relative to the canal (*RelElev*), where villages below canal in elevation (to the right of the vertical line) can receive irrigation. The dots plot the coefficient on each *RelElev* bin, black lines plot the fitted line through coefficients, and grey lines represent 95% confidence intervals. * 0.1 ** 0.05 *** 0.01. Conceptually, the "canal treatment effect" in the regression tables captures the vertical jump in the outcome (y-axis) exactly at *RelElev* = 0. Village and canal elevation are both noisily measured, so we drop a "donut hole" of villages within 2.5 meters elevation of the canal, leading to the small gap around the $RelElev = 0$ line. The outcomes are: (a) share of adult women who report agriculture or paid non-agricultural work as their primary activity, (b) the analog of (a) for adult men, (c) the (F - M) gender gap labor force participation, i.e. a minus b.



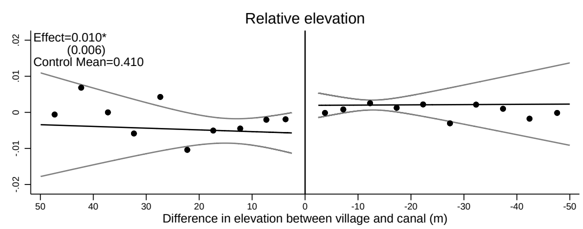
(a) Female agriculture



(b) Male agriculture



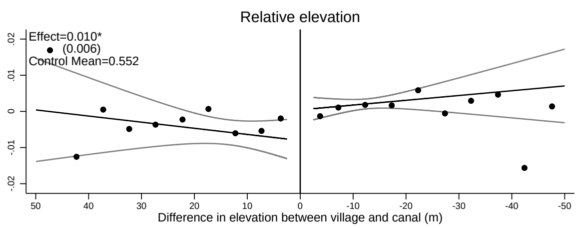
(c) Female nonfarm paid work



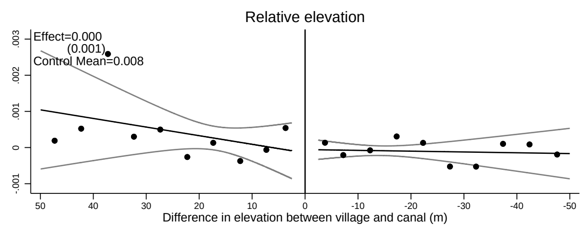
(d) Male nonfarm paid work

Figure A3: Paid work outcomes by sector

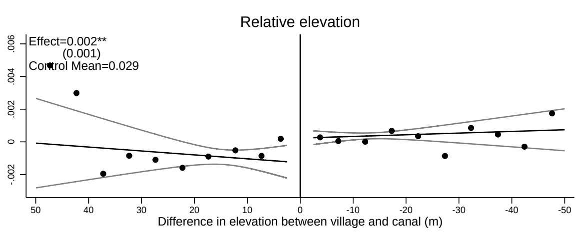
Notes: The regression-discontinuity plots visualize the canal treatment effect. The running variable (x-axis) is the village's elevation relative to the canal ($RelElev$), where villages below canal in elevation (to the right of the vertical line) can receive irrigation. The dots plot the coefficient on each $RelElev$ bin, black lines plot the fitted line through coefficients, and grey lines represent 95% confidence intervals. * 0.1 ** 0.05 *** 0.01. Conceptually, the "canal treatment effect" in the regression tables captures the vertical jump in the outcome (y-axis) exactly at $RelElev$ = 0. Village and canal elevation are both noisily measured, so we drop a "donut hole" of villages within 2.5 meters elevation of the canal, leading to the small gap around the $RelElev = 0$ line. The outcomes are: (a) female agricultural LFP, the share of adult women who report their primary activity is agriculture, (b) male agricultural LFP, (c) female non-agricultural LFP, the share of adult women whose primary activity is paid labor in service, manufacturing, or generic day labor, and (d) male non-agricultural LFP.



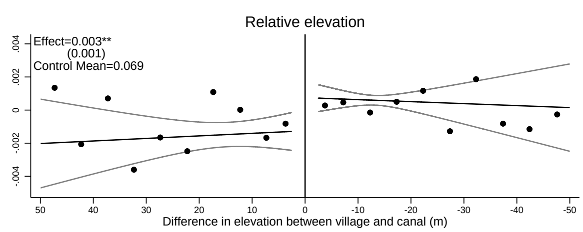
(a) Female home production



(b) Male home production



(c) Female studying



(d) Male studying

Figure A4: Unpaid work outcomes by sector

Notes: The regression-discontinuity plots visualize the canal treatment effect. The running variable (x-axis) is the village's elevation relative to the canal (*RelElev*), where villages below canal in elevation (to the right of the vertical line) can receive irrigation. The dots plot the coefficient on each *RelElev* bin, black lines plot the fitted line through coefficients, and grey lines represent 95% confidence intervals. * 0.1 ** 0.05 *** 0.01. Conceptually, the "canal treatment effect" in the regression tables captures the vertical jump in the outcome (y-axis) exactly at $RelElev = 0$. Village and canal elevation are both noisily measured, so we drop a "donut hole" of villages within 2.5 meters elevation of the canal, leading to the small gap around the $RelElev = 0$ line. The outcomes are: (a) the share of adult women who report their main activity is home production, i.e. house chores or unpaid household worker, (b) male home production share, (c) the share of adult women who are full time students, and (d) share of adult men who are full time students.

A2 Disaggregating agricultural labor effects by pay status

Table A1: Canal treatment effects on agricultural activities by pay status

	(1)	(2)	(3)	(4)	(5)	(6)
	Ag work, definitely paid		Ag work, unclear if paid		Unpaid work on own plot	
	Females	Males	Females	Males	Females	Males
Below Canal	0.0001	0.0008	-0.0079***	-0.0079*	-0.0076***	-0.0051
	(0.0031)	(0.0036)	(0.0027)	(0.0046)	(0.0021)	(0.0047)
Geographic controls	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓
Mean in control group	0.047	0.094	0.042	0.122	0.033	0.189
R-squared	0.383	0.345	0.266	0.320	0.259	0.427
Obs	37774	37774	37774	37774	37774	37774

Notes: This table reports the main RDD estimates following equation 1, using labor force participation outcomes from the SECC dataset. Each outcome variable is estimated separately, with the coefficient capturing the direct effect of canal irrigation, weighted by land area. For each reported coefficient, stars indicate the coefficient's significance, and the standard error clustered at the sub-district level is reported below in parentheses. For all columns, the outcome is the share of gender-specific village population aged 20-64 who self-report that their main job is the given activity. In Cols 1 and 2, the activity is agricultural work which is clearly paid, translating to "agricultural laborer" or "farm hired hand." In Cols 3 and 4, the activity is all agricultural work with unclear connotations about pay, translating to "agriculture" or similarly vague. In Cols 5 and 6, the activity is clearly unpaid work on ones own plot, translating to "cultivator" or "tenant farmer." To exclude areas of low-quality data, the sample of villages is restricted to districts where the overall female labor force participation rate is at least 2%, and the overall male LFP rate is at least 40%. * 0.1 ** 0.05 *** 0.01.

A3 Landownership disaggregation

Table A2: Education outcomes by household landownership

	Years of education		Education year gap (F - M)	
	(1)	(2)	(3)	(4)
	Female	Male	Within-village	Within-household
Below Canal: landless	0.0533*	0.1399***	-0.0842***	-0.0851***
	(0.0315)	(0.0424)	(0.0268)	(0.0266)
Below Canal × Landowner	0.0846***	0.1399***	-0.0530**	-0.0521**
	(0.0317)	(0.0369)	(0.0240)	(0.0236)
Geographic controls	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓
Control group mean, landless	2.382	4.147	-1.762	-1.769
Control group mean, landowner	2.895	5.075	-2.180	-2.205
R-squared	0.554	0.465	0.403	0.403
Obs	110037	110024	110018	109925

Notes: The table reports the main RDD estimates on adult educational attainment. Panel A reports results for households that own any land, while Panel B for landless households. For all education outcomes, we restrict to women and men aged 18 and over. Columns 1 and 2 report average years of education for women and men aged 18 and above. Columns 3 and 4 report the female-minus-male gap at the within-village and within-household levels. * 0.1 ** 0.05 *** 0.01

Table A3: Labor market outcomes by household landownership

	Overall LFP		Agricultural LFP		Nonag LFP		HH Production		Studying	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Below Canal: landless	-0.0150**	-0.0056**	-0.0109**	-0.0114*	-0.0045	0.0056	0.0108*	0.0010	0.0015	0.0032**
	(0.0065)	(0.0024)	(0.0050)	(0.0065)	(0.0054)	(0.0064)	(0.0063)	(0.0007)	(0.0011)	(0.0016)
Below Canal × Landowner	-0.0100**	-0.0028	-0.0095**	0.0057	-0.0005	-0.0084	0.0071	-0.0006	0.0025**	0.0032*
	(0.0050)	(0.0025)	(0.0043)	(0.0051)	(0.0041)	(0.0052)	(0.0049)	(0.0007)	(0.0013)	(0.0017)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Control group mean, landless	0.355	0.887	0.115	0.334	0.243	0.562	0.545	0.009	0.020	0.048
Control group mean, landowner	0.267	0.853	0.143	0.507	0.126	0.350	0.617	0.010	0.034	0.079
R-squared	0.511	0.385	0.372	0.486	0.448	0.516	0.522	0.115	0.361	0.388
Obs	36336	36336	36336	36336	36336	36336	36336	36336	36336	36336

Notes: Panel A reports the main RDD estimates following equation 1, using labor force participation outcomes from the SECC dataset, restricted to households that own agricultural land. Panel B reports the same regression, but restricting the sample to households that do not own agricultural land. Each outcome variable is estimated separately, with the β_1 coefficient capturing the direct effect of canal irrigation, Weighted by land area. For each reported coefficient, stars indicate the coefficient's significance, and the standard error clustered at the sub-district level is reported below in parentheses. For all columns, the outcome is the share of gender-specific village population aged 20-64 who self-report that their main activity is the given activity. In columns 1 and 2, the activity is any work, including jobs that could not be classified by sector, e.g. "laborer". In columns 3 and 4, the activity is any agricultural activity. In columns 5 and 6, the activity is work in services. In columns 7 and 8, the activity is unpaid housework, predominately housewives and to a lesser degree, unpaid positions in household enterprises. In columns 9 and 10, the activity is any form of higher education. To exclude areas of low-quality data, the sample of villages is restricted to districts where the overall female labor force participation rate is at least 2%, and the overall male LFP rate is at least 40%. * 0.1 ** 0.05 *** 0.01

A4 Heterogeneity by marital status

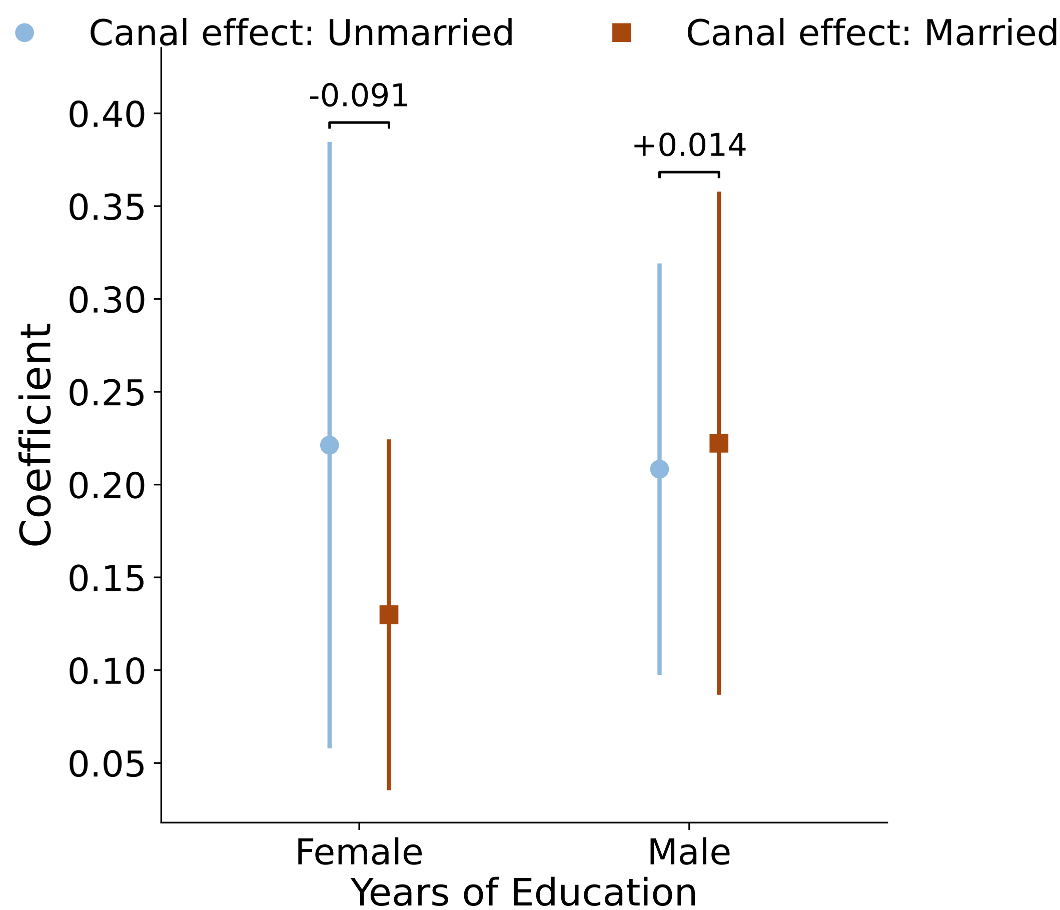
In this setting, households are primarily patrilocal, meaning that a woman moves out of her birth village to her husband's village after marriage. A village-level estimate of canal effects on women overall therefore bundles effects on married women (selective in-migration of brides) with effects on unmarried women (child investment decisions of treated parents). To separate these two channels, we re-estimate a person-level analogue of Equation 1, fully interacted with an indicator for marital status. Let $Married_p$ be in an indicator equaling 0 if a person is never married, and 1 if married. Divorced, separated, and widowed individuals are dropped because these outcomes are rare and highly selected. Because marriage is nearly universal after age 30, we restrict the sample to people aged 18-29. For person p residing in village v, subdistrict s, we estimate

$$\begin{aligned} y_{p,v,s} = \beta_0 + \beta_1 BelowCanal_v + \beta_2 Married_p + \beta_3 (BelowCanal_v \times Married_p) \\ + \beta_4 W_{v,s,p} + \beta_5 (Married_p \times W_{v,s,p}) + v_{s,Married} + \epsilon_{p,v,s}, \end{aligned} \tag{A1}$$

where $W_{v,s,p}$ collects the running variable, all geographic controls from Equation 1: the relative-elevation running variable (allowed a separate slope above and below the canal cutoff), the geographic controls $X_{v,s}$, and person-level demographic controls (caste category, land ownership, age, and age squared). Every element of $W_{v,s,p}$ is interacted with $Married_p$, so married and unmarried individuals are allowed fully separate control functions – β_3 is therefore not confounded by marital-status differences in, for example, the age or caste gradient of the outcome. As in Equation 1, v_s is a subdistrict fixed effect (also interacted with $Married_p$ and standard errors are clustered at the subdistrict level.

β_1 recovers the effect of canal access for unmarried individuals, and $\beta_1 + \beta_3$ recovers the effect for married individuals. β_3 is therefore the coefficient of interest in this section: it captures how much the estimated effect of canal access differs for individuals who married into their current village relative to those who did not, and so how much in-migration of brides could be biasing the baseline estimate in Equation 1.

Figure A5: Coefficient plot: heterogeneity in education outcomes by marital status



Notes: The coefficient plot displays the canal treatment effect interacted with individual marital status for adults aged 18-29. Blue circles plot the canal treatment effects and 95% CI on the baseline group (never married), corresponding to β_1 in Equation A1. Brackets above each cluster display β_3, the coefficient on $(BelowCanal_v \times Married_p)$ capturing the additional change in the outcome for married, compared to never married, households. Stars on brackets indicate significance of β_3 (* $p < 0.10$, ** 0.05, *** 0.01). Red squares plot the treatment effect and 95% CI on married adults, where standard error of the treatment effect ($\hat{\beta}_1 + \hat{\beta}_3$) is the square root of:

$$\widehat{Var}(\hat{\beta}_1 + \hat{\beta}_3) = \widehat{Var}(\hat{\beta}_1) + \widehat{Var}(\hat{\beta}_3) + 2\widehat{Cov}(\hat{\beta}_1, \hat{\beta}_3)$$

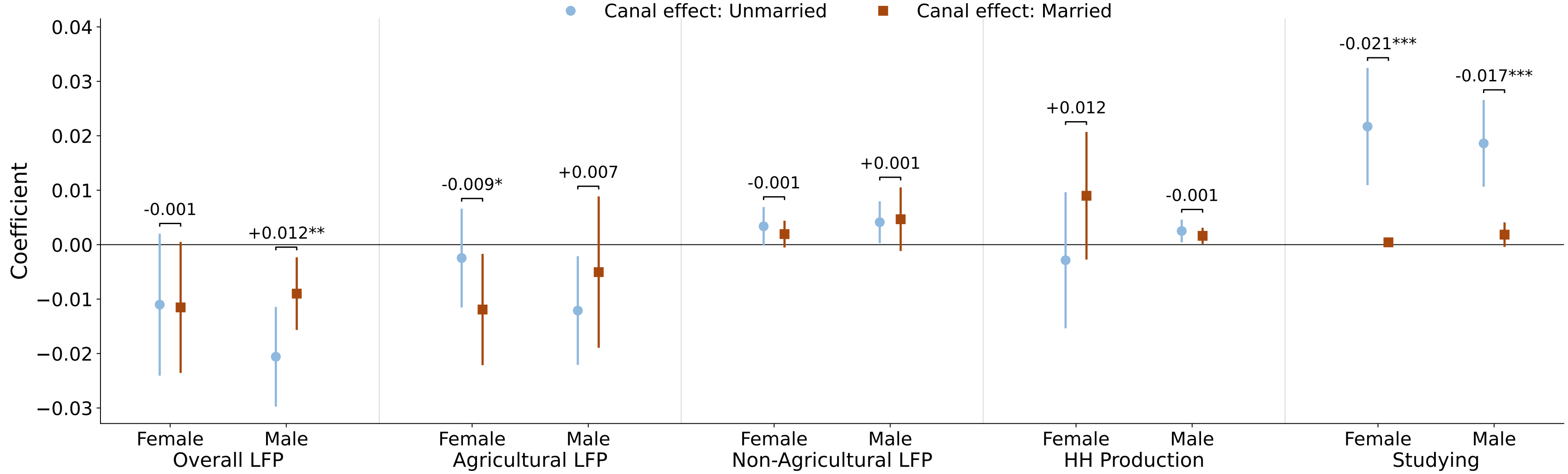
.

Table A4: Education outcomes by marital status: population 18-29 y.o.

	Years of education	
	(1)	(2)
	Female	Male
Below Canal: never married	0.2212***	0.2083***
	(0.0832)	(0.0565)
Below Canal × Married	-0.0914	0.0140
	(0.0728)	(0.0597)
Geographic controls	✓	✓
Subdistrict FE	✓	✓
Person demographic controls	✓	✓
Control group mean, unmarried	7.265	8.021
Control group mean, married	4.631	5.670
R-squared	0.289	0.234
Obs	3451502	3719506

Notes: The table reports the marital status heterogeneity estimates on adult educational attainment, following Equation A1. The sample is adults aged 18-29. The regression is run at the individual-level, with individual demographic controls (caste, age, landownership status), village-level geographic controls, and subdistrict fixed-effects (also interacted with marital status). * 0.1 ** 0.05 *** 0.01

Figure A6: Coefficient plot: heterogeneity in labor outcomes by marital status



Notes: The coefficient plot displays the canal treatment effect interacted with individual marital status for adults aged 18-29. Blue circles plot the canal treatment effects and 95% CI on the baseline group (never married), corresponding to β_1 in Equation A1. Brackets above each cluster display β_3, the coefficient on $(BelowCanal_v \times Married_p)$ capturing the additional change in the outcome for married, compared to never married, households. Stars on brackets indicate significance of β_3 (* $p < 0.10$, ** 0.05, *** 0.01). Red squares plot the treatment effect and 95% CI on married adults, where standard error of the treatment effect ($\hat{\beta}_1 + \hat{\beta}_3$) is the square root of:

$$\widehat{Var}(\hat{\beta}_1 + \hat{\beta}_3) = \widehat{Var}(\hat{\beta}_1) + \widehat{Var}(\hat{\beta}_3) + 2\,\widehat{Cov}(\hat{\beta}_1, \hat{\beta}_3)$$

.

Table A5: Labor outcomes by marital status: population 18-29 y.o.

	Overall LFP		Ag LFP		Nonag LFP		HH Prod (unpaid)		Study (unpaid)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Below Canal: never married	-0.0110*	-0.0206***	-0.0025	-0.0121**	0.0034*	0.0041**	-0.0029	0.0025**	0.0217***	0.0186***
	(0.0066)	(0.0047)	(0.0046)	(0.0051)	(0.0018)	(0.0020)	(0.0064)	(0.0011)	(0.0055)	(0.0041)
Below Canal × Married	-0.0005	0.0116**	-0.0095*	0.0071	-0.0014	0.0005	0.0118	-0.0009	-0.0213***	-0.0168***
	(0.0072)	(0.0049)	(0.0050)	(0.0055)	(0.0016)	(0.0024)	(0.0072)	(0.0011)	(0.0054)	(0.0039)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Person demographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Control group mean, unmarried	0.190	0.529	0.058	0.184	0.024	0.082	0.285	0.013	0.347	0.326
Control group mean, married	0.264	0.882	0.109	0.372	0.018	0.103	0.627	0.011	0.010	0.031
R-squared	0.155	0.196	0.097	0.158	0.050	0.067	0.232	0.013	0.360	0.212
Obs	4030958	4296944	4030958	4296944	4030958	4296944	4030958	4296944	4030958	4296944

Notes: The table reports the marital status heterogeneity estimates on adult educational attainment, following Equation A1. The sample is adults aged 18-29. The regression is run at the individual-level, with individual demographic controls (caste, age, landownership status), village-level geographic controls, and subdistrict fixed-effects (also interacted with marital status). * 0.1 ** 0.05 *** 0.01

A5 Robustness checks

A5.1 Verifying main labor results in alternative data source

Our main labor results rely on manual coding of free-text descriptions of adults' primary activity into standardized labor activity codes in the SECC dataset. To address concerns that our activity coding choices could generate the main results, we replicate the main labor findings in an entirely different village-level dataset: the 2011 Village Primary Census Abstract (PCA). The PCA is constructed as part of India's decennial census. Village bureaucrats reports population counts by gender and labor status. We calculate "labor force participation rate" as the share of adult (18+) population who work for more than 6 months of the year (called “main” as opposed to “marginal” workers in the PCA). The PCA does not have information about participation in household production or higher education, so we cannot directly verify our claims about reallocation into home production or studying.

Village bureaucrats determining work status in the PCA have different definitions of labor force participation compared to people who self-label their main activity in the SECC, especially for women. As a result, the baseline LFP rates in the control group are quite different. Control group average FLFP is 30.5% in the SECC versus 66% in the PCA, and MLFP in the SECC is 87% in the SECC versus 92% in the PCA. Despite these differences in measurement, we observe the same basic canals result in the PCA: in canal villages, FLFP is is 2.7PP lower ($p < 0.01$) and MLFP is 1.0 PP lower ($p < 0.01$). The drop in overall LFP is driven entirely by exit from agriculture, and is stronger for women.

Table A6: Canal treatment effects on labor using alternative data source

	Total Employment		Agriculture		Other Sectors	
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Male	Female	Male	Female	Male
Below Canal	-0.0267***	-0.0103***	-0.0269***	-0.0125**	0.0001	0.0022
	(0.0061)	(0.0030)	(0.0066)	(0.0058)	(0.0027)	(0.0046)
Geographic controls	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓
Mean in control group	0.662	0.918	0.561	0.711	0.098	0.205
R-squared	0.448	0.311	0.464	0.384	0.297	0.369
Obs	80120	80135	80120	80135	80120	80135

Notes: The table displays the canal treatment effect on village-level fertility and marriage outcomes. The specification is the same as the main results. Columns 1 and 2 measure the share of adult population who work more than 6 months per year in any sector. Columns 3 and 4 measure the share of adults working 6+ months per year as cultivators (own-field) or agricultural laborers (paid labor on others' fields). Columns 5 and 6 reports share of adult population working in manufacturing/processing in household industries, and other services. * 0.1 ** 0.05 *** 0.01

A5.2 Null effects on fertility and marriage outcomes

Table A7: Canal treatment effects on marriage and fertility

	(1)	(2)	(3)	(4)	(5)	(6)
	Children per HH	Sex ratio (F/M) 0-5	Sex ratio (F/M) 1st-born	Married Men 18+	Married Women 18+	Married Women 16-17
Below Canal	-0.0030	-0.0117*	-0.0037	-0.0017	-0.0018	-0.0018
	(0.0093)	(0.0061)	(0.0049)	(0.0017)	(0.0015)	(0.0030)
Geographic controls	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓
Mean in control group	1.799	0.964	0.936	0.708	0.781	0.090
R-squared	0.704	0.092	0.133	0.441	0.509	0.306
Obs	57811	57304	57081	56960	57173	54102

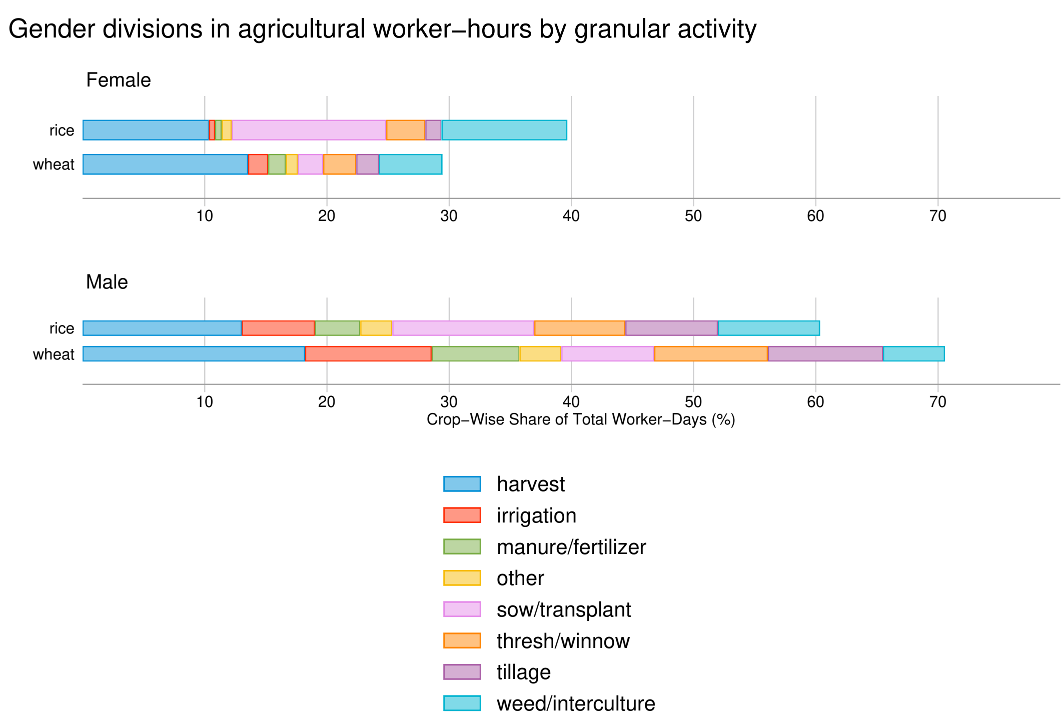
Notes: The table displays the canal treatment effect on village-level fertility and marriage outcomes. The specification is the same as the main results. Column 1 measures number of children aged 0-17 living in the household and proxies fertility, because we do not observe women's complete birth history. Column 2 uses the female-over-male sex ratio among infants aged 0-5. A sex ratio less than 1 indicates that birth ratio is male-skewed. Column 3 is the analogous sex ratio for the oldest child aged 0-17 still living in the household. Column 4 is the share of men aged 18-plus who have ever been maried. Column 5 is the analogous marriage rate for women aged 18-plus. Column 6 is a proxy for underage marriage of girls and measures share of girls aged 16-17 who are ever-married. Source: SECC. * 0.1 ** 0.05 *** 0.01

A5.3 Ruling out labor demand story: agricultural tasks by gender

In Figure A7, we use the 1999 India Rural Economic Development Survey (REDS) dataset to plot the gendered division of agricultural tasks for India's staple grains: rice and wheat. For each crop, the "female" and "male" bars total to 100% of the total worker-days used in cultivating the crop, disaggregated by activity type (harvesting, irrigation, etc). The data combines household member time spent as a casual worker, permanent worker, or work on family farm. The critical result is that irrigation tasks that would be directly displaced by canals (i.e. irrigation) are performed overwhelmingly by men, not women. Thus, if canals reduced agricultural labor force participation through direct technological displacement of irrigation-related tasks, men would leave agriculture at higher rates.

Source: ARIS-REDS data collected by the National Council of Applied Economic Research in Delhi. The data is representative of the rural population of India in 17 major states. It comprised 4,257 households in 259 villages, as part of a panel of four rounds between 1971 and 2006. We use the 1999 round only. It is the only dataset that dis-aggregates agricultural time use by both crop and gender.

Figure A7: Workers-hours by gender



A5.4 Spillover canal effects to control settlements

A key concern in our empirical strategy is the possibility that labor markets are integrated such that canal construction affects not only treated villages but also nearby control areas. For example, migration from villages just above the canal toward canal-irrigated villages below could reduce labor supply in the untreated areas, potentially increasing wages and female labor force participation there.

To assess the possibility of spillovers through integrated local labor markets, we estimate equation A5.4, comparing our control group (villages above and nearby a canal) to villages at the same elevation but more than 20 kilometers from a canal, and therefore unlikely to be affected by its presence.

$$y_{i,s} = \alpha_0 + \alpha_1(AboveCanal_{i,s} > 0) + \alpha_2 X_{i,s} + \gamma_d + \epsilon_{i,s}$$

where $AboveCanal_{i,s}$ is an indicator equal to 1 if settlement i lies just above a canal (within 10 km), and 0 if it lies at least 20 km away from the nearest canal but within 50 meters of elevation. Xi, s is the same vector of time-invariant geophysical controls used in the main RDD specification, including relative elevation, which here does not serve as a running variable. γ_d denotes district fixed effects.

We find no evidence of spillover effects on labor force participation or gender gaps: labor market outcomes are statistically indistinguishable between the near- and distant-control groups (Table A8), and there are no labor effects by sector (Table A9). Similarly, we find no evidence of spillover effects on educational attainment (Table A10). This suggests that our control villages are not significantly affected by proximity to the canal. [6]

[6] We have also run the analysis comparing our control villages to villages farther than 30km and found similar results.

Table A8: Spillover: Labor Market Outcomes

	(1) Female LFP	(2) Male LFP	(3) Gender Gap
1 control, 0 distant village	0.0139	0.0002	0.0137
	(0.0093)	(0.0027)	(0.0086)
Geographic controls	✓	✓	✓
District FE	✓	✓	✓
Mean in control group	0.258	0.871	-0.613
R-squared	0.473	0.300	0.509
Obs	24923	24923	24923

Notes: This table reports the main RDD estimates following equation A5.4, using labor force participation outcomes from the SECC dataset. Each outcome variable is estimated separately, with the β_1 coefficient capturing the direct effect of being just above a canal relative to distant from canals, weighted geographic characteristics. For each reported coefficient, stars indicate the coefficient's significance, and the standard error clustered at the district level is reported below in parentheses. In Columns 1 and 2, the outcome is the share of the gender-specific village population aged 20–64 who self-report being engaged in any work activity. In Column 3, the outcome is the ratio of the female to male work participation rates. To exclude areas of low-quality data, the sample of villages is restricted to districts where the overall female labor force participation rate is at least 2%, and the overall male LFP rate is at least 40%.

Table A9: Spillover: Labor Market Outcomes by Sector

	LFP				Not LFP			
	Agriculture		Non-agriculture		HH production		Studying	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female	Male	Female	Male	Female	Male	Female	Male
1 Above Canal, 0 Distant Village	0.010	0.007	0.004	-0.007	-0.012	0.001	0.002	0.001
	(0.008)	(0.010)	(0.006)	(0.010)	(0.010)	(0.001)	(0.001)	(0.002)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	0.126	0.448	0.183	0.434	0.580	0.008	0.024	0.058
R^2	0.30	0.25	0.33	0.23	0.48	0.06	0.26	0.34
Obs	24,923	24,923	24,923	24,923	24,923	24,923	24,923	24,923

Notes: Notes: This table reports the main RDD estimates following equation A5.4, using labor force participation outcomes from the SECC dataset. Each outcome variable is estimated separately, with the β_1 coefficient capturing the direct effect of being just above a canal relative to distant from canals, weighted by geographic characteristics. For each reported coefficient, stars indicate the coefficient's significance, and the standard error clustered at the district level is reported below in parentheses. For all columns, the outcome is the share of gender-specific village population aged 20-64 who self-report that their main job is the given activity. In columns 1 and 2, the activity is any agricultural activity. In columns 3 and 4, the activity is work in services. In columns 5 and 6, the activity is work in manufacturing. In columns 7 and 8, the activity is jobs that could not be classified by sector, e.g., "laborer". In columns 9 and 10, the activity is any form of education. In columns 11 and 12, the activity is unpaid housework, including housewives and unpaid positions in household enterprises. To exclude areas of low-quality data, the sample of villages is restricted to districts where the overall female labor force participation rate is at least 2%, and the overall male LFP rate is at least 40%.

Table A10: Spillover: Education Outcomes

	(1)	(2)	(3)	(4)
	Years of Education		Gaps in Education	
	Females	Males	Village	Household
1 control, 0 distant village	-0.0747	-0.0730	-0.0003	0.0015
	(0.0655)	(0.0846)	(0.0365)	(0.0366)
Geographic controls	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Mean in control group	2.858	4.677	-1.817	-1.816
R-squared	0.472	0.363	0.385	0.388
Obs	26497	26496	26495	26469

Notes: This table reports the main RDD estimates following equation A5.4, using education outcomes from the SECC dataset. Each outcome variable is estimated separately, with the β_1 coefficient capturing the direct effect of being just above a canal relative to distant from canals, weighted by geographic characteristics. For each reported coefficient, stars indicate the coefficient's significance, and the standard error clustered at the district level is reported below in parentheses. In Columns 1 and 2 the outcome is the average years of education for females and males. In Column 3 the outcome is the gender gap in average years of education (females minus males) within the village. In Column 4 the outcome is the gender gap within the household.

A5.5 Migration across treatment villages and demographic composition changes

Next we directly assess if migration in response to canal construction altered the demographic composition of treated areas in ways that could bias our results. Canal settlements experienced substantial population growth, with 22% higher population density (Asher *et al.*, 2022). We examine whether this expansion led to systematic differences in population characteristics between canal and non-canal villages.

We find no significant differences in age, gender, or marital status composition, but some differences in caste composition. Population pyramids in Appendix Figure A8 show age and gender population distributions extremely similar for villages above and below canals. Likewise, composition of marital status is balanced, with a similar share of the adult population being married in canal-irrigated and control villages (Appendix Table A11). Appendix Table A12 indicates that villages below the canals have a higher share of Scheduled Caste (SC) population and a lower share of Scheduled Tribe (ST) population compared to above-canal villages. SC and ST are both government categories for historically and socio-economically disadvantaged castes or tribes.

Although we find limited evidence of demographic change, we use the SECC microdata to re-estimate our main results at the individual level, which allows us to control for these demographic factors. The regressions include roughly 12 million adult women and 13 million adult men who live in the settlements in the main analysis sample, and are reweighted by inverse settlement adult population (by gender) to ensure that each settlement is equally weighted in the regression. The specification for person p, living in settlement v, in subdistrict s

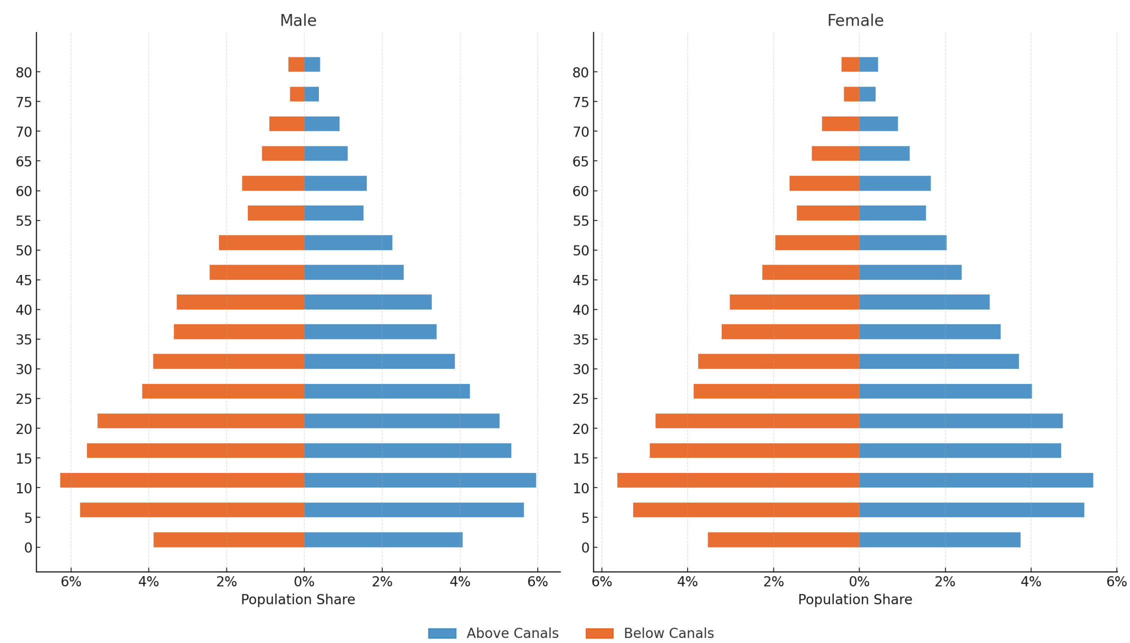
$$\begin{aligned} y_{p,v,s} = \beta_0 + \beta_1(BelowCanal_{v,s}) + \beta_2 RelElev_{v,s} + \\ + \beta_3(RelElev_{v,s} \times BelowCanal_{v,s}) + \beta_4 X_{v,s} + \beta_5 W_{p,v,s} + v_s + \epsilon_{i,s} \end{aligned} \tag{A2}$$

Here, $y_{p,v,s}$ represents the person-level outcome (e.g. working of not) in settlement v within subdistrict s. These regressions control for individual demographic controls $W_{p,v,s}$: age, age-squared, marital status, and caste category (SC, ST, Other). If our main results were driven by changes in caste or age composition, then the "Below Canal" coefficient would become null after controlling for individual demographic traits.

We report the individual-level regressions with demographic controls for labor (Appendix

Table A13) and education (Appendix Table A14), and find results highly consistent with our village-aggregated main results in magnitude and significance, essentially unchanged. This consistency suggests that although migration may have occurred, our main findings are not driven by selective migration or compositional changes, which explain the observed patterns in labor force participation or education.

Figure A8: Population by age and gender below and above canals



Notes: The figure plots the age distribution of treatment (blue) and control (orange) villages. The purpose is to allow visual inspection of whether the greater population density of treatment villages translates to a generational demographic shift that could drive results. Each bar plots the population share of a 5-year age bin (e.g. the "0" bar represents ages 0-4). Source: SECC.

Table A11: Balance by Marital Status within Gender

	Never married		Married		Widow		Sep'd/Divorced	
	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male	(7) Female	(8) Male
Below Canal	0.0010	0.0006	-0.0014	-0.0010	0.0002	-0.0001	0.0003	0.0005
	(0.0016)	(0.0019)	(0.0019)	(0.0020)	(0.0009)	(0.0004)	(0.0006)	(0.0006)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	0.084	0.231	0.842	0.747	0.062	0.015	0.011	0.007
N shrids	49203	49624	49203	49624	49203	49624	49203	49624
R-squared	.0333	.0283	.0211	.0264	.00784	.00409	.00739	.0096
Obs	1.03e+07	1.09e+07	1.03e+07	1.09e+07	1.03e+07	1.09e+07	1.03e+07	1.09e+07

Notes: The table tests for differences in marital status composition across treatment and control villages. The sample is the roughly 12 million adult women and 13 million adult men in the main analysis village sample. If treatment villages had a significantly higher population share of never married women, then the coefficient in Column 1 would be positive and significant. Source: SECC. * 0.1 ** 0.05 *** 0.01

Table A12: Balance by caste composition within gender

	Scheduled Caste		Scheduled Tribe		Other		Declared no caste	
	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male	(7) Female	(8) Male
Below Canal	0.0151***	0.0133***	-0.0142***	-0.0134***	-0.0016	-0.0003	0.0007	0.0004
	(0.0044)	(0.0043)	(0.0054)	(0.0052)	(0.0061)	(0.0059)	(0.0014)	(0.0014)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	0.143	0.144	0.184	0.176	0.658	0.666	0.015	0.015
N shrids	55579	55648	55579	55648	55579	55648	55579	55648
R-squared	.0599991	.0584392	.3361395	.3348756	.1330238	.1318642	.0857077	.0833504
Obs	1.34e+07	1.43e+07	1.34e+07	1.43e+07	1.34e+07	1.43e+07	1.34e+07	1.43e+07

Notes: The table tests for differences in broad caste composition across treatment and control villages. The sample is the roughly 12 million adult women and 13 million adult men in the main analysis village sample. The outcome in Columns 1-2 is an indicator the individual self-reports as a member of a Scheduled Caste. In Columns 3-4, the individual self-reports as Scheduled Tribe. In Columns 5-6, the individual self-reports as "Other" caste, including Forward and Other Backward Caste. In Columns 6-7, the individual chooses to declare no caste. If treatment villages had a significantly higher population share of Scheduled Caste women, then the coefficient in Column 1 would be positive and significant. Source: SECC. * 0.1 ** 0.05 *** 0.01

Table A13: Labor Market Outcomes by Sector, controlling for demographics

	Agriculture		Non-ag		HH production		Studying	
	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male	(7) Female	(8) Male
Below Canal	-0.0130***	-0.0114*	0.0023*	0.0081***	0.0109**	0.0008*	0.0018***	0.0032***
	(0.0047)	(0.0060)	(0.0012)	(0.0025)	(0.0054)	(0.0005)	(0.0005)	(0.0010)
Scheduled Caste	-0.0065**	-0.1733***	-0.0035***	-0.0415***	-0.1044***	-0.0010***	-0.0036***	-0.0081***
	(0.0030)	(0.0056)	(0.0007)	(0.0020)	(0.0042)	(0.0002)	(0.0004)	(0.0008)
Scheduled Tribes	0.0288***	-0.0584***	-0.0073***	-0.0604***	-0.1288***	-0.0011***	-0.0092***	-0.0171***
	(0.0041)	(0.0085)	(0.0009)	(0.0028)	(0.0062)	(0.0003)	(0.0006)	(0.0011)
Declared no caste	-0.0023	-0.0353***	0.0002	0.0047	-0.0084	0.0007	-0.0018	-0.0043**
	(0.0082)	(0.0099)	(0.0018)	(0.0058)	(0.0114)	(0.0011)	(0.0012)	(0.0018)
Married	0.0371***	0.1404***	-0.0109***	0.0016	0.3111***	-0.0024***	-0.2825***	-0.1861***
	(0.0029)	(0.0031)	(0.0010)	(0.0012)	(0.0060)	(0.0003)	(0.0071)	(0.0038)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓
Person demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	0.117	0.403	0.021	0.121	0.583	0.009	0.030	0.072
N shrids	34675	34675	34675	34675	34675	34675	34675	34675
R-squared	.0979796	.1549569	.0360053	.0550064	.2262032	.0063708	.2691678	.2371756
Obs	1.21e+07	1.28e+07	1.21e+07	1.28e+07	1.21e+07	1.28e+07	1.21e+07	1.28e+07

Notes: The table re-estimates the canal treatment effect, using individual-level outcomes from the SECC microdata, rather than village-level averages. The regressions include roughly 12 million adult women and 13 million adult men who live in the settlements in the main analysis sample, and are reweighted by inverse settlement adult population (by gender) to ensure that each settlement is equally weighted in the regression. These regressions control for individual age, age-squared, marital status fixed effects, and caste category fixed effects (SC, ST, Other, None). In Columns 1-2, the outcome is an indicator for whether the adult's self-reported main activity is agriculture. Cols 3-4 use the analogous indicator for non-agricultural paid work. Cols 5-6 use unpaid home production (e.g. housewife, house chores). Column 7-8 uses an indicator that the adult's full-time activity is "student." To exclude areas of low-quality data, the sample of villages is restricted to districts where the overall female labor force participation rate is at least 2%, and the overall male LFP rate is at least 40%. Source: SECC. * 0.1 ** 0.05 *** 0.01

Table A14: Education Outcomes, controlling for demographics

	Years of ed.		Primary school		Middle school		Secondary school	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female	Male	Female	Male	Female	Male	Female	Male
Below Canal	0.0615*	0.1497***	0.0044	0.0129**	0.0058*	0.0133***	0.0046*	0.0106**
	(0.0373)	(0.0484)	(0.0048)	(0.0053)	(0.0035)	(0.0049)	(0.0026)	(0.0042)
Scheduled Caste	-1.0998***	-1.1608***	-0.1221***	-0.1127***	-0.0914***	-0.1031***	-0.0581***	-0.0807***
	(0.0403)	(0.0496)	(0.0045)	(0.0051)	(0.0037)	(0.0046)	(0.0025)	(0.0036)
Scheduled Tribes	-1.7926***	-2.2084***	-0.2016***	-0.2189***	-0.1471***	-0.1983***	-0.0972***	-0.1527***
	(0.0833)	(0.0948)	(0.0091)	(0.0094)	(0.0075)	(0.0085)	(0.0057)	(0.0073)
Declared no caste	0.0706	-0.3235***	0.0072	-0.0313**	-0.0053	-0.0304***	-0.0098*	-0.0343***
	(0.0889)	(0.1085)	(0.0122)	(0.0130)	(0.0088)	(0.0112)	(0.0058)	(0.0090)
Married	-1.7393***	-1.2169***	-0.1172***	-0.0683***	-0.1429***	-0.1046***	-0.1567***	-0.1227***
	(0.0467)	(0.0503)	(0.0047)	(0.0048)	(0.0046)	(0.0043)	(0.0043)	(0.0041)
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓
Person demographic controls	✓	✓	✓	✓	✓	✓	✓	✓
Mean in control group	3.380	5.540	0.395	0.604	0.242	0.430	0.146	0.294
N shrids	49189	49568	49189	49568	49189	49568	49189	49568
R-squared	.2841479	.2212726	.2432205	.1885904	.2109289	.1789489	.1659539	.1445969
Obs	1.03e+07	1.09e+07	1.03e+07	1.09e+07	1.03e+07	1.09e+07	1.03e+07	1.09e+07

Notes: The table re-estimates the canal treatment effect, using individual-level outcomes from the SECC microdata, rather than village-level averages. The regressions include roughly 12 million adult women and 13 million adult men who live in the settlements in the main analysis sample, and are reweighted by inverse settlement adult population (by gender) to ensure that each settlement is equally weighted in the regression. These regressions control for individual age, age-squared, marital status fixed effects, and caste category fixed effects (SC, ST, Other, None). In Column 1-2, the outcome is the adult's completed years of education. Columns 3-4, 5-6, and 7-8 report indicators that the adult completed primary, middle, and secondary school, respectively. Source: SECC. * 0.1 ** 0.05 *** 0.01