Internal Migration and the Spatial Reorganization of Agriculture *

Raahil Madhok†
University of Minnesota

Frederik Noack‡
University of British Columbia

Ahmed Mushfiq Mobarak§
Yale University

Olivier Deschenes¶
University of California, Santa Barbara

September 30, 2024

Abstract

This paper studies how food production responds to the loss of agricultural labor during the process of urbanization and structural transformation. Using household microdata from India and exogenous variation in migration opportunities induced by urban income shocks, we show that migrant-sending households do not replace labor with capital when facing labor loss. Instead, they *lower* agricultural technology use, cultivate less land, and reduce crop production. In contrast, non-migrant households in those villages and in more remote regions growing similar crop mixes expand agricultural investments and production as land prices and crop prices adjust. In aggregate, these indirect market spillovers mitigate over three-fourths of the direct agricultural losses driven by urbanization. This leads to a spatial reorganization in which food production moves away from land near urban areas with high emigration potential towards remote areas with low emigration.

Keywords: Internal Migration, Agricultural Development, India
JEL Codes: O13, O15, Q15, Q16, R11, R12, J43

*We thank Werner Antweiler, Marc Bellemare, Francisco Costa, Bruno Esposito, and Carol McAusland for insightful comments. We also thank participants at the UCSD 21CIC Seminar, CREEA, Occasional Workshop, OSWEET, PacDev, AERE, Berkeley ARE, Aix-Marseille School of Economics, Yale Environmental Economics seminar and Paris Dauphine University. Noack thanks the Social Sciences and Humanities Research Council of Canada for their financial support (NFRFE-2018-01353). All errors are our own.

†Department of Applied Economics; rmadhok@umn.edu
‡Food and Resource Economics Group; frederik.noack@ubc.ca
§Department of Economics and Yale School of Management; ahmed.mobarak@yale.edu
¶Department of Economics; olivier@econ.ucsb.edu

1 Introduction

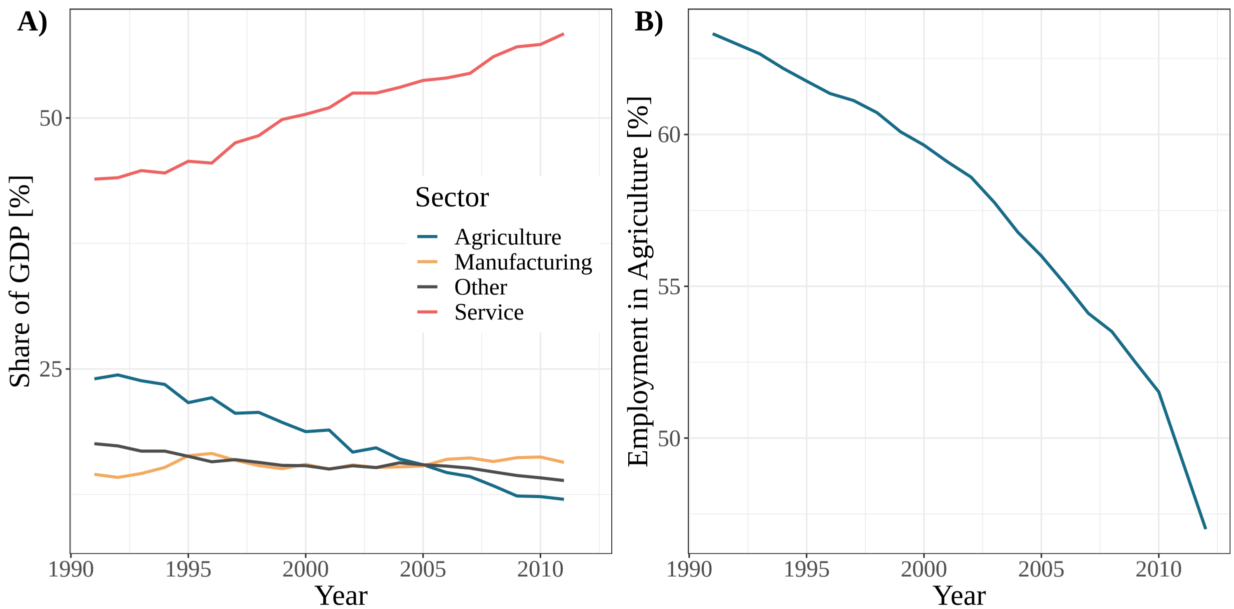
Internal labor migration is perhaps the most pervasive feature of structural transformation. America's rise in the early 20th century was powered by dramatic worker migrations from farms to cities in search of manufacturing and service jobs (Alvarez-Cuadrado and Poschke, 2011). China's 20th-century economic boom witnessed rural workers moving *en masse* to urban factories (Gao et al., 2022). Likewise, starting in 1991, India liberalized its economy to stimulate the service sector (Figure 1A, red). Labor shifted away from agriculture thereafter, dropping by over 30% in the following two decades (Figure 1B). This raises a central question we aim to answer: how does food production cope with such massive labor loss?

Maintaining the same level of food production after labor shifts away from agriculture requires some form of agricultural transformation. A *direct effect* of farming households losing labor through migration could be to substitute workers with capital. The literature documents many examples of labor-capital substitution during structural transformation. The notion that farms modernize in response to labor loss has become somewhat of a stylized fact (Manuelli and Seshadri, 2014; Hornbeck and Naidu, 2014; Alvarez-Cuadrado and Poschke, 2011). Alternatively, as factor markets and prices readjust following labor migration, an *indirect effect* could be for non-migrant farming households to alter their agricultural operations in response to the new factor prices. This indirect channel has been largely overlooked in the literature. We explore both these direct and indirect effects because ignoring such market adjustments could bias estimates of the effects of structural transformation on agricultural output, and potentially lead to misinformed policy.

We answer two questions: (1) what are the direct effects of internal migration on agricultural investment and food production in the largest developing economy as it underwent large-scale transformation? In revisiting this question, we provide credible evidence using a plausibly exogenous shift-share instrument for migration; and (2) what are the indirect effects on farmer investment strategies as crop and land markets readjust following labor migration? To answer this, we develop and validate two new measures of land and crop market adjustments in response to aggregate labor loss. This enables us to empirically test the overall impact of internal migration on agricultural development inclusive of these direct and indirect effects.

Our empirical setting is India between 2005 and 2012, a period of rapid urbanization. Using a panel survey of 42,000 households, we provide three new insights about how agricultural development unfolds during structural transformation. First, contrary to conventional wisdom, we find no evidence that farmers replace workers with capital when facing agricultural labor loss; instead, they reduce crop output and downsize

Figure 1: Economic Growth and Reallocation of Labor (1991-2012)



Note: Panel A reports sectoral shares of India's GDP in 2005 prices using data from the Planning Commission of India. Panel B reports the percent of total employment in agriculture using data from ILOSTAT.

their farms. Second, we develop a spatial model to motivate a regression equation that estimates farmer responses to the land- and crop-market adjustments in a unified empirical framework. The third insight is that indirect effects of labor migration matter: other farmers not directly benefiting from the new migration opportunities expand production in response to higher crop prices and lower land values when migrant households divest from agriculture, and this compensates over three-fourths of food production losses induced by the direct effect. Without accounting for these crop and land market adjustments to aggregate labor loss, we would have incorrectly diagnosed India as becoming food-scarce (or more import-dependent) as structural transformation plays out. Instead, we find that local markets play a quantitatively significant role for spatially redistributing food production.

The first part of the paper develops an empirical strategy for quantifying the direct effects of internal migration on agricultural development. The key empirical challenge is that individual migration is co-determined with origin-based push factors (e.g., poverty). We address this with a shift-share instrument where the "shift" consists of income shocks at potential migration destinations. The "share" consists of the inverse distance to these potential destinations combined with the migration potential of the household. Since mostly men migrate for work in India, households with working-age men have more potential to react to destination wage changes. In our main specification, we therefore measure household migration potential by the number of working-age males (i.e., "potential migrants") at home during the baseline period. Combining both share variables allows us to compare households with the same migration potential exposed to close and far income shocks as well as households exposed to the same income shocks but with high and low migration potential.

We find no evidence of labor-capital substitution in response to migration through the direct channel. Instead, migrant-sending households invest *less* in agrochemicals, irrigation water, and work animals. This, in turn, drives a reduction in crop output and farm size. Households losing one additional migrant cultivate 1.7 fewer hectares of land. Downsizing is mainly driven by larger farms becoming smaller, whereas smaller and medium-sized farms remain similarly sized. This is consistent with the notion that labor constraints increase in farm size in India (Foster and Rosenzweig, 2022).

The second part of the paper investigates equilibrium effects on agricultural investment and food production, as markets react to labor reallocation. To set the stage, we map aggregate changes in agricultural labor, crop output, and economic growth during the study period. The visual evidence yields an interesting spatial pattern: while rural areas near urban centers experience labor loss and output decline, remote areas experience agricultural booms (Figure 4). We hypothesize that this spatial pattern is explained by crop and land market spillovers. Declining output and farm size through the direct channel trigger higher crop prices and lower land prices, respectively. Households in remote areas, where emigration is low, benefit from these indirect market spillovers by expanding farm activity. Together, this leads to a spatial reorganization of agriculture from migrant-sending areas toward remote areas.

We formalize this idea with a spatial model of migration and household production (Appendix C) designed to guide our empirical analysis. The key theoretical insight is that the direct effects of labor loss and the indirect effects through markets can be conveniently estimated via a single additive estimating equation. Our approach draws on Adao et al. (2019) and Borusyak et al. (2022b), who develop this additive reduced form estimation equation from a theoretical gravity framework.[1] The net effect of migration on agriculture is theoretically ambiguous and depends on which channels dominate.

Based on these insights, we expand our empirical framework to quantify indirect effects. We assume land markets operate at the village level: when other farming households divest from agriculture as their sons emigrate in response to urban opportunities, local land prices fall, which enables non-migrant households in the village to expand production. Since India restricts cross-state agricultural trade, we conceive crop markets as areas *within a state* that are equally suitable for growing a fixed set of crops. The crop mar-

[1] See Huber (2023) for additional support for this empirical framework.

ket adjustment is measured as aggregate emigration among other households in those crop markets defined using gridded crop suitability indices. Intuitively, if urban shocks trigger mass emigration from a rice-suitable region, thereby reducing aggregate rice output and increasing rice prices, then non-migrant households elsewhere in the state will expand rice production if they also live in a rice-suitable region.

In this expanded framework, the direct effect remains instrumented with the shift-share variable whereas indirect crop and land market measures are un-instrumented. We make this choice mainly because spatially aggregated instruments (e.g., at the land and crop market level) are less credible, since more unobserved confounding differences arise when comparing across distant geographic units. Instead, we design placebo tests to demonstrate the validity of our estimates of indirect effects. To validate our land market measure, we show that household farm size does not respond to aggregate emigration by *non-agricultural* households in the village, which suggests that our measure is not picking up other non-land factor market changes. To validate our crop market measure, we show that households do not respond to aggregate emigration from outside the crop market, nor from areas growing unrelated crops.

We find that farmers increase production in response to crop market adjustments, partially offsetting the migration-induced output loss from the direct channel. We build a crop price index from ICRISAT to test mechanisms and find that mass emigration from the crop market increases crop prices, in line with our model. The price increase leads farming households with few or no migrants to spend more on agricultural technologies, such as agrochemicals and equipment rentals, ultimately leading to higher production.

We also find that other farmers in the village who are less exposed to the migration shock expand their farms in response to land market adjustments, again counteracting the downsizing of farms among migrant-sending households. We use data on household land transactions from IHDS to test the mechanisms. As expected, aggregate emigration from the land market (village) depresses land prices. "Left-behind" farmers in the village exploit this opportunity and expand cultivation.

These results embed an important spatial component: the (negative) direct effect of migration on crop output dominates for households near cities that face low migration costs. The (positive) market-driven effects dominate for households who live further away and face high migration costs. Although these remote households do not participate directly in structural transformation, they still experience production and technology benefits through indirect market and price spillovers triggered by those who do migrate.

More broadly, our findings add important and policy-relevant value to our understanding of structural transformation. Labor reallocation *does* drive agricultural development, but unlike received wisdom, not through capital substitution or other direct responses to labor loss. Instead, agricultural development spatially shifts through market forces, in the form of increased technology adoption and production among remote, non-migrant households. This implies that structural transformation can promote income redistribution toward those who do not directly participate in it.

The paper concludes with a simple accounting exercise based on our estimates to determine how much of the migration-induced agricultural losses are mitigated by the crop and land market adjustments. We consider special cases of the empirical specification and predict aggregate crop value in the absence of migration, with migration but no market adjustments, with migration and land market but no crop market adjustments, and so on. We find that market adjustments mitigated 80% of aggregate agricultural production losses due to emigration. These results do not imply that food production declined in India during this period of emigration and structural adjustment, only that we are able to quantify the partial effects of these two margins of adjustment.

1.1 Literature Contributions

This paper contributes to both the micro- and macro-development literature on structural transformation and its impact on the agricultural sector. We discuss each in turn.

Micro-Development Contributions First, we contribute to the micro-development literature on the direct effects of labor loss on agriculture. Prior studies use cross-sectional data from several villages to show that individual migration reduces crop income in China (Rozelle et al., 1999; Taylor et al., 2003) and Bangladesh (Mendola, 2008). We advance this work with household panel data across India and an instrument for migration. In Uganda, Brewer et al. (2022) use a shift-share instrument to document labor-capital substitution among farmers, the opposite of our findings. We advance this work by considering spatial spillovers.

Second, we are able to decompose agricultural responses to labor reallocation into direct (labor-driven) and indirect (market-driven) channels. In doing so, we advance the micro-development literature on structural transformation that uses data at coarser spatial scales. Hornbeck and Naidu (2014) and Manuelli and Seshadri (2014) also study migration and agricultural development, but with county-level data from the early 1900s in the United States. Unlike our findings, they document a process of agricultural modernization whereby counties facing emigration shocks adopted new labor-replacing technologies. Similarly, Clemens et al. (2018) find that reduced agricultural immigration spurs adoption of labor-saving technologies. In these studies, technology adoption may be a direct response to labor loss or an indirect response to changing sectoral compositions across the broader economy. By quantifying direct and indirect effects together, our micro data also allow us to paint a richer picture of the post-migration adjustment process.[2]

Our third contribution is to characterize direct and indirect effects from structural transformation in a single empirical framework. Previous micro-development studies have done this at small spatial scales (Blakeslee et al., 2023) or studied spillovers at different geographies separately (Asher et al., 2023). Our approach unifies this literature by identifying the spatial incidence of direct and indirect effects of structural change together in one framework. This enables us to compare the two forces and estimate the volume of agricultural losses compensated through spatial market-driven spillovers.

Macro-Development Contributions This paper is unique in that it uses household microdata and applied microeconomic techniques to elaborate long-standing issues in macro-development.[3] Our analysis serves as an empirical test of classic models of structural transformation. Lewis (1954) and Harris and Todaro (1970) present models where urban capital accumulation raises wages, attracts rural migrants, and drives structural transformation. Recent work shows that this process increases agricultural productivity as labor moves to relatively higher-productivity sectors (Gollin et al., 2014; Duarte and Restuccia, 2010; Lagakos and Waugh, 2013; Restuccia et al., 2008; Vollrath, 2009). We further advance this literature by using spatially explicit household data to empirically identify the micro-mechanisms—including farmer-level labor-capital substitution and changes in local land and crop prices—underlying the aggregate agricultural impacts of structural transformation.

Our paper is especially related to Kaboski et al. (2024), who study direct and indirect effects of structural change on firms in India. They posit a direct channel, whereby lower trade costs leads to fewer firms, and an indirect channel whereby cheaper agricultural imports releases agricultural labor leading to more firms. Whereas we also find that direct and indirect effects move in opposite directions, we study impacts on agriculture rather than firms. Our paper is also related to Garriga et al. (2023), who study direct and indirect effects of structural transformation on housing in China.

Methodologically, we connect to a growing macro-development literature on spatial general equilibrium estimation. We build on Adao et al. (2019), which develops a reduced form framework for estimating spatial equilibrium effects of economic shocks.

[2] A related literature studies the effect of agricultural innovation on labor migration and finds that improvements in agricultural technology is labor-saving and leads to industrial growth (Caprettini and Voth, 2020; Bustos et al., 2016; Caunedo and Kala, 2021; Emerick, 2018; Moscona, 2019). We extend this literature by studying the reverse scenario and also by documenting market spillovers within the same sector.

[3] Lagakos and Waugh (2013) provide an overview of the importance of micro data for deepening our understanding of macro-development structural transformation.

The next section describes the data and presents stylized facts about migration and agriculture in India. Section 3 outlines the shift-share research design and Section 4 presents estimates of the direct effects of migration on agriculture. Section 5 estimates indirect effects in response to market adjustments. Section 6 incorporates both estimates in an accounting exercise that calculates net changes in food supply. Section 7 concludes.

2 Data and Stylized Facts

We estimate the impact of internal migration on agricultural development using detailed household panel data. This section describes the data and presents four stylized facts about migration and agriculture. These facts motivate the empirical strategy in Section 3.

2.1 IHDS Household Panel

Highly disaggregated household-level migration data are sparse in India. An exception is the two-wave IHDS panel, which is nationally representative and covers 384 districts (out of 594 at the time of the survey) across all states[4]. Wave I (2004-05) surveyed 41,554 households, of which 83% were located again in Wave II (2011-12).

There are at least three advantages of IHDS. First, it is among the few Indian surveys documenting both labor mobility and agricultural production. Second, the same households are interviewed twice, enabling the inclusion of household fixed effects to control for time-invariant unobserved heterogeneity across households, such as caste or baseline poverty. Lastly, IHDS disaggregates income into several categories, including agricultural income, according to a standardized procedure.

Yet IHDS is not without limitations. First, there is a 17% attrition rate. While attrition is common in household panels, it poses minimal concern for us since 80% of dropouts were landless in Wave I and would have been excluded anyway. Second, seasonal migration is not reported in Wave II, restricting our analysis to medium-term migration.

2.2 Main Variables

2.2.1 Migration

IHDS defines household residents as those living together for at least 6 months of the past year. It defines migrants as the opposite: household members who have lived elsewhere *for over 6 months of the past year*. Residents report age, sex, and location of migrant family

[4]Except Andaman and Nicobar Islands and Lakshadweep, which contain $< 1\%$ of the population.

members. The IHDS definition thus characterizes longer-term spells[5], such as a son who has been away for the past five years. If he instead returned after five months, he would be a household resident, not a migrant. We exclude international migrants, which represent $< 10\%$ of migrants, as our paper focuses on internal migration.

The main explanatory variable, and key migrant demographic, are working-age males between age 15 to 60. To justify this choice of age window, Table A1 shows migrant types by age group. The share of migrants that leave for education drops after age 14, the same age at which employed sons jumps five-fold. This suggests that migrant males transition from school to work around age 15. Similarly, the share of husband migrants sharply drops after age 59, suggesting they return home around this age.

2.2.2 Agricultural Activity

Farming households report agricultural capital, labor, and crop income. The survey divides capital into input expenses, which include seeds, fertilizer, pesticides, irrigation water, and hired animals, as well as machinery, which includes tubewells, electric/diesel pumps, bullock carts, tractors, and threshers. There is also a labor expense sheet that documents wages paid for hired labor and person-days of unpaid family labor in the past year. We deflate all expenses to 2005 prices using the rural or urban Consumer Price Index, depending on household location[6].

Since capital and machinery are each described by five variables, we collapse them into two indices to allay concerns of multiple hypothesis testing. We follow Anderson (2008) whereby each capital or machinery variable influences its index proportional to the information it adds. Intuitively, if seed expenses are highly correlated with other input expenses, then it adds little to the expense index. To operationalize this idea, we first demean the five capital and five machinery variables and divide them by the standard deviation of non-migrant households. This converts each variable to a unitless measure that can be easily aggregated. The capital index is then computed as a weighted sum of the five standardized capital variables with weights equal to the row sum of the inverse covariance matrix. The expense index is computed in the same way.

Crop production is measured as crop income in 2005 prices[7]. Crop income is defined as crop revenue minus expenses, where revenue is price $\times$ quantity. For crops sold at market, farmers report quantity sold and sales price. For crops grown for personal consumption, they report quantity and the price they *would* have received at market.

[5] Short-term (<6 months) circular migration is only documented in Wave II.

[6] The price deflator is a pre-constructed variable distributed by IHDS.

[7] Five percent of households report negative farm income, which we recode as missing

2.2.3 Droughts

Droughts are a key covariate in our analysis. Accounting for droughts is crucial because they affect both labor movement and agricultural decisions. We measure drought intensity using the gridded (0.5° resolution) Standardized Precipitation-Evapotranspiration Index (SPEI), which measures the difference between potential evapotranspiration and precipitation. We extract the mean over cells within districts and then compute annual averages for 2005 and 2012 to match with the IHDS.

2.3 Summary Statistics

The sample frame comprises 40,018 households interviewed in both periods. Nine percent of households ($N = 3747$) had a migrant in Wave I, and 23% ($N = 9112$) had a migrant in Wave II. Table A2 profiles the typical migrant: 80% are male, and nearly 60% are labor migrants in the first period. The labor migration rate jumps to 71% in the second period. This supports our focus on working-age males as the key demographic. The bulk of remaining migrants leave for education. Our working age window of 15-60 excludes the majority of these student migrants (Table A1).

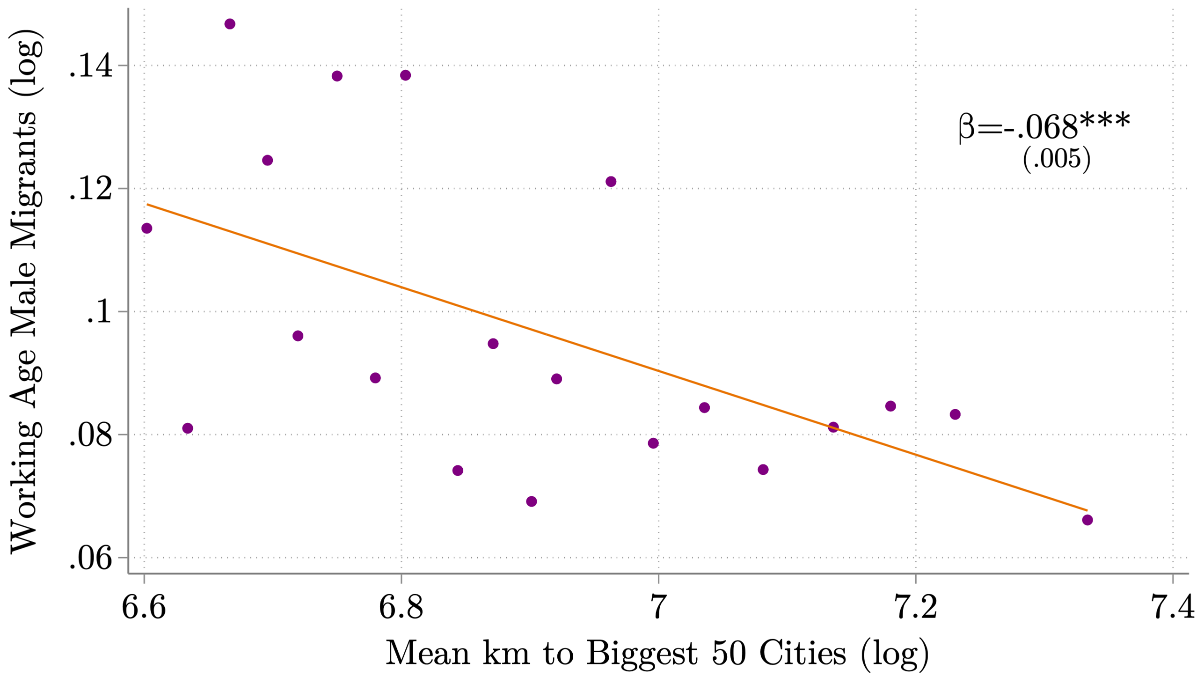
Most migration is within-state (Table A2 Panel C). Interestingly, rural-rural migration accounts for a large share of migration in both waves. Figure B1 splits migration streams by inter- and intra-state travel, revealing an interesting fact: among inter-state migrants, rural-urban migrants dominate rural-rural by 7-to-1, whereas rural-rural migration dominates among intra-state migrants. We, therefore, include both rural and urban destinations in the choice set when building our instrument for migration in Section 3.2.

2.4 Four Stylized Facts About Migration and Agriculture

Next, we describe four stylized facts from the data that motivate our empirical strategy. The first fact is that entire-family migration is rare. Table A3 describes family migration between surveys, defined as households surveyed in Wave II (2012) that reported having moved after Wave I (2005). Only 1% of households moved as a family and, among them, 85% migrated within the district. We, therefore, treat migration as a continuous variable (number of migrants sent) rather than a binary decision (move or stay) in the analysis.

The second fact is that remote households send fewer migrants. Figure 2 shows a binscatter plot of district remoteness, measured as mean distance to the 50 biggest cities, against number of household migrants, residualized on year fixed effects. The negative slope is the regression coefficient. Since both axes are on log scale, changes can be interpreted in percentages: households twice as remote (100 percent increase) send 7% fewer

Figure 2: Binscatter Plot of Migration and Distance to Nearest Large Cities



Note: Y-axis is log number of working age male migrants in a household plus one, from IHDS. X-axis is log of mean distance from each district to the 50 biggest districts by population. Year fixed effects are partialled out. Orange line is the best linear fit, constructed from an OLS regression of y-residuals on x-residuals.

migrants. This provides basic evidence that migration costs increase with distance. We incorporate this fact into our instrument for migration, which leverages household exposure to destination income shocks where exposure declines with distance.

The third fact is that land markets are thin, yet still active in India. Table 1 summarizes land, labor, and capital among farming households. Farming is small-scale, with the typical household cultivating about three acres (Panel A). The share of rented land increases from 8% to 9% between surveys, indicating that land market transactions, including rentals, are relatively uncommon. Cultivation yields an annual income of approximately 8000 Rs./acre across both surveys. Fertilizer is the largest expense (Panel B), amounting to 10% of income on average. Water pumps are relatively common, while large equipment like tractors are rare (Panel C).

The fourth fact is that rural labor markets are malfunctioning. Family contributes over 15 times more person-days of farm labor than hired workers (Panel D). In fact, only 6% of total farm labor is hired. This is evidence of malfunctioning rural labor markets in India, in line with the literature (Fernando, 2022; Foster and Rosenzweig, 2022).

Table 1: Summary Statistics of Land, Labor, and Capital

	IHDS-I (2004-05)		IHDS-II (2011-12)	
	Mean	SD	Mean	SD
A: Land				
Area cultivated (ac.)	3.04	3.40	3.62	3.78
Area rented in (ac.)	0.24	0.69	0.32	0.86
Rental price (Rs./ac./yr)	3914.57	16730.12	3105.50	11927.10
Yield (Rs./ac.)	8640.49	26863.14	6815.23	24092.13
B: Expenses/Acre				
Seeds	712.46	1246.18	625.02	3083.32
Fertilizer	1012.31	2342.70	850.17	4263.19
Pesticides	264.07	916.97	247.34	1922.83
C: Machinery (Num./Acre)				
Pumps	0.17	1.05	0.11	0.80
Tractors	0.02	0.14	0.01	0.11
Bullock Carts	0.07	0.27	0.04	0.21
D: Labour (past yr.)				
Family person-days/acre	253.02	596.90	193.38	630.44
Hired person-days/acre	15.52	40.22	10.44	78.55
Wages Paid/Acre	773.92	2168.93	767.53	5920.71

Note: Data are a household-panel for land-owning households. Pumps include electric and diesel water pumps. Wages refers to total paid labor wage bill in the past year. All monetary values are in 2005 prices.

3 Empirical Framework

This section develops an empirical strategy to estimate the *direct* effects of internal migration on agricultural development. Direct effects materialize through a labor-capital substitution channel, where crop and land prices are fixed. We develop a shift-share instrument for migration based on exposure to destination income shocks, and test instrument credibility through various validation exercises. Section 5 extends this framework to investigate the *indirect* effects of migration through land and crop market adjustments.

3.1 Baseline OLS Equation

The relationship between agricultural outcomes and migration can be written as:

$$Y_{idt} = \beta_1 M_{idt}^{labor} + \Gamma X'_{dt} + \alpha_i + \gamma_t + \varepsilon_{idt}$$

where Y_{idt} are agricultural outcomes for household i in district d at time t (e.g., crop income). M_{idt}^{labor} is the number of working-age male migrants sent from household i. X'_{dt} is a vector of covariates that jointly influence migration incentives and agricultural outcomes, such as drought conditions. Household fixed effects, α_i, absorb time-invariant differences between households, such as distance to cities or land quality, both which may affect migration and agricultural decisions. Demand-side effects from increased urban productivity are captured by year fixed effects, γ_t.

β_1 captures households' direct response to labor loss, and is likely biased when estimated via OLS due to the endogeneity of M_{idt}^{labor}. Although α_i absorbs baseline wealth, changes in income could jointly influence migration and agricultural decisions. Improvements in agricultural technology can also release surplus labor, leading to reverse causality. We next introduce a shift-share instrument to mitigate these threats to identification.

3.2 Shift-Share Instrument Design

3.2.1 Measurement

To identify β_1, we construct a shift-share instrument for M_{idt}^{labor} that combines income shocks at each potential destination (the shift) with measures of household exposure to the shocks (the share). When combined, the shift and share yield an instrument that isolates the pull stream of migration that is plausibly exogenous to origin push factors.

Income shocks, $inc_{d't}$, the "shift" of the shift-share design, are measured by mean income of households in each destination district $d' \in \Theta/d$ at time t, where Θ is the set of all districts. We interact this with district population from the 2001 Census, $pop_{d'}$, to incorporate the fact that migration propensity increases toward urban destinations (Table A2).

Unlike previous studies, we use two "shares" that measure exposure to the shift. The first reflects the fact that potential migrants are more exposed to nearby shocks (Fact 2, Section 2.4). We measure this force by the inverse distance from each origin district d to every potential destination district $d' \in \Theta/d$, denoted as $\frac{1}{\tau_{dd'}}$[8], where τ is distance. Since Θ spans all districts, households also consider rural destinations in their choice set, in line with the observed prevalence of rural-rural migration in India (Table A2). We use a distance elasticity of one based on similar values from the literature (Bryan and Morten,

[8]Distances are measured by kilometers between district centroids based on 2001 Census shapefiles.

2019; Schwartz, 1973). The second share proxies urban productivity of the household, φ_i. We measure this as the number of baseline working-age males living in household i, which determines household i's *potential* to benefit from destination wage increases.

Finally, we combine the shift and the share to form an instrument for migration, z_{idt}:

$$z_{idt} = \underbrace{\varphi_i}_{\text{productivity (share I)}} \times \underbrace{\sum_{d' \in \Theta / d} \frac{1}{\tau_{dd'}}}_{\text{inverse distance (share II)}} \times \underbrace{(inc_{d't} \cdot pop_{d'})}_{\text{shift}} \tag{1}$$

3.2.2 Instrument Validity

z_{idt} is a valid instrument if it strongly predicts labor migration, M^{labor}_{idt}, and fulfills the exclusion restriction. We test the first criteria directly with the first stage equation (Section 3.3). The exclusion restriction is that z_{idt} affects agricultural outcomes *only* through labor migration, conditional on fixed effects and controls. While we cannot test this directly, we turn to a discussion of potential violations through each component of z_{idt}.

The first concern is that φ_i can affect agricultural outcomes directly if households with more working-age males at baseline have more income, land or other production factors. Household fixed effects address this. The second concern is that households close to economic centers (formally, those with high values of $s_{dt} := \sum_{d' \in \Theta / d} \frac{inc_{d't} \cdot pop_{d'}}{\tau_{dd'}}$), may face different input and output markets, opening other "backdoor" channels through which z_{idt} can affect household agricultural decisions. We address this by including s_{dt} directly in both the first and second stage regressions, leaving identification to rely on differences in household exposure to the shock, conditional on the shock itself. Third, urban income shocks, captured by $inc_{d't} \times pop_{d'}$, can increase aggregate food demand and affect household crop production through output prices. Year fixed effects help address this.

Despite these remedies, the main remaining concern is that even after controlling for household fixed effects, year fixed effects, and s_{dt}, φ_i may still be endogenous if household composition differentially predicts *changes* in agricultural outcomes. For example, if households with more working-age males are more exposed to the shock, and also experience different agricultural productivity trends, then the exclusion restriction is violated.

We address this concern in two ways. First, we rely on the recent literature showing that shift-share instruments with endogenous shares are valid if the shock is as-good-as-randomly assigned (Borusyak et al., 2022a). Figure B2A shows a histogram of the shock, s_{dt}, across households with high and low values of φ_i. The distributions are nearly identical, especially when residualizing on household and year fixed effects (Panel B). This means that both groups are similarly exposed to the shock, suggesting that the shock is unlikely to pick up correlated characteristics of more and less exposed households.

Table 2: Shock Balance Tests

	(1) Males (φ_i)	(2) Educated	(3) Farm Size	(4) Ag. HH	(5) Landowner
Δ Wt. Income	0.086	0.219***	-0.356	0.034	-0.076
	(0.076)	(0.085)	(0.392)	(0.040)	(0.052)
Δ Origin Income	Yes	Yes	Yes	Yes	Yes
State FEs	✓	✓	✓	✓	✓
Observations	38589	38589	18213	38589	38588
R^2	0.021	0.031	0.147	0.049	0.086

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Data are a cross section of households in 2005. The explanatory variable is the change in the inverse distance-weighted income shock between 2005 and 2012. The outcome in Column 1 is the number of resident working age males. Column 2 is number of household members with at least secondary education. Column 3 is farm size in acres. Column 4 an indicator for primary income coming from agriculture. Column 5 is a dummy for land ownership. Standard errors clustered by PSU.

Second, we conduct formal falsifications tests by regressing a set of baseline confounders, X_{id} on the *change* in the shock, $\Delta s_d := s_{dt} - s_{dt-1}$, in a pooled cross-section:

$$X_{id} = \phi \cdot \Delta s_d + \Gamma \cdot \Delta inc_d + \gamma_s + \varepsilon_{ijd}. \tag{2}$$

where ϕ is the balance coefficient. If Δs_d is as-good-as-randomly assigned, it should not predict X_{id} and $\phi = 0$. Δinc_d is the change in origin district income, which accounts for demand effects, and γ_s is a state fixed effect. Table 2 reports estimates of ϕ from regressions with number of working age males (φ_i), education, farm size, sector, and land ownership as outcomes. The key result is that the shock is uncorrelated with the share (column 1). We control for education in a robustness check since the shock is disproportionately felt by more educated households, who may experience different outcome trends even in the absence of migration. Otherwise, all other estimates are consistent with the as-good-as-random hypothesis.

3.3 Two-Stage Least Squares

Equipped with our shift-share instrument, z_{idt}, and having shown its validity, we specify the effect of labor migration on agricultural outcomes in a 2SLS framework as follows:

$$M_{idt}^{labor} = \mu_1 z_{idt} + \mu_2 s_{dt} + \mu_3 inc_{dt} + \Gamma X'_{dt} + \alpha_i + \gamma_t + \varepsilon_{ijdt} \tag{3}$$

$$Y_{idt} = \beta_1 \widehat{M_{idt}^{labor}} + \beta_2 s_{dt} + \beta_3 inc_{dt} + \Gamma X'_{dt} + \alpha_i + \gamma_t + \eta_{ijdt} \tag{4}$$

where $s_{dt} := \sum_{d' \in \Theta/d} \frac{inc_{d't} \times pop_{d'}}{\tau_{dd'}}$, the distance-weighted income shock. Equation 3 is the first stage regression, which relates labor outflows from the origin, M_{idt}^{labor}, to the instrument, z_{idt}, controlling for income shock itself, s_{dt}. We control for per-capita incomes in the origin district, inc_{dt}, to account for spatial correlation between destination and origin income shocks. α_i and γ_t are household and year fixed effects, respectively. To the extent that z_{idt} is plausibly exogenous (Section 3.2.2) conditional on fixed effects and controls, $\widehat{M_{idt}^{labor}}$ represents the pull stream of migration that is orthogonal to push incentives at the origin, and μ_1 identifies the labor response to destination income shocks.

Equation 4 estimates the second stage impact of labor migration on agricultural activity, Y_{idt}, of household i in district d at time t. The main outcome variables are crop incomes, technology adoption, and farm size. The coefficient of interest is β_1, which captures the direct effect of emigration on agricultural development through the labor channel. $\beta_1 > 0$ indicates that households respond to labor loss by adopting better technology and and increasing production, in line with the idea of labor-capital substitution. $\beta_1 < 0$ implies that households divest from technology and/or downsize their operation.

Since s_{dt} is a component of z_{idt}, and also included directly as a covariate, β_1 is identified off of differences in household i's exposure to *changes* in destination incomes, where exposure declines with distance and increases in the number of resident working-age males. Inclusion of origin district income, inc_{dt}, as a covariate ensures that identifying variation captures the bilatural nature of migration decisions (Borusyak et al., 2022b). In particular, identifying variation captures migrant responses to *differences* between origin and potential destination incomes.

4 Results: The Direct Effect of Migration on Agriculture

We now present evidence on the direct effect of labor emigration on agriculture. In contrast to the idea of labor-capital substitution, we find that Indian households do not replace labor with capital. Instead, they *reduce* technology and output in response to labor migration. Although crop and land prices are assumed to be unaffected by labor migration here, Section 5.3 extends the analysis to enable market adjustments.

4.1 Main Estimates

Table A4 presents first stage estimates (Equation 3). Column 1 excludes controls, column 2 controls for the direct shock, s_{dt}, and column 3 controls for both s_{dt} and origin income, inc_{dt}. The instrument strongly predicts labor outflows across all specifications. Column 3 is the preferred specification. To interpret the coefficient, note that the average household

Table 3: Second Stage Estimates—Direct Effect of Migration on Agricultural Development

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Man-days	(6) Crops
Male Migrants (σ)	-1.046***	-0.709***	-1.218***	0.076	-1.877***	-1.347***
	(0.205)	(0.182)	(0.229)	(0.158)	(0.311)	(0.224)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes	Yes
Outcome SD	-	-	3.621	150.176	211.329	21071.048
Explanatory SD	0.524	0.521	0.533	0.518	0.518	0.513
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25928	24970	29346	20910	20910	25854
F-Stat	58.9	55.1	53.3	55.4	55.4	63.4

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The explanatory variable is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. All outcomes are in standard deviations. All specifications control for district drought conditions. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.2). Wage bill (column 4) is total wages paid to all workers in the past year. Person-days (column 5) includes both household and hired labor. Crop profits (column 6) are net of expenses. Standard errors clustered by PSU.

has 1.76 male working-age residents. Therefore, a 1σ increase in destination incomes pulls 22% (=0.383/1.76) of them away to join the destination labor force. F-statistics are well above rule-of-thumb levels in tests for weak instruments (Kleibergen and Paap, 2006).

Table 3 presents second stage estimates of Equation 4. All variables are reported in standard deviations to facilitate comparisons of coefficient magnitudes across outcomes. The coefficient of interest is negative and statistically significant for technology, farm size, and crop income, suggesting that labor loss prompts agricultural decline, not modernization. The point estimate in column 1 implies that a 1σ reduction in farm labor causes households to reduce agricultural expenses by 1.05σ. When decomposing the index by individual expenses (Table A5, columns 1-5), the effect is driven by lower spending on seeds, agrochemicals, irrigation water, and rented equipment. Column 2 of Table 3 shows that labor loss also causes households to reduce machinery stock. This is driven by divestment of tubewells, water pumps, bull carts, and threshers (Table A5, columns 6-10).

Columns 3-6 of Table 3 explore additional margins of response. Column 3 implies that labor loss causes households to downsize cultivated area by 1.22σ, equivalent to about four acres. This is consistent with the idea that declining labor and technology use reduces the marginal productivity of land, which in turn prompts farm contraction (Appendix C). As we show in the next subsection, the large coefficient magnitude is driven primarily by large farms downsizing, whereas small and medium farms remain similar size.

In contrast to land markets, labor markets respond imperfectly. There is no impact on wages (column 4), likely because the loss of one household laborer is insufficient to affect equilibrium wages. Column 5 shows a decline in total person-days worked. If the migrant were replaced with family labor, or from the labor market, there would be no effect. The negative effect on person-days is therefore indicative that households are labor-constrained and unable to replace lost labor, consistent with evidence of India's malfunctioning rural labor markets (Fact IV, Section 2).

Lastly, declining inputs (columns 1-5) in response to labor migration leads to a contraction of output (column 6). The point estimate implies that 1σ of labor loss leads to a 1.35σ decline in agricultural profits over the past year, corresponding to about Rs. 28,000.

4.2 Estimates by Farm Size

The fact that farms do not compensate family labor loss with hired labor is unsurprising given imperfect rural labor markets in India (Fact 4, Section 2.4). It also supports the idea that the results in Table 3 are a direct response to labor loss. A more formal test of the labor channel would compare the responses of labor-constrained farms, where farmers are sensitive to labor loss, to surplus-labor farms, where marginal workers are unproductive and their loss has little effect on farmer decision-making.

Our setting is well-suited for this test since small farms in India are known to have surplus labor (Foster and Rosenzweig, 2022). Table A6 shows labor-to-land ratios for small (0-2 acres), medium (2-4 acres), and large farms (4+ acres) in our sample. The labor-to-land ratio for family labor is five times larger for small farms compared to large farms. It is nearly two times larger for hired labor. We therefore expect farmer responses to emigration to increase in farm size as labor constrains become tighter.

Figure 3 estimates Equation 4 separately by farm size. As expected, the main results appear to be driven primarily by larger farms (green). Wide confidence intervals are due to fewer large farms in the sample. While large farms consistently respond more to emigration than the pooled sample, the response of small and medium farms (blue, red) to emigration is much smaller, often statistically indistinguishable from zero. Since emigration from these labor-unconstrained farms has little effect on agricultural outcomes, these results indicate that our baseline findings in Table 3 are likely driven by larger, labor-

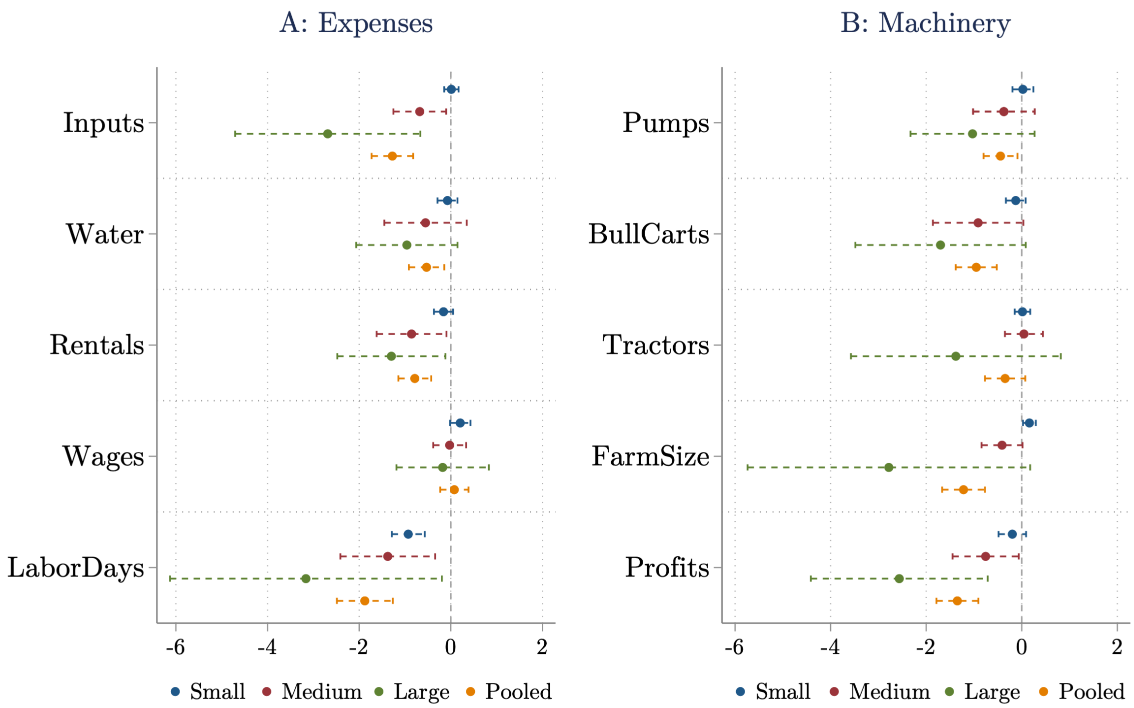


Figure 3: Second Stage Estimates by Farm Size

Note: The explanatory variable is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. Outcome variables are on the y-axis. Blue, red, green, and orange coefficients are for small (0-2 ac.), medium (2-4 ac.), large (4+ ac.),and pooled farm sizes, respectively. All outcomes are in standard deviations. All specifications control for district drought conditions. In Panel A, inputes, water, rental and wages are in Rs./year. In panel B, pumps, bull carts, and tractors are quantities owned. Profits are in Rs./year. Standard errors clustered at village level.

constrained farmers directly responding to labor loss, as opposed to another channel.

4.3 Robustness Checks

We now probe the sensitivity of our estimates of direct effects to a variety of robustness tests. After establishing robustness, we turn to an exploration of indirect effects.

Controlling for Observables: Panel A of Table A7 tests robustness to using state-by-year instead of year fixed effects. State-year fixed effects account for unobserved factors that are constant across households within a state-year that jointly determine migration propensity and agricultural outcomes, such as state GDP or state-level agricultural policies. Results are very similar to the baseline estimates.

Panel B adds a control for household education, since we found evidence in the instrument validity test (Table 2) that educated households are disproportionately exposed to destination income shocks. Estimates remain stable, improving our confidence that the instrument is not picking up differential trends experienced by educated households. Panel C controls for the baseline shock interacted with year fixed effects. This accounts for trends in the pre-period shock, which is a concern if the trend in agricultural outcomes across households in areas with high baseline shock severity differs from that of households in areas with low baseline shock severity, regardless of whether they have resident working age males. The resulting estimates are again very similar to the baseline ones.

Extensive Margin: Panel D tests robustness to including the extensive margin in the sample—households that left or joined agriculture between the two survey waves. To do so, we zero-impute outcome variables for landless households. This adds to the sample landless households in period 1 who owned land in period 2 as well as households who owned land in period 1 and then left agriculture in period 2. The main results are robust to this alternative sample.

Alternative Shifts and Shares: Next, we show that our estimates are robust to alternative shift-share instruments. First, we explore alternative specifications of the "shift". Panel E redefines z_{idt} to isolate rural-urban migration only. To do so, we let the destination choice set (Θ in Equation 1), consist of urban destinations only, and restrict the estimation sample to rural households. The results are virtually unchanged, suggesting that our baseline estimates are not driven by any particular migration stream but rather reflect internal migration in general. In Panel F, we use inverse-distance weighted nightlights as the "shift"[9]. Again, results remain stable, suggesting that our baseline estimates are robust to different measures of the shift.

Second, we explore alternative measures of the "share". Panel G tests robustness to redefining the urban productivity parameter, φ_i, as the number of working age male *or female* household members. M_{idt}^{labor}, the number of migrants sent from household i in the first stage, also covers both genders. The results remain similar, which is unsurprising since women make up less than 10% of migrants (Table A2).

Inference: Lastly, Table A8 reports estimates from alternative methods of statistical inference. The baseline estimates are clustered at the PSU level, a fundamental geographic unit defined by IHDS for the initial stage of sampling, and within which households likely share unobserved characteristics. Since the shock is at the district level, we also report standard errors clustered by district to account for correlation between shocks across

[9]Satellite-detected nightlights are considered a strong proxy for local GDP (Henderson et al., 2012). Data are obtained from the DMSP-OLS satellite product at 30 arc second resolution.

households in the same district. We also cluster by state in case shocks are correlated across districts within the same state. Lastly, we investigate spatial correlation more systematically by estimating Hsiang (2010)'s implementation of Conley (1999) standard errors for kernel cut-off distances ranging from 200km to 500km[10]. Overall, the statistical significance of our estimates of the direct effct of migration on agricultural outcomes is robust to these alternative methods of accounting for spatial and serial correlation.

5 The Indirect Effect of Internal Migration on Agriculture

5.1 Conceptual Framework

We have shown that Indian households do not substitute labor with capital when facing labor loss. Instead, they reduce agricultural technology adoption, downsize their farms, and produce less food. Given the wide-sweeping urbanization in India (Figure 1), this implies a potentially concerning reduction in domestic food supply compared to a counterfactual world with no urbanization. Of course, we have thus far only studied *direct* effects of internal migration through a labor channel. The equilibrium effect on food production requires understanding how land and crop markets adjust following labor migration, and, how farmers respond *indirectly* to these market channels. These insights are important for understanding how much of the migration-induced food shortage, if any, is compensated through markets.

To build intuition about the nature of these indirect effects, first consider crop markets. If the decline in food supply through the direct channel (Column 6, Table 3) increases crop prices, then we expect increases in crop production, especially among households with no migrants who do not experience direct effects but still benefit from higher crop prices. Since these non-migrant households are more remote (Fact 2, Section 2.4), we expect production booms in remote areas. Similarly, if the decline in farm size through the direct channel depresses land prices, we may also see agricultural expansion through the indirect land market channel.

Figure 4 visually illustrates such indirect responses. Panels A, B, and C map the change in agricultural labor, crop output, and nightlight intensity, respectively, between 2001-2011. Southern, Eastern, and Northern districts experiencing labor loss (Panel A, red) also experience output contraction (Panel B, red). A local analysis would thus imply that emigration leads to agricultural decline, as we saw in the previous section. However, the aggregate view reveals compensating production increases in Central and Western India (Panel B, blue). Economic activity is a key driver of this spatial redistribution: whereas

[10] We adapt the Hsiang (2010) implementation to an IV setting using the method of Colella et al. (2019).

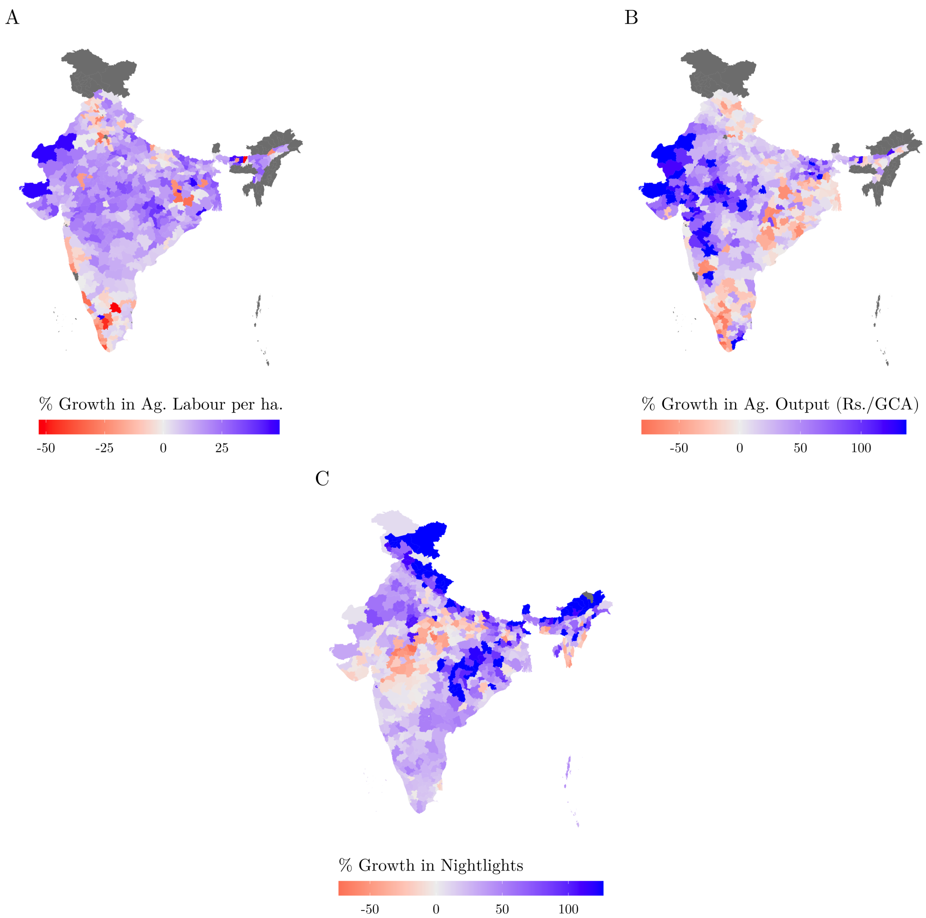


Figure 4: Labor Reallocation and Output Growth Across Districts (2001-2011)

Note: The time period for each map is 2001-2011. Panel A maps % change in agricultural workers per hectare of cultivated area using data from the 2001 and 2011 Census. Panel B maps the % change in crop output using data from the ICRISAT database. Output is measured as value of all crops produced in a district divided by gross cropped area. Values (in both panels) are truncated at the 5th and 95th percentile. Panel C maps percent change in nightlight intensity using data from the VIIRS satellite product.

labor exit and output contraction (Panels B, C; red) are concentrated in high-growth areas (Panel C, blue), agriculture thrives in low-growth areas (Panel C, red).

We formalize these patterns into a simple model in Appendix C and dedicate the remainder of this section to testing its predictions. The main model insight is that agricultural contraction among migrant-sending households is counterbalanced by increased production among non-migrant households, who tend to live in areas with low emigration rates. The model shows that land and crop market adjustments are the main mechanisms driving this spatial reorganization of agriculture. The declining farm size of migrant households increases land supply and allows non-migrant households to expand their farms and produce more crops. Lastly, similar to Adao et al. (2019) and Borusyak et al. (2022b), the model shows that the overall effect of migration on agricultural outcomes is composed of additively separable direct and indirect effects. We use this feature to extend our 2SLS setup to incorporate both channels.

5.2 Measurement

5.2.1 The Land Market Channel

To measure land market adjustments, we first conceive land markets at the village level. The land market adjustment is then measured by aggregate emigration from all other households i' of village j, in district d, excluding household i:

$$M_{ijdt}^{land} = \frac{\sum_{i' \in N_{jdt}/i} Migrants_{i'jdt}}{|N_{jdt}/i|} \tag{5}$$

Intuitively, M_{ijdt}^{land} measures average migration from other households $i' \neq i$ in the land market. $Migrants_{i'jdt}$ is the number of working-age male migrants sent from households i'. N_{jdt}/i is the set of households in village j excluding household i, and $|N_{jdt}/i|$ is the number of elements in N_{jdt}/i. We divide the number of migrants (except household i) by the number of households in village j (except household i) to account for heterogeneous village populations and land endowments. Household i's response to changes in M_{ijdt}^{land} captures responses through the indirect land channel. In Section 5.6, we show that their response to M_{ijdt}^{land} materializes through changes in village land prices (Table 5).

5.2.2 The Crop Market Channel

To measure crop market adjustments, we conceive crop markets as the convex hull around households in the state growing similar crops. To quantify crop similarity, we consider two vectors $\mathbf{x}_i = (x_{i1}, ..., x_{iK})$ and $\mathbf{x}_{i'} = (x_{i'1}, ..., x_{i'K})$ that list the K dominant crops available to households i and i'. The order of elements is identical for both households. $x_{ik} \in [0, 1]$ is the suitability of crop k in district d. Crop similarity between household i and i' is measured as the inverse Euclidean distance between their crop vectors[11]:

$$d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1} = \left(\sqrt{(x_{i1} - x_{i'1})^2 + (x_{i2} - x_{i'2})^2 + ... + (x_{iK} - x_{i'K})^2} + 1 \right)^{-1}$$

Using this idea of crop similarity, the crop market adjustment is measured as the weighted average emigration from other households i' in the crop market, with weights equal to the crop similarity between household i and i':

$$M_{ijdt}^{crop} = \sum_{i' \in N_s/i} d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1} Migrants_{i't} \tag{6}$$

[11] We use d(x)=1/(x+1) to avoid divide-by-zero issues.

where N_s/i denotes all households in state s excluding household i. Since agricultural markets are heavily regulated in India and cross-state trade is restricted, preventing crop prices from equilibrating across states (Chatterjee, 2023), we assume separate crop markets by state. Nevertheless, we let the market span the whole country in a robustness test. Since migration by working-age men from other households growing equally suitable crops to i receives more weight in the aggregation, M^{crop}_{ijdt} effectively measures aggregate emigration from the same crop market as i. Household responses to changes in M^{crop}_{ijdt} thus capture adjustments through crop markets. We show in Section 5.6 that their response captures an underlying relationship between M^{crop}_{ijdt} and crop prices (Table 5). Intuitively, if urban shocks trigger widespread labor migration from a region suitable for rice, thereby affecting aggregate rice supply and rice prices, then households elsewhere in the state will respond more to the price change if they also live in a rice-suitable region.

We use gridded crop suitability (1km resolution) from the GAEZ FAO portal (Fischer et al., 2021) to measure x_{ik}. Suitability is an index over the 1980-2010 period, normalized to one. We obtain separate suitability rasters for each major crop in India[12]. District suitability for each crop is computed by extracting means over all cells within a district.

An advantage of measuring crop markets in suitability space is that we avoid issues of endogenous crop choice between household i and i', which may be jointly determined by unobserved correlates of migration and crop production. In contrast, crop suitability is fixed and beyond the control of households, providing a higher-quality measure of the crop-growing region. Nevertheless, we test sensitivity to measuring crop similarity with actual production at baseline (Section 5.5).

5.3 Estimating Equation

In order to jointly estimate the direct and indirect effects of internal migration, we extend the 2SLS framework in Section 3.3 to include the land and crop market channels:

$$M^{labor}_{ijdt} = \mu_1 z_{ijdt} + \mu_2 s_{dt} + \mu_3 inc_{dt} + \mu_4 M^{land}_{ijdt} + \mu_5 M^{crop}_{ijdt} + \Gamma X'_{dt} + \alpha_i + \gamma_t + \varepsilon_{ijdt} \quad (7)$$

$$Y_{ijdt} = \beta_1 \widehat{M^{labor}_{ijdt}} + \beta_2 M^{land}_{ijdt} + \beta_3 M^{crop}_{ijdt} + \beta_4 s_{dt} + \beta_5 inc_{dt} + \Gamma X'_{dt} + \alpha_i + \gamma_t + \eta_{ijdt} \quad (8)$$

where M^{land}_{ijdt} captures the land market adjustment and M^{crop}_{ijdt} is the crop market adjustment. All other terms and subscripts are the same as Equation 4. The total response of farmers to emigration comprises the direct effect through household labor loss, β_1, as well as the indirect effects as land and crop markets adjust, captured by β_2 and β_3, respectively.

[12] Major crops are defined by the Department of Agriculture, Cooperation & Farmers Welfare as rice, wheat, nutri-cereals, pulses, foodgrains, oilseeds, sugarcane, cotton, and jute & mesta.

In our 2SLS setup, only the direct labor channel, M_{ijdt}^{labor}, is instrumented with z_{ijdt}, whereas the indirect market channels are not. This assumes that migration of other households in the village and crop market is unrelated to the decision of household i, conditional on controls and fixed effects. For land markets, the main threat to this assumption is that households i and i' may respond to village-level shocks (e.g., a local drought) that co-determine agricultural outcomes and migration. We address this by controlling for droughts, and, to the extent that these shocks manifest as economic shocks, these are captured by destination income shocks, s_{dt}, and own-district income changes, inc_{dt}. Another threat is that emigration from the land market may affect non-land factor markets, in which case β_2 is biased. We rule this out by controlling directly for non-land factor market changes in a placebo test (Section 5.5.2). For crop markets, we are less worried about endogeneity since the market is defined in crop suitability space, within which households span many villages, districts, and regions. We still validate the measure by showing that households do not respond to aggregate emigration from outside the crop market nor to emigration among households growing unrelated crops. Despite these validation exercises, the fact that M_{ijdt}^{land} and M_{ijdt}^{crop} are un-instrumented mean that we interpret estimates of β_2 and β_3 (indirect effects) with more caution than β_1 (direct effects).

5.4 Results: Direct and Indirect Effects

Table 4 presents estimates of β_1, β_2, and β_3 from Equation 8. The estimates of the direct and indirect effects are reported in standard deviations for ease of comparison. The main result is that direct and indirect effects draw in opposing directions, as suggested by the visual evidence and the model. The direct effect (row 1) remains negative, statistically significant, and similar in magnitude to Table 3 for all agricultural outcomes except agricultural wages. However, column 1 shows that migration-induced declines in agricultural technology due to the direct effect are partially offset by household responses to land (row 2) and crop (row 3) market adjustments. Table A9 columns 1-5 show that technology uptake through crop market adjustments materializes through spending more on seeds, agrochemicals, water, and rental equipment. There is little change in machinery ownership (columns 6-10). We show in Section 5.6 that agricultural expansion through crop market adjustments captures household responses to higher crop prices.

Column 3 of Table 4 shows that downsizing farms through the direct channel is partially offset by the indirect land market channel (row 2). This corroborates the model prediction that falling land prices induced by the direct effect prompt non-migrant households to expand farm size. We provide evidence for this price mechanism in Section 5.6. Columns 4 and 5 describe labor market responses. Wages are unaffected by aggregate

Table 4: Direct and Indirect Effects of Internal Migration

	Technology Index		Land (ac.)	Labor		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage bill	(5) Man-days	(6) Crops
Male Migrants	-1.078***	-0.735***	-1.269***	0.082	-2.009***	-1.358***
(direct labor channel)	(0.207)	(0.189)	(0.237)	(0.167)	(0.332)	(0.224)
Village emigration	0.237***	0.191***	0.259***	-0.022	0.419***	0.247***
(indirect land channel)	(0.049)	(0.044)	(0.052)	(0.037)	(0.082)	(0.054)
Crop region emigration	0.207***	0.040	0.242***	-0.015	0.422***	0.178***
(indirect crop channel)	(0.045)	(0.039)	(0.046)	(0.029)	(0.066)	(0.043)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25924	24966	29342	20906	20906	25852
F-Stat on Direct Effect	63.9	56.8	55.7	55.8	55.8	67.5

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All variables are standardized. *Male Migrants* is the number of working-age male migrants sent from the household, instrumented with inverse-distance population-weighted income interacted with a number of baseline working-age males. *Village emigration* is the leave-one-out average number of working-age male migrants in the village. *Crop region emigration* is the leave-one-out number of migrants in the state weighted by inverse Euclidean distance between crop suitability (Section 5.2). Columns 1 and 2 are indices for expenses and machinery, respectively. The wage bill (column 4) is wages paid to hired labor in the past year. Person-days (column 5) include family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

emigration at the level of the crop market.

Lastly, the estimates for crop output[13] in column 6 show the same directionality through each channel. Point estimates are precise and economically significant. Whereas households reduce production by 1.36σ in response to their own labor loss, output expansion through crop market adjustments offsets this by 18% (= 0.247/1.358).

These results implicitly feature a spatial component. The direct effect predominantly impacts households sending many migrants, for whom large output declines are not offset by the indirect effects. In contrast, households sending zero migrants exclusively experience the indirect positive effects in the form of greater technology adoption and production. Since remote households send fewer migrants (Fact 2, Section 2.4), we expect a

[13]Since nationwide price changes are absorbed by year fixed effects, the coefficients in Column 6 can be interpreted as output changes even though the measure is profits.

spatial reorganization of agriculture away from urban areas (high-migration) and toward remote areas (low-migration). We observed this pattern in aggregated data in Figure 4.

5.5 Robustness and Placebo Tests

Next, we test the robustness of our estimates of the indirect effects of internal migration. We also probe our crop and land market measures with placebo tests. Our estimates are remarkably stable, and our market measures appear robust and well-defined.

5.5.1 Robustness Tests

Alternative Crop Market Measure: Table A10 tests robustness to an alternative definition of the crop market. For this test, crop similarity weights are defined in crop production space, not suitability space. Crop-wise production data is available in the 2005 IHDS survey. We define M_{ijdt}^{crop} as in Equation 6, except that $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ is constructed such that K is the number of possible crops that households report growing, and x_{ik} is crop k's share of actual production for household i. $x_{ik} = 0$ if crop k is not grown. Reassuredly, the estimates β_3 are virtually identical to Table 4 when we use this alternative measure of M_{ijdt}^{crop}.

Agricultural Regulation: Table A11 reports estimates when we allow crop markets to span the whole of India. This enables household i to be affected by the emigration of households growing similarly suitable crops anywhere in India rather than only within the same state. The coefficients on the crop channel remain positive across most specifications, but the magnitudes increase. Estimate sensitivity is likely driven by the fact that agricultural markets in India function at the state level, in which case this robustness test may be misspecified. We therefore prefer our main specification.

Extensive Margin: The estimates in Table 4 describe the intensive margin since the sample comprises households that owned land in both periods. Households who left or entered agriculture between surveys had missing outcomes in the period when they were landless. We add them to the sample by zero-filling their agricultural technology expenses and machinery. Table A12 shows that our estimates of the direct and indirect effects remain remarkably stable, suggesting that including the extensive margin of farmer exit and entry in response to labor migration does not change our main conclusions.

5.5.2 Placebo Tests

In the absence of instruments for the indirect effects, we conduct placebo tests designed to validate the measures of the crop and land market adjustments and help support the causal interpretation of our estimates. First, we compute a placebo crop market measure by replacing $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ in Equation 6 with $d(\mathbf{x}_i, \mathbf{x}_{i'})$. Since these weights are not inverted, households in the state growing *different* crops receive more weight in the aggregation. If our estimates of household responses to crop market adjustments are truly driven by household i reacting to aggregate supply shifts induced by emigration from the crop market, then the emigration of households outside the crop market should have no effect on household i's output. In column 1 of Table A13, we control for the placebo in Equation 8, and its coefficient turns negative and converges to zero. Therefore, the positive coefficient on M_{ijdt}^{crop} indicates that our estimates of β_2 are indeed driven by household i's response to crop market adjustments.

Next, we compute a more sophisticated crop market placebo that accounts for crop substitutability. Even if households i and i' grow different crops, their production may still be linked through prices if they grow substitutes or complements. We, therefore, measure $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ as the inverse cross-price elasticity between the main crop of household i and i' obtained from Anand et al. (2016)[14]. Taking the inverse means that households growing unrelated crops receive more weight in the aggregation. Column 2 of Table A13 controls for this placebo, and, once again, the placebo coefficient is near-zero. This suggests that households' reaction to crop market adjustments in response to aggregate emigration does not only depend on supply shifts of their main crops within the crop market, but also supply shifts of substitutes and complements within the market.

Lastly, column 3 presents placebo estimates that validate the land market measure. One concern is that the existing measure—emigration by other households in the village—also picks up adjustments through non-land factor market changes in the village. We disentangle these channels by controlling for emigration among *non-agricultural* households in the village in Equation 8. When farm size is the outcome, the coefficient on the land market placebo is statistically indistinguishable from zero, whereas the main land market coefficient remains stable. Thus, our land market measure, M_{ijdt}^{land}, is indeed picking up farm size adjustments through land markets and not other indirect channels.

5.6 Mechanisms

We have argued throughout that β_2 and β_3 in Equation 8 capture households' indirect responses to land and crop market adjustments, respectively. The model in Appendix C shows that prices are the guiding force behind these indirect effects. We provide evidence

[14]Compensated elasticities for India are provided for four relevant crop categories: cereals, pulses, veggies/fruits, and other. We first identify the main crop of each household as the one with the highest share of total production. Second, we categorize each crop into one of the four categories. Lastly, we fill the weight matrix with the absolute value of elasticities for each pair of households.

Table 5: Mechanisms: Impact of Aggregate Emigration on Prices

	(1) Crop Price	(2) Land Price (rent in)	(3) Land Price (rent out)
Crop region emigration	0.137***		
(indirect crop channel)	(0.052)		
Village emigration		-0.052**	-0.063*
(indirect land channel)		(0.025)	(0.034)
Aggregation	District-year	Village-year	Village-year
District FEs	✓		
Village FEs		✓	✓
Year FEs	✓	✓	✓
Observations	586	1222	1296
R^2	0.892	0.667	0.633

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All outcomes are in standard deviations. In column 1, data are at the district-year level. The outcome is a district crop price index. In columns 2 and 3, data are at the village-year level. Regressions are weighted by the number of households renting in our out, respectively. *Crop region emigration* is a district average of the leave-one-out number of migrants *within the state* weighted by crop suitability. *Village emigration* is mean number of working age male migrants in the village. All columns control for drought conditions. Errors clustered by district in column 1 and by village in columns 2-3.

of the price mechanism using data on crop an land prices. Crop prices are at the district-year level from ICRISAT. We construct a price index for each district as the weighted average crop price for major crops, weighted by normalized suitability for each crop in the district (Section 5.2 for data details). Land prices are reported in IHDS for households that buy or sell land. We measure land price at the village level, and weight regressions by the number of households trading land in the village. This ensures that coefficients are more influenced by observations measured with better precision.

We estimate the price effects of emigration from the crop and land market as follows:

$$CropPrice_{dt} = \Phi \cdot M_{dt}^{crop} + \alpha_d + \gamma_t + \epsilon_{dt}$$

$$LandPrice_{jt} = \psi \cdot M_{jt}^{land} + \alpha_j + \gamma_t + \epsilon_{jt}$$

where d, j, and t index districts, villages, and time, respectively. $CropPrice$ is the crop price index in district d and $LandPrice$ is land price in village j . M_{dt}^{crop} and M_{jt}^{land} are measured as before, but aggregated at the district and village level, respectively. The coefficients of interest are Φ and ψ. If $\Phi > 0$, then emigration from the crop market raises crop prices.

If $\psi < 0$, then emigration from the land market reduces land prices. All specifications include location and time fixed effects.

Table 5 presents the price effects. Outcomes are reported in standard deviations. Starting with crop price in column 1, the coefficient is positive and statistically significant. The point estimate implies that a 1σ increase in aggregate emigration from the crop market increases crop prices by 0.14σ. This suggests that crop prices are a key mechanism behind the positive crop market effect in Table 4; declining crop output through the labor channel raises crop prices, leading non-migrant farmers to increase production.

Column 2 documents the impact of aggregate emigration from the land market on the price of land rented in. The coefficient is negative and significant, implying that emigration from the land market depresses land prices. This suggests that land prices are a key mechanism behind the positive land market effects in Table 4; the reduction in farm size through the labor channel increases the supply of land, thereby reducing land prices and leading to farm expansion among non-migrant farmers. The land price effect is robust to using seller price as the outcome (column 3).

6 Accounting Exercise

In aggregate, how much of the agricultural losses are compensated by market adjustments? We conduct a simple accounting exercise to answer this. About 80% of the migration-induced decline in food supply is compensated by domestic market adjustments. Given that these indirect effects materialize in remote areas, this implies a spatial reorganization of agriculture away from peri-urban areas and toward remote regions.

6.1 Methods

For each household i in year t, we use the coefficients from Equation 8 to predict household i's crop production as a function of migration realizations. Let t_1 and t_2 denote IHDS survey wave I and II, respectively. We first define a baseline *No Migration* (NM) scenario in which all migration variables are fixed at t_1:

$$Y^{NM}_{ijdt_2} = \hat{\beta}_1 \widehat{M^{labor}_{ijdt_1}} + \hat{\beta}_2 M^{land}_{ijdt_1} + \hat{\beta}_3 M^{crop}_{ijdt_1} + \alpha_i + \gamma_{t_2}$$

We also define a *Labor Only* (LO) scenario in which both indirect effects are fixed at t_1:

$$Y^{LO}_{ijdt_2} = \hat{\beta}_1 \widehat{M^{labor}_{ijdt_2}} + \hat{\beta}_2 M^{land}_{ijdt_1} + \hat{\beta}_3 M^{crop}_{ijdt_1} + \alpha_i + \gamma_{t_2}$$

In the same way, we define a *Labor and Land* (LL) scenario where the indirect crop market adjustment is fixed at t_1, as well as a *Labor and Crop* (LC) scenario where the indirect land market adjustment is fixed at t_1. We then sum household-level predicted values to construct aggregate counterfactuals: $TotY^{NM}_{t_2}$, $TotY^{LO}_{t_2}$, $TotY^{LL}_{t_2}$ and $TotY^{LC}_{t_2}$, the total value of crop production without migration, with migration but no market adjustments, with migration plus land market adjustments, and with migration plus crop market adjustments, respectively. Comparisons with in-sample fitted values $TotY_{t_2}$ (the scenario with all channels operational) yield three statistics of interest. The first and second are aggregate values of crop production with (GE) and without (PE) market adjustments relative to the total value absent migration:

$$PctChange^{GE} = 100 \cdot \left(\frac{TotY_{t_2} - TotY^{NM}_{t_2}}{TotY^{NM}_{t_2}} \right) \quad PctChange^{PE} = 100 \cdot \left(\frac{TotY^{LO}_{t_2} - TotY^{NM}_{t_2}}{TotY^{NM}_{t_2}} \right)$$

The third is the amount of agricultural decline through the direct channel that is offset by markets as a percent of the counterfactual change absent market adjustments:

$$PctOffset = 100 \cdot \left(\frac{PctChange^{PE} - PctChange^{GE}}{PctChange^{PE}} \right) \tag{9}$$

We follow the same steps to decompose the land channel and crop channel separately.

6.2 Back-of-the-Envelope Estimates

Figure 5A shows estimates of the change in aggregate crop output with and without indirect market spillovers. To account for uncertainty in the underlying point estimates, we repeat the exercise on 100 bootstrapped samples and plot the mean (orange) and confidence intervals (purple error bars). Under the *Labor Only* scenario, with indirect channels shut off, aggregate migration would have caused a 53% reduction in agricultural output compared to the *No Migration* counterfactual. This amounts to Rs. 118 million worth of food. When all channels operate, the supply contraction becomes ten times smaller.

Panel B shows how much of the migration-induced output decline is mitigated by market adjustments (Equation 9). 38% of agricultural losses through the labor channel are mitigated through land markets and 42% through crop markets. Both indirect channels together mitigate 80% of the direct effects of migration. The recovered crops are worth Rs. 95 million, or 18% of the in-sample total crop value in 2012.

It is important to note that these results do not imply a net loss of food across India, only that we are able to quantify the partial effects of these two observed margins of adjustment. Other unobserved determinants of output, such as climate, are captured by the time fixed effects in our empirical model. The comparison is, therefore, with a hypothetical counterfactual in which all the other determinants of production are considered but migration is reduced to zero.

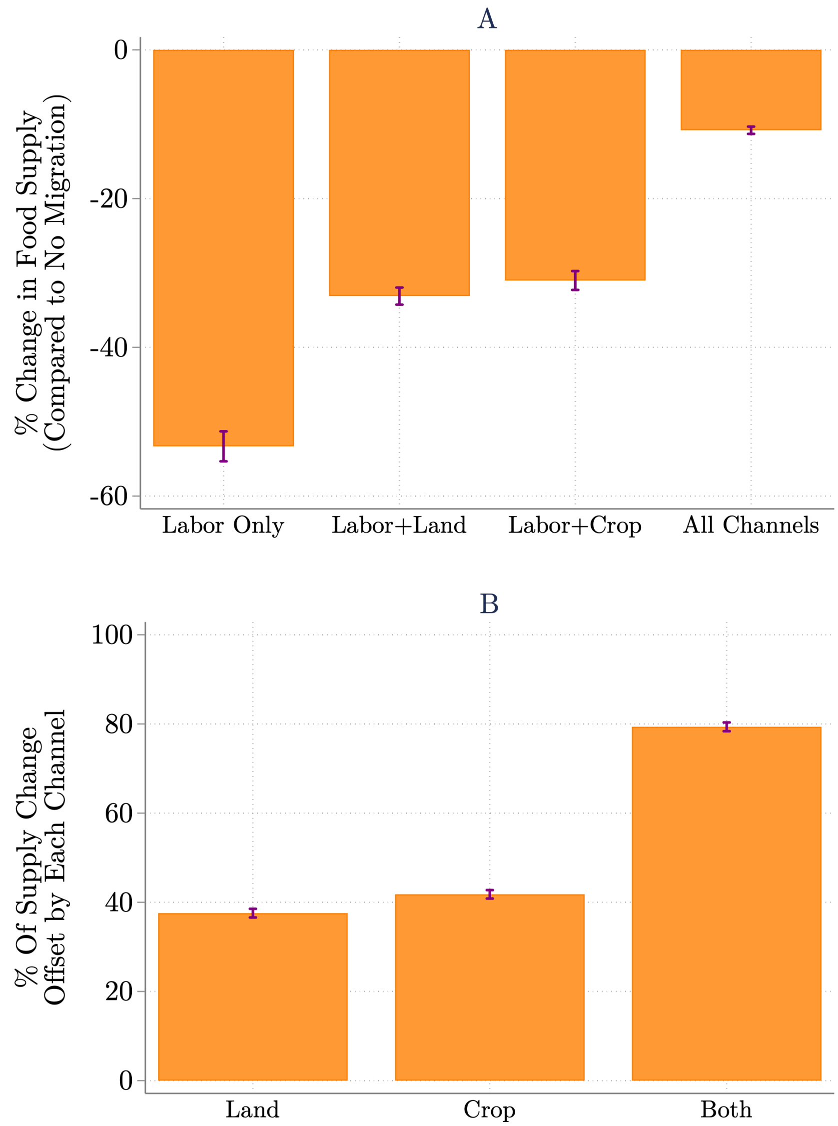


Figure 5: Aggregate Extent of Indirect Effects

Note: Panel A displays the aggregate change in agricultural supply from migration under the four scenarios. *Labor Only* means the crop and land channels are held constant, *Labor + Crop* means the land channel is held constant, and so on. Panel B shows the percent of the *Labor Only* agricultural decline mitigated by the indirect land and crop forces (Equation 9). Confidence intervals computed from 100 bootstrap draws.

7 Conclusion

This paper studies how internal migration affects agricultural development and the spatial organization of agriculture. The reallocation of labor from farms to cities is an emblematic feature of economic development. While this process is well-studied, its effects on agriculture are relatively unexplored. We track labor migration and agricultural activities with detailed household panel data from India between 2005-2012, a period of rapid economic modernization. We address the endogeneity of migration choice using a shift-share instrument based on distance-weighted destination income shocks interacted with households' potential to benefit from these shocks.

In contrast to commonly held beliefs, we find that Indian farmers do not replace labor with capital when facing agricultural labor loss. Instead, they adopt less technology, downsize farms, and reduce crop output. Yet this does not mean that we should expect food scarcity as India urbanizes. Equilibrium effects on food supply depend on how crop and land markets adjust. We document rising crop prices in response to migration-induced output declines, and falling land prices as more land becomes available.

Putting together the direct and indirect effects, we find evidence of a spatial reorganization of agriculture. While households in peri-urban areas face agricultural losses, households with no migrants in remote areas benefit from higher crop prices and lower land prices. These households increase production, expand farmland, and adopt technology. In aggregate, our estimates suggest that indirect market spillovers mitigate 80% of the migration-induced food shortage between 2005-2012. The spatial redistribution of agriculture through markets is, therefore, economically significant but not a panacea.

Many studies find the opposite result: that agriculture modernized more in areas experiencing labor emigration (Hornbeck and Naidu, 2014; Manuelli and Seshadri, 2014). One reason for the discrepancy is that previous literature uses aggregate data at the county or district level. Our household-level results suggest that non-migrant households in high-migration areas invest more in agriculture than non-migrant households in low-migration areas. This interpretation reconciles our findings with the literature.

Our results have important distributional implications. We showed that while agricultural development declines in peri-urban areas, it surges in remote areas where low migration rates and poverty are widespread. We thus expect structural transformation to promote income redistribution toward those who do not directly participate in it. Evidence on this phenomenon is an open area for future research.

References

Acemoglu, Daron, "When does labor scarcity encourage innovation?," *Journal of Political Economy*, 2010, *118* (6), 1037–1078.

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito, "General equilibrium effects in space: Theory and measurement," Technical Report, National Bureau of Economic Research 2019.

Allen, Treb and Costas Arkolakis, "Economic Activity across Space," *The Journal of Economic Perspectives*, 2023, *37* (2), 3–28.

Alvarez-Cuadrado, Francisco and Markus Poschke, "Structural change out of agriculture: Labor push versus labor pull," *American Economic Journal: Macroeconomics*, 2011, *3* (3), 127–58.

Anand, Rahul, Naresh Kumar, and Mr Volodymyr Tulin, "Understanding India's food inflation: The role of demand and supply factors," *IMF Working Paper No. 16/2*, 2016.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

Asher, Sam, Alison Campion, Douglas Gollin, and Paul Novosad, "The Long-run Development Impacts of Agricultural Productivity Gains: Evidence from Irrigation Canals in India," *Mimeo*, 2023.

Blakeslee, David, Aaditya Dar, Ram Fishman, Samreen Malik, Heitor S Pellegrina, and Karan Singh Bagavathinathan, "Irrigation and the spatial pattern of local economic development in India," *Journal of Development Economics*, 2023, *161*, 102997.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 2022, *89* (1), 181–213.

__, Rafael Dix-Carneiro, and Brian Kovak, "Understanding Migration Responses to Local Shocks," *Available at SSRN 4086847*, 2022.

Brewer, Julia, Ashley Larsen, and Frederik Noack, "The land use consequences of rural to urban migration," *American Journal of Agricultural Economics*, 2022.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural productivity and structural transformation: Evidence from Brazil," *American Economic Review*, 2016, *106* (6), 1320–65.

Caprettini, Bruno and Hans-Joachim Voth, "Rage against the machines: Labor-saving technology and unrest in industrializing England," *American Economic Review: Insights*, 2020, *2* (3), 305–20.

Caunedo, Julieta and Namrata Kala, "Mechanizing Agriculture," Working Paper 29061, National Bureau of Economic Research July 2021.

Chatterjee, Shoumitro, "Market Power and Spatial Competition in Rural India," *The Quarterly Journal of Economics*, 2023.

Clemens, Michael A, Ethan G Lewis, and Hannah M Postel, "Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion," *American Economic Review*, 2018, *108* (6), 1468–87.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with arbitrary clustering," *IZA Discussion Paper No. 12584*, 2019.

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of econometrics*, 1999, *92* (1), 1–45.

Duarte, Margarida and Diego Restuccia, "The role of the structural transformation in aggregate productivity," *The quarterly journal of economics*, 2010, *125* (1), 129–173.

Emerick, Kyle, "Agricultural productivity and the sectoral reallocation of labor in rural India," *Journal of Development Economics*, 2018, *135*, 488–503.

Fernando, A Nilesh, "Shackled to the soil? Inherited land, birth order, and labor mobility," *Journal of Human Resources*, 2022, *57* (2), 491–524.

Fink, Günther, B Kelsey Jack, and Felix Masiye, "Seasonal liquidity, rural labor markets, and agricultural production," *American Economic Review*, 2020, *110* (11), 3351–3392.

Fischer, Gunther, FO Nachtergaele, HT Van Velthuizen, F Chiozza, G Franceschini, M Henry, D Muchoney, and S Tramberend, *Global Agro-Ecological Zones v4–Model documentation*, Food & Agriculture Org., 2021.

Foster, Andrew D and Mark R Rosenzweig, "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size," *Journal of Political Economy*, 2022, *130* (3), 636–680.

Gao, Xuwen, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song, "Restrictions on Migration Create Gender Inequality: The Story of China's Left-Behind Children," *NBER Working Paper No. 30990*, 2022.

Garriga, Carlos, Aaron Hedlund, Yang Tang, and Ping Wang, "Rural-urban migration, structural transformation, and housing markets in China," *American Economic Journal: Macroeconomics*, 2023, *15* (2), 413–440.

Gollin, Douglas, David Lagakos, and Michael E. Waugh, "THE AGRICULTURAL PRODUCTIVITY GAP," *The Quarterly Journal of Economics*, 2014, *129* (2), 939–994.

Harris, John R and Michael P Todaro, "Migration, unemployment and development: a two-sector analysis," *The American economic review*, 1970, pp. 126–142.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring economic growth from outer space," *American economic review*, 2012, *102* (2), 994–1028.

Hornbeck, Richard and Suresh Naidu, "When the levee breaks: black migration and economic development in the American South," *American Economic Review*, 2014, *104* (3), 963–90.

Hsiang, Solomon M, "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America," *Proceedings of the National Academy of sciences*, 2010, *107* (35), 15367–15372.

Huber, Kilian, "Estimating general equilibrium spillovers of large-scale shocks," *The Review of Financial Studies*, 2023, *36* (4), 1548–1584.

Kaboski, Joseph P, Jianyu Lu, Wei Qian, and Lixia Ren, "Melitz Meets Lewis: The Impacts of Roads on Structural Transformation and Businesses," *NBER WP No. 32448*, 2024.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry, "Agricultural decisions after relaxing credit and risk constraints," *The Quarterly Journal of Economics*, 2014, *129* (2), 597–652.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of econometrics*, 2006, *133* (1), 97–126.

Lagakos, David and Michael E Waugh, "Selection, agriculture, and cross-country productivity differences," *American Economic Review*, 2013, *103* (2), 948–80.

Lewis, William Arthur, "Economic development with unlimited supplies of labour," 1954.

Manuelli, Rodolfo E and Ananth Seshadri, "Frictionless technology diffusion: The case of tractors," *American Economic Review*, 2014, *104* (4), 1368–91.

Mendola, Mariapia, "Migration and technological change in rural households: Complements or substitutes?," *Journal of Development Economics*, 2008, *85* (1-2), 150–175.

Morten, Melanie and Jaqueline Oliveira, "The effects of roads on trade and migration: Evidence from a planned capital city," *NBER Working Paper*, 2018, *22158*, 1–64.

Moscona, Jacob, "Agricultural development and structural change within and across countries," Technical Report, Mimeo, MIT 2019.

Pellegrina, Heitor S and Sebastian Sotelo, "Migration, Specialization, and Trade: Evidence from Brazil's March to the West," Technical Report, National Bureau of Economic Research 2021.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu, "Agriculture and aggregate productivity: A quantitative cross-country analysis," *Journal of monetary economics*, 2008, *55* (2), 234–250.

Rozelle, Scott, J Edward Taylor, and Alan DeBrauw, "Migration, remittances, and agricultural productivity in China," *American Economic Review*, 1999, *89* (2), 287–291.

Schwartz, Aba, "Interpreting the effect of distance on migration," *Journal of political economy*, 1973, *81* (5), 1153–1169.

Taylor, J Edward, Scott Rozelle, and Alan De Brauw, "Migration and incomes in source communities: A new economics of migration perspective from China," *Economic Development and Cultural Change*, 2003, *52* (1), 75–101.

Veljanoska, Stefanija, "Do remittances promote fertilizer use? The case of Ugandan farmers," *American Journal of Agricultural Economics*, 2022, *104* (1), 273–293.

Vollrath, Dietrich, "How important are dual economy effects for aggregate productivity?," *Journal of development economics*, 2009, *88* (2), 325–334.

Appendix - For Online Publication Only

A Appendix Tables

Table A1: Migrant Distribution by Type by Age Cohort

Age Cohort	Student	Employed Son	Husband
0-4	0.69	0.09	0.06
5-9	0.96	0.01	0.00
10-14	0.91	0.05	0.00
15-19	0.61	0.29	0.02
20-24	0.33	0.44	0.10
25-29	0.07	0.44	0.25
30-34	0.01	0.37	0.33
35-39	0.01	0.32	0.36
40-44	0.00	0.30	0.40
45-49	0.00	0.22	0.46
50-54	0.00	0.19	0.41
55-59	0.00	0.11	0.47
60+	0.00	0.02	0.18
Total	0.27	0.30	0.20

Note: Data from IHDS Wave I (2004-05). Each row denotes an age cohort. Values denote the share of migrants in each age group belonging to each migrant type.

Table A2: Summary Statistics: Migrant Profiles

	IHDS-1 (2004-05)			IHDS-II (2011-12)		
	# Migrants	Share	SD	# Migrants	Share	SD
A: Gender						
Male	3516	0.80	0.40	10478	0.79	0.41
Female	861	0.20	0.40	2795	0.21	0.41
B: Status						
Student	1153	0.26	0.44	3602	0.27	0.44
Working	2501	0.57	0.49	9427	0.71	0.45
Neither	723	0.17	0.37	244	0.02	0.13
C: Destination						
Within State	2787	0.64	0.48	8639	0.65	0.48
Out of State	1590	0.36	0.48	4634	0.35	0.48
D: Stream						
Rural-Rural	1108	0.25	0.43	4287	0.32	0.47
Rural-Urban	2065	0.47	0.50	5833	0.44	0.50
Urban-Rural	542	0.12	0.33	1519	0.11	0.32
Urban-Urban	475	0.11	0.31	707	0.05	0.22

Note: The table describes data on migrants in each period. There are 4,377 migrants in Wave I and 13,273 in wave II. Subgroups are mutually exclusive. ”Rural-Rural” indicates the origin and destination were rural, ”Rural-Urban” indicates the origin is rural and the destination is urban, and so on.

Table A3: Entire-Family Migration

	# HH	%	SD
Moved within last 7 yrs	421	1.1	10.2
Place of Origin			
Same state, same district	360	84.9	35.8
Same state, another district	49	11.6	32.0
Another State	15	3.5	18.5
Type of Origin			
Village	240	56.3	49.7
Town	186	43.7	49.7
Total	40,018		

Note: Data from IHDS wave 2. The first row reports households that moved between survey waves. % refers to the percent of all households (see Total). Remaining rows pertaining only to those households that moved between waves.

Table A4: First Stage Results

Outcome: # of working-age male migrants	(1)	(2)	(3)
Wt. Income $\times$ Working Age Males (z_{idt})	0.325***	0.381***	0.383***
	(0.046)	(0.048)	(0.048)
Wt. Income (s_{dt})	No	Yes	Yes
Origin Income (inc_{dt})	No	No	Yes
HH FEs	✓	✓	✓
Year FEs	✓	✓	✓
KP (2006) F-Stat	48.89	62.64	63.41
Observations	25854	25854	25854

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The outcome is number of working age male migrants sent from household i. "Wt. Income" is inverse-distance, population-weighted income (s_{dt}). "Working Age Males" is the number of working-age male household residents (φ_i). "Origin Income" is mean per capita district income, inc_{dt}. The units of z_{idt} are standard deviations. All specifications control for drought conditions. Standard errors clustered by PSU.

Table A5: Second Stage—Direct Effect of Migration on Technology Adoption

	Agricultural Expenses (Rs.)					Machinery (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants (σ)	-1.023***	-1.420***	-0.755***	-0.529***	-0.787***	-0.411***	-0.445**	-0.952***	-0.348	-0.576***
	(0.201)	(0.246)	(0.175)	(0.197)	(0.183)	(0.145)	(0.181)	(0.219)	(0.215)	(0.192)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outcome SD	2076.457	2802.867	1138.541	709.926	1583.843	0.409	0.523	0.382	0.257	0.220
Explanatory SD	0.524	0.523	0.519	0.521	0.524	0.521	0.520	0.519	0.520	0.519
HH FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	25852	25454	24054	23496	24602	24692	23862	24042	23882	23746
F-Stat	58.6	56.8	53.0	51.9	56.4	54.3	55.3	53.8	53.1	52.9

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The explanatory variable is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. Outcomes are in standard deviations. Columns 1-5 are measured in Rs. spent in the past year. Water refers to purchases of irrigation water. Rentals include both hired equipment and animals. Column 6-10 are measured in quantities. Pumps include both electric and diesel pumps. All specification control for district drought conditions. Standard errors clustered by PSU.

Table A6: Labor-to-Land Ratio by Farm Size

	Small (0-2 ac.)		Medium (2-4 ac.)		Large (4+ ac.)	
	Mean	SD	Mean	SD	Mean	SD
Family person-days/acre	335.22	777.37	128.43	421.86	70.45	119.03
Hired person-days/acre	14.83	46.98	12.82	112.23	8.92	23.98

Note: Data are a panel for all land-owning households. Farms are placed into size bins based on baseline farm size. Family and hired person-days of labor refer to the past year.

Table A7: Robustness Checks: Direct Effects

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Man-days	(6) Crops
Panel A: State-Year FEs						
Male Migrants (σ)	-0.830***	-0.584***	-1.181***	0.036	-1.781***	-1.438***
	(0.175)	(0.166)	(0.213)	(0.112)	(0.293)	(0.235)
Observations	25928	24970	29346	20910	20910	25854
Panel B: Control for Education						
Male Migrants (σ)	-1.171***	-0.846***	-1.335***	0.061	-2.124***	-1.514***
	(0.264)	(0.241)	(0.298)	(0.168)	(0.400)	(0.288)
Observations	25928	24970	29346	20910	20910	25854
Panel C: Pre-period Trends						
Male Migrants (σ)	-1.096***	-0.697***	-1.296***	0.089	-1.965***	-1.429***
	(0.221)	(0.189)	(0.248)	(0.150)	(0.336)	(0.239)
Observations	25928	24970	29346	20910	20910	25854
Panel D: Extensive Margin						
Male Migrants (σ)	-1.261***	-0.915***	-1.151***	-0.047	-1.874***	-1.214***
	(0.208)	(0.173)	(0.170)	(0.130)	(0.238)	(0.181)
Observations	69756	68698	75334	63970	63970	66040
Panel E: Rural-Urban Migration						
Male Migrants (σ)	-1.077***	-0.721***	-1.269***	0.071	-2.016***	-1.386***
	(0.218)	(0.201)	(0.235)	(0.179)	(0.329)	(0.237)
Observations	25148	24288	28176	20256	20256	24836
Panel F: Nightlights Shift-Share						
Male Migrants (σ)	-1.119***	-0.661***	-1.369***	-0.015	-2.151***	-1.554***
	(0.239)	(0.207)	(0.269)	(0.143)	(0.378)	(0.265)
Observations	25928	24970	29346	20910	20910	25854
Panel G: All Gender Migration						
Migrants (σ)	-1.934***	-1.412***	-2.416***	0.064	-3.259***	-2.312***
	(0.481)	(0.406)	(0.574)	(0.227)	(0.733)	(0.468)
Observations	25928	24970	29346	20910	20910	25854

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All outcome variables are standardized. The explanatory variable is the number of working age male migrants in the household. See tables notes of Table 3 for details on the migration instrument and other specification details). Panel A includes state-year FEs. The remaining panels only include year FEs. Panel B includes a control for whether the household head is educated (beyond 10th grade). Panel C controls for baseline trends in the distance-weighted income shock. Panel D adds households who left and joined agriculture between survey waves to the estimation sample. Panel E restricts to rural-urban migration only. Panel F uses distance-weighted nightlights as the shift. In Panel G, the explanatory variable and the IV "share" includes both males and females. All specifications control for district drought conditions. Standard errors are clustered at the PSU level.

Table A8: Robustness: Alternative Standard Error Clustering

	Machinery	Expenses	Cultivated	Wage Bill	Labor	Profits
Coefficient on Migration	-1.046	-0.709	-1.218	0.076	-1.877	-1.347
SE: Robust	0.160	0.135	0.181	0.132	0.257	0.188
SE: District	0.223	0.215	0.253	0.170	0.316	0.243
SE: State	0.290	0.237	0.332	0.199	0.401	0.324
SE: Conley (200km radius)	0.229	0.160	0.195	0.130	0.315	0.225
SE: Conley (500km radius)	0.175	0.100	0.222	0.032	0.307	0.268

Note: The first row replicates the main coefficient estimates from Table 3. Conley standard errors are estimated in a 2SLS setting using the Colella et al. (2019) implementation.

Table A9: Second Stage—Direct and Indirect Effects of Migration on Agricultural Development

	Agricultural Expenses (Rs.)					Machinery (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants	-1.053***	-1.464***	-0.792***	-0.557***	-0.814***	-0.418***	-0.469**	-0.986***	-0.377*	-0.600***
(direct labor channel)	(0.201)	(0.249)	(0.183)	(0.206)	(0.188)	(0.153)	(0.193)	(0.230)	(0.228)	(0.203)
Village emigration	0.230***	0.318***	0.154***	0.134***	0.202***	0.142***	0.109**	0.231***	0.089*	0.123**
(indirect land channel)	(0.046)	(0.057)	(0.044)	(0.046)	(0.044)	(0.038)	(0.044)	(0.054)	(0.052)	(0.048)
Crop region emigration	0.201***	0.266***	0.199***	0.130***	0.123***	-0.062	0.048	0.040	0.085**	0.036
(indirect crop channel)	(0.041)	(0.052)	(0.040)	(0.044)	(0.043)	(0.039)	(0.041)	(0.046)	(0.040)	(0.038)
Wt. Income	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HH FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	25848	25450	24050	23492	24598	24690	23860	24040	23880	23744
F-Stat on Direct Effect	63.6	61.1	56.0	54.1	59.7	55.4	55.5	54.7	53.0	52.8

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All variables are standardized. *Male Migrants* is number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with number of baseline working age males. *Village emigration* is the leave-one-out average number of working age male migrants in the village. *Crop region emigration* is the leave-one-out number of migrants in the state weighted by crop suitability (see section 5.2 for measurement). Columns 1-5 are measured in Rs. spent in the past year. Water refers to purchases of irrigation water. Rentals include both hired equipment and animals. Column 6-10 are measured in quantities. Pumps include both electric and diesel pumps. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table A10: Robustness: Crop Markets Measured in Crop Production Space

	Technology Index		Land (ac.)	Labor		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Man-days	Crops
Male Migrants	-1.052***	-0.754***	-1.276***	0.092	-1.977***	-1.332***
(direct labour channel)	(0.205)	(0.192)	(0.253)	(0.170)	(0.331)	(0.259)
Village emigration	0.227***	0.193***	0.269***	-0.026	0.411***	0.238***
(indirect land channel)	(0.048)	(0.045)	(0.056)	(0.037)	(0.082)	(0.062)
Crop region emigration	0.241***	0.044	0.239***	-0.010	0.460***	0.249***
(indirect crop channel)	(0.047)	(0.044)	(0.052)	(0.037)	(0.072)	(0.052)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25662	24632	27102	20644	20644	20886
F-Stat on Direct Effect	63.8	56.2	50.5	54.5	54.5	49.4

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All variables are standardized. *Male Migrants* is number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with number of baseline working age males. *Village emigration* is the leave-one-out average number of working age male migrants in the village. *Crop region emigration* is the leave-one-out number of migrants across households in the state weighted by inverse euclidean distance between actual crop portfolios between household i and i'. Columns 1 and 2 are indices for expenses and machinery, respectively. Wage bill (column 4) is wages paid to hired labor in the past year. Person-days (column 5) includes family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table A11: Robustness: Direct and Indirect Effects (Nationwide Crop Markets)

	Technology Index		Land (ac.)	Labor		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Man-days	(6) Crops
Male Migrants	-1.075***	-0.736***	-1.263***	0.084	-1.975***	-1.370***
(direct labour channel)	(0.208)	(0.187)	(0.235)	(0.165)	(0.325)	(0.226)
Village emigration	0.282***	0.199***	0.308***	-0.024	0.499***	0.288***
(indirect land channel)	(0.055)	(0.050)	(0.059)	(0.041)	(0.088)	(0.061)
Crop region emigration	0.591***	0.254	1.230***	-0.261*	0.781**	0.770***
(indirect crop channel)	(0.178)	(0.155)	(0.219)	(0.144)	(0.309)	(0.197)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25924	24966	29342	20906	20906	25852
F-Stat on Direct Effect	64.5	58.0	56.8	57.2	57.2	67.1

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All variables are standardized. *Male Migrants* is number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with number of baseline working age males. *Village emigration* is the leave-one-out average number of working age male migrants in the village. *Crop region emigration* is the leave-one-out number of migrants across India weighted by crop suitability (see Section 5.2 for details). Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively. Wage bill (column 4) is wages paid to hired labour in the past year. Person-days (column 5) includes family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table A12: Robustness: Direct and Indirect Effects (Extensive Margin)

	Technology Index		Land (ac.)	Labor		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Man-days	Crops
Male Migrants	-1.402***	-0.995***	-1.300***	-0.056	-2.126***	-1.332***
(direct labour channel)	(0.236)	(0.195)	(0.200)	(0.146)	(0.283)	(0.208)
Village emigration	0.271***	0.201***	0.255***	0.018	0.386***	0.231***
(indirect land channel)	(0.046)	(0.038)	(0.038)	(0.027)	(0.057)	(0.040)
Crop region emigration	0.184***	0.048	0.190***	0.007	0.300***	0.160***
(indirect crop channel)	(0.037)	(0.032)	(0.031)	(0.020)	(0.042)	(0.030)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	69744	68686	75320	63958	63958	66030
F-Stat on Direct Effect	89.5	84.6	80.8	79.7	79.7	75.2

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Estimation sample includes households that owned land in both periods as well as those who did not. All variables are standardized. *Male Migrants* is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with number of baseline working age males. *Village emigration* is the leave-one-out average number of working age male migrants in the village. *Crop region emigration* is the leave-one-out number of migrants *within the state* weighted by crop suitability. Columns 1 and 2 are indices for agricultural expenses and machinery, respectively (see Section 5.2 for index construction). Wage bill (column 4) is in Rupees and describes total expenses on hired labor in past year. Column 5 is total person-days of family and hired labor in past year. Crop profits are net of expenses (Column 6). All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table A13: Placebo Tests of Indirect Effects

	Crop Market Placebo		Land Market Placebo
	(1)	(2)	(3)
	Ag. Profits	Ag. Profits	Farm Size
Male Migrants	-1.347***	-1.325***	-0.842***
(direct labour channel)	(0.222)	(0.257)	(0.205)
Village emigration	0.248***	0.234***	0.136**
(indirect land channel)	(0.053)	(0.062)	(0.057)
Crop region emigration	0.249***		0.210***
(indirect crop channel)	(0.055)		(0.047)
Crop region emigration	-0.091**		
(placebo crop channel)	(0.039)		
Crop region emigration		0.227***	
(elasticity wt.)		(0.047)	
Crop region emigration		0.063*	
(placebo: inv. elasticity wt.)		(0.035)	
Village emigration			0.067
(placebo land channel)			(0.052)
HH FEs	✓	✓	✓
Year FEs	✓	✓	✓
Observations	25852	20886	16116
F-Stat on Direct Effect	67.9	49.7	33.5

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The outcome crop profits in columns 1-2 and farm size in column 3. All variables are standardized. *Male Migrants* is number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with number of baseline working age males. *Village emigration* is the leave-one-out average number of working age male migrants in the village. *Crop region emigration* is the weighted, leave-one-out number of migrants *within the state*. Column 1 adds a placebo where weights on *Crop region emigration* are not inverted. In column 2, weights on *Crop region emigration* are cross-price elasticities between the main crop grown by household i and i' and the placebo measure is the inverse elasticity. Column 3 adds a placebo measured as aggregate emigration among other non-agricultural households in the village. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

B Appendix Figures

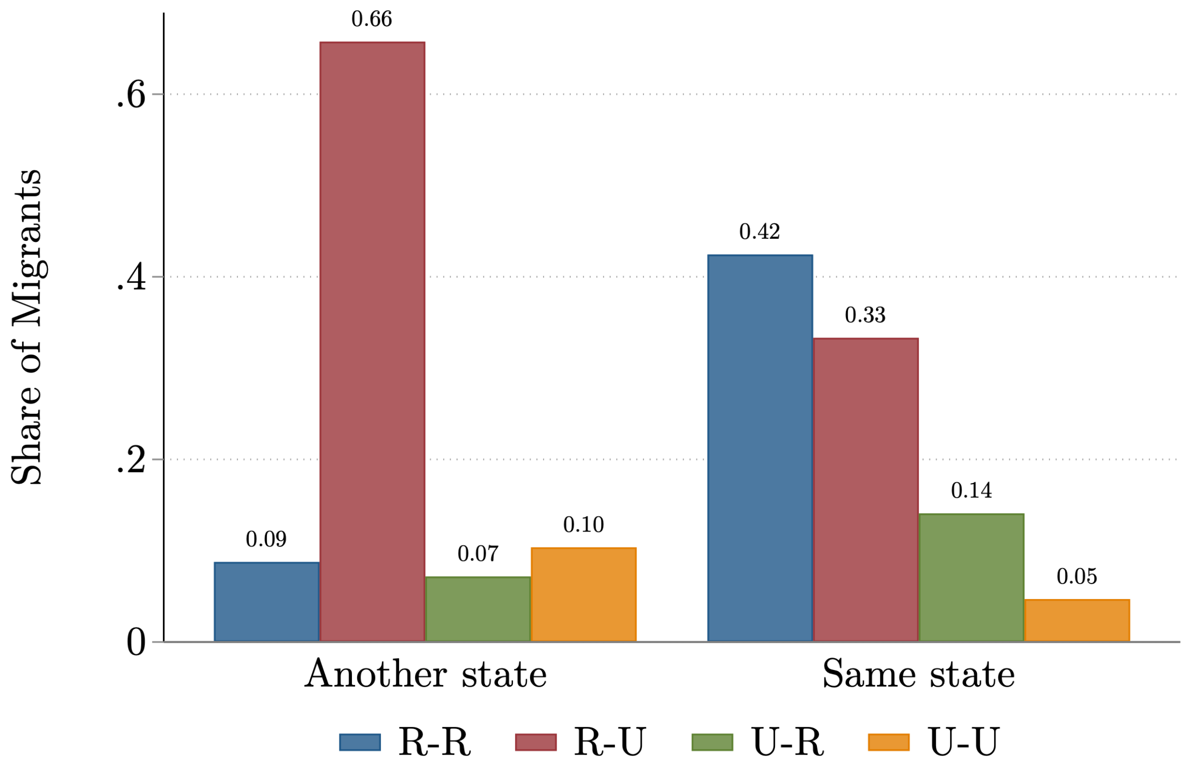


Figure B1: Migration Streams by Distance

Note: Data is at the migrant level. ”R-R” denotes migrants with rural origin and rural destination, ”R-U” denotes rural origin and urban destination, and so on.

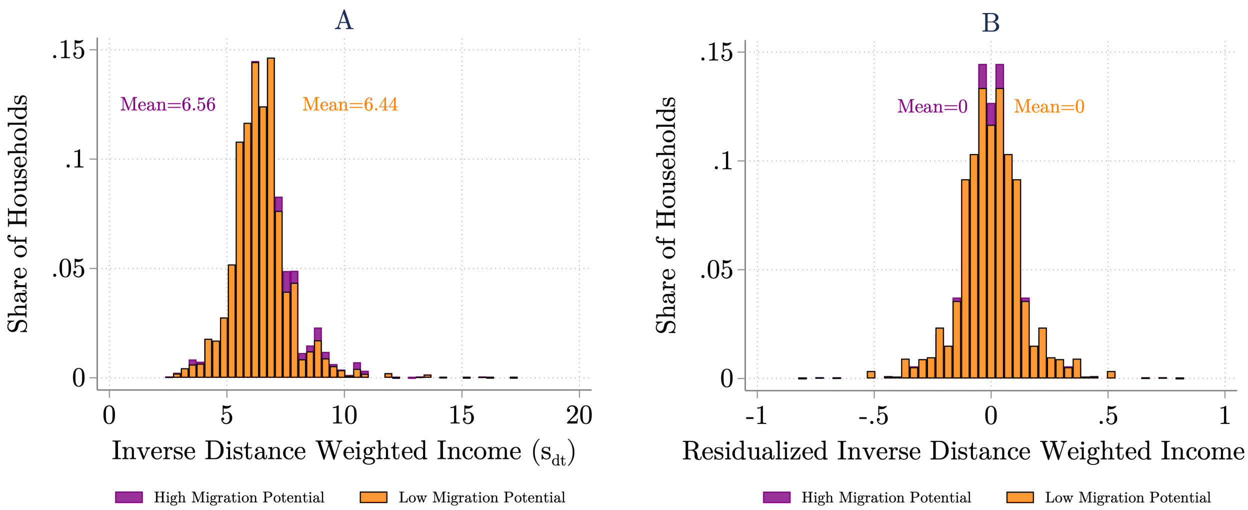


Figure B2: Distribution of Shocks

Note: Figures show the distribution of inverse-distance weighted income shocks (s_{dt}) for households with below median number males living at home during baseline (low migration potential) and above-median for the same (high migration potential). Panel A plots the raw data. Panel B plots values residualized on household and year fixed effects.

Online Appendix

C Conceptual Framework

This appendix develops a model of rural production and migration with market spillovers. We use the model predictions to motivate our estimation framework, including the instrument and spillover estimation. Although the model is adapted to our context, it builds on general spatial trade models (see Allen and Arkolakis (2023) for a recent overview).

C.1 Set-up

Environment: Consider an economy that comprises many villages indexed by $j \in [1, ..., J]$ and one urban center. Villages produce crops, while urban centers produce services. Each village is endowed with fixed land A_j and is inhabited by heterogeneous households. These households draw their crop mix from a vector of crops $k \in [1, ..., K]$ based on local crop suitability, ω_{jk}. Crop suitability is homogeneous within villages and uncorrelated with distance to the urban center. While this is a simplifying assumption, none of our results is driven by it. Multiple villages can be equally suitable for the same crops. The urban center uses labor from surrounding rural areas to produce the service goods. Space is characterized by distance to the urban center, which determines the opportunity cost of agricultural labor. Residents of remote villages face lower opportunity costs of agriculture than those in villages close to the urban center.

Markets: Rural labor markets are absent, while land, product, and technology markets are frictionless. We recognize that this is a strong abstraction from reality. Although none of our results hinge critically on this assumption, it allows us to make clear predictions. The labor market assumption further draws on the evidence of malfunctioning rural labor markets in India (Foster and Rosenzweig, 2022). The simplifying assumption of frictionless land, product, and technology markets allows us to focus on the impact of rural-to-urban migration on agriculture while abstracting from other market imperfections. Frictionless product markets equalize the crop price, p_k, across villages, similar to the assumptions in Borusyak et al. (2022b).[15] Crop prices are endogenously determined by aggregate supply within the economy.

We use the term land in a wider sense to include other immobile production factors such as irrigation water or immobile capital. Land prices, ρ_j, are endogenously deter-

[15] This assumption contradicts the agricultural trade barriers across Indian states. However, we treat each state in our empirical approach as an economy in this model to reconcile the theoretical assumptions with the empirical facts.

mined within villages and equal the marginal productivity of land. The marginal productivity of land increases with the total village agricultural labor, L_j, such that $\frac{\partial \rho_j}{\partial L_j} > 0$. Land markets equalize the marginal productivity of land across crops and households within one village.

Labor and Migration: Households supply one unit of labor inelastically, which they allocate to agriculture and the service sector. The allocation of rural labor to urban services describes rural-to-urban migration. We treat this process as a continuous rather than a binary decision in line with the large fraction of temporary migration and the migration of individual household members instead of whole households (Fact 1, Section 2.4).

Working in the service sector involves a migration cost, τ, which increases with distance to the urban center. This is a recurring cost since migrants eventually return to their origin. Similar to Morten and Oliveira (2018) and Pellegrina and Sotelo (2021), we assume iceberg migration costs (a fraction of the wage is lost) such that farmers earn net urban wage $\frac{w}{\tau}$. We assume $\tau \geq 1$, with $\tau = 1$ describing zero-cost migration since the farmer earns the full wage.

Agricultural labor productivity is homogeneous across households, but urban labor productivity, φ_{ij}, is heterogeneous. This generates heterogeneous responses to urban productivity shocks in line with the observed migration patterns.

Demand: We do not impose a functional form on the demand system to allow non-homothetic preferences per Engel's law. We assume downward-sloping demand for crop k that weakly increases in the total income of the economy and declines in aggregate supply of crop k:

$$\partial p_k / \partial Y \geq 0 \quad \text{and} \quad \partial p_k / \partial Y_k < 0$$

where Y is aggregate income and Y_k is aggregate supply of crop k. Unlike agricultural goods, demand for urban goods is perfectly elastic. This reflects India's highly restricted agricultural markets alongside an export-oriented service and manufacturing sector. The urban good is the numeraire throughout the model.

Agricultural Production: Production of crop k is an increasing and concave function of labor, land, and technology defined by:

$$y_{ijk} = \omega_{jk} f(l_{ijk}, a_{ijk}, \theta_{ijk})$$

where i indexes households, j indexes villages, and k indexes crops. l_{ijk} thus describes labor of household i in village j allocated to crop k. a_{ijk} denotes land and θ_{ijk} denotes technology in the same way. ω_{jk} is village j's agricultural suitability for crop k. We further assume that land increases the marginal productivity of labor and technology, i.e.

$$\frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk} \partial a_{ijk}} > 0 \tag{10}$$

$$\frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk} \partial a_{ijk}} > 0 \tag{11}$$

This assumption is built into the most common production functions, including Cobb-Douglas and CES. It also conveniently reduces the number of possible cases for the results. Unlike land, technology can be labor-saving or labor-complementary. It is labor-saving if it reduces the marginal productivity of labor and vice versa for labor-complementary technologies (Acemoglu, 2010).

Definition 1. *A technology is labor-saving if*

$$\frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk} \partial \theta_{ijk}} < 0.$$

It is labor-complementary if

$$\frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk} \partial \theta_{ijk}} > 0$$

To simplify the analysis further, we assume that the effect of technologies on labor is homogeneous across crops, i.e., a labor-saving technology is labor-saving for all crops, and a labor-complementary technology is labor-complementary for all crops.

Urban Production: Urban production is a constant returns-to-scale function with labor as the only input. Urban output is the numeraire such that the marginal productivity of effective urban labor is given by the wage, w. Urban productivity of households, φ_{ij}, converts household labor to effective urban labor units. Although households face the same urban wage, w, opportunity costs of agriculture are heterogeneous across households because of differences in φ_{ij}.

Household Problem: Rural household i in village j maximizes profits, taking prices as given:

$$\max_{a_{ijk}, \theta_{ijk}, l_{ijk}} \sum_{k=1}^{K} \left[p_k \omega_{jk} f(l_{ijk}, a_{ijk}, \theta_{ijk}) - \rho_j a_{ijk} - \nu \theta_{ijk} \right] + \frac{w \varphi_{ij}}{\tau_j} l_{ijw} \quad \text{s.t.} \quad \sum_{k=1}^{K} l_{ijk} + l_{ijw} = 1 \tag{12}$$

where p_k is the price of crop k, ρ_j is the land price in village j, and v is the exogenously given rental rate of capital. φ_{ij} is urban labor productivity of household i from village j,

w is the urban wage, $\tau_j \geq 1$ is the migration cost, which increases with distance, and l_{ijw} is the labor allocated to urban production. Optimality requires

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \geq \frac{w \varphi_{ij}}{\tau_j},$$
$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} \geq \rho_j,$$
$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \geq \nu,$$

where the conditions hold with equality for all crops produced by the household (all k with $l_{ijk}, a_{ijk}, \theta_{ijk} > 0$) and with inequality for all crops not produced by the household (all $k' \neq k$ with $l_{ijk'}, a_{ijk'}, \theta_{ijk'} = 0$).

C.2 Theoretical Predictions

Equipped with the model primitives, we now characterize the economy's optimal labor allocation and spatial distribution of production. We then show how urbanization affects this distribution. This section mainly provides intuition for our theoretical results. The proofs are in Appendix D.

C.2.1 Spatial Organization of the Economy

The share of agriculture in village-level production increases with remoteness. The mechanism is through declining opportunity cost of agricultural labor with distance to the urban center, $\frac{w\varphi_{ij}}{\tau_j}$. For labor complementary technologies, this also implies that more technology is used with increasing distance to the urban center. A formal proof of this result is provided in Appendix D.1.[16]

C.2.2 Urban Productivity Shocks and the Reorganization of Agriculture

We now use the model to ask how urban productivity shocks affect the spatial organization of agriculture. The total effect is composed of two forces: 1) a direct effect through labor reallocation and the adjustment of the other production factors, and 2) an indirect effect through aggregate emigration and the corresponding changes in crop and land prices. To characterize these effects, we first present two predictions about the effect of urban productivity shocks on agricultural labor:

[16]There are special cases with labor-saving technologies where the loss of agricultural labor close to urban centers could lead to more technology adoption and possibly also to increased agricultural production if the productivity increases from technology adoption overcompensate the labor losses.

Prediction 1. *The direct effect of a positive urban productivity shock on agricultural labor is negative. The magnitude of the effect declines with the distance to the urban center, and it increases with the urban productivity of the household.*

Proof. The proof is given in Appendix D.2. □

Intuitively, increasing urban productivity raises the opportunity cost of agriculture and prompts labor reallocation. Proximity to the urban center enhances this opportunity cost effect. We will use this to motivate the first stage of our estimation strategy in Section 3.2.

In the rest of the model, we define "rural-to-urban migration" as labor reallocation from agriculture to urban production in response to positive urban productivity shocks. We also call all villages in which crop k is grown the "crop-growing region" for crop k. The next prediction describes how the effect of migration differs based on the technology type:

Prediction 2. *The direct effect of rural-to-urban migration on household-level agricultural technology use is negative for labor-complementary technologies and ambiguous for labor-saving technologies. Aggregate village emigration increases household-level agricultural technology use, everything else equal. Aggregate crop-growing region emigration, which reduces aggregate production of crop k, increases household-level technology use for crop k, everything else equal.*

Proof. The proof is given in Appendix D.3. □

With these two predictions, we characterize the market-driven indirect effects of urban productivity shocks on production. The indirect effects of aggregate emigration on production are independent of distance to the urban center. Instead, they depend on aggregate emigration from the crop-growing region and on aggregate emigration from the same village. These are the boundaries within which crop and land prices are determined, respectively.

Urban productivity does not affect land and technology directly. It does so only through labor reallocation and corresponding changes in crop and land prices. We, therefore, describe the impact of urban productivity shocks on crop production through labor and price adjustments alone.[17]

[17] Urban productivity shocks may affect agricultural production through alternative channels. For example, rural households in developing countries are often credit-constrained (Karlan et al., 2014; Fink et al., 2020). Urban income shocks could relax these constraints and provide liquidity for technology investment (Veljanoska, 2022). We refer the reader to Rozelle et al. (1999) for a simple and intuitive model of liquidity constraints and migration.

Prediction 3. *The response of household crop production to urban income shocks can be decomposed into*

$$\frac{dy_{ijk}}{dw} = \frac{dl_{ijk}}{dw}\phi_{1ijk} + \frac{d\rho_j}{dw}\phi_{2ijk} + \frac{dp_k}{dw}\phi_{3ijk} \tag{13}$$

where ϕ_{1ijk}, ϕ_{2ijk} and ϕ_{3ijk} are composite coefficients defined in Appendix D.4.

Proof. The proof is given in Appendix D.4. □

The key insight of Prediction 3 is that urban productivity shocks trigger a direct effect (ϕ_{1ijk}) on crop production that may or may not be offset by an indirect land (ϕ_{2ijk}) or crop price effect (ϕ_{3ijk}) induced by aggregate emigration.[18] While $\phi_{2ijk} < 0$ and $\phi_{3ijk} > 0$, the sign of ϕ_{1ijk} possibly depends on whether the technology is labor-saving or labor-complementary (see Appendix D.4). ϕ_{1ijk} is positive for labor complementary technologies. Rural-to-urban migration that reduces agricultural labor and land prices and increases crop prices due to reduced crop production can, therefore, have opposing direct and indirect effects on agricultural production. The direction and magnitude of the individual effects become an empirical question.

D Proofs

D.1 Spatial Distribution of Agriculture

Here, we show the impact of distance to the urban center on crop production. Because there are many households in each location, the individual household has no market power and takes crop and land prices as given. Therefore, the first-order conditions for crop $k \in [1, ..., K]$ are given by

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} - \frac{w\varphi_{ij}}{\tau_j} = 0, \tag{14}$$

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} - \rho_j = 0 \quad \text{and} \tag{15}$$

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} - v = 0. \tag{16}$$

To simplify notation, we use subscripts in the following to denote partial derivatives.[19] We also drop crop, household and village subscripts because the conditions are identical across all crops, households and villages. Next, we fully differentiate the first-order

[18] In Appendix D.4, we show further that the effect of rural-to-urban migration on total household-level agricultural production is a linear combination of the crop-household-level responses.

[19] For example $f_l := \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}}$ and $f_{al} := \frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk} \partial a_{ijk}}$ etc.

conditions with respect to the distance to the urban center, τ:

$$p\omega(f_{ll}l_\tau + f_{la}a_\tau + f_{l\theta}\theta_\tau) + \frac{w\varphi}{\tau^2} = 0$$
$$p\omega(f_{al}l_\tau + f_{aa}a_\tau + f_{a\theta}\theta_\tau) - \rho_\tau = 0$$
$$p\omega(f_{\theta l}l_\tau + f_{\theta a}a_\tau + f_{\theta\theta}\theta_\tau) = 0$$

Since we have assumed no trade costs for goods, crop prices are constant across space. However, the land price changes in response to the local labor allocation decisions. Solving the equations yield

$$l_\tau = \left[\frac{w\varphi}{\tau^2}(f_{aa}f_{\theta\theta} - f_{\theta a}^2) + \rho_\tau(f_{la}f_{\theta\theta} - f_{l\theta}f_{\theta a})\right] D^{-1}$$
$$a_\tau = \left[\frac{w\varphi}{\tau^2}(f_{l\theta}f_{\theta a} - f_{la}f_{\theta\theta}) + \rho_\tau(f_{l\theta}^2 - f_{ll}f_{\theta\theta})\right] D^{-1}$$
$$\theta_\tau = \left[\frac{w\varphi}{\tau^2}(f_{la}f_{\theta a} - f_{l\theta}f_{aa}) + \rho_\tau(f_{ll}f_{\theta a} - f_{la}f_{l\theta})\right] D^{-1}$$
$$D := \omega p(f_{ll}f_{\theta a}^2 - f_{ll}f_{aa}f_{\theta\theta} + f_{la}^2 f_{\theta\theta} - 2f_{la}f_{l\theta}f_{\theta a} + f_{l\theta}^2 f_{aa}) > 0$$

The denominator, D is the determinant of the Hessian matrix times $(-\omega p)$. It is positive because of the concavity assumption of the production function. The direction of the effect is, therefore, determined by the enumerator.

The total effect of distance on agricultural inputs is composed of a direct effect and an indirect effect through the response of the land price to labor reallocation to urban production. For individual farmers with low urban productivity, the indirect effect can dominate (e.g. for $\varphi = 0$). In aggregate, the indirect effect cannot dominate because it is the response to the direct effect.

Crop-level and household-level responses: l_τ, a_τ, and θ_τ describe the crop- and household-specific responses of agricultural inputs to distance from the urban center. The household-level response is the sum of crop-level responses. Crop-specific and household-specific results are analogous because we assume homogeneous effects of technology on labor across crops.

Labor: The direct effect of distance on agricultural labor is positive because of the concavity assumption. Less labor migrates to the urban center with increasing distance to the urban center, everything else equal. The land price increases in response. The indirect effect is, therefore, primarily negative except for special cases with strong effects of labor-saving technology i.e. $f_{la}f_{\theta\theta} > f_{l\theta}f_{\theta a}$.

Land: The amount of land per household is constant if we abstract from corner solutions of labor reallocation i.e. no complete household migration. This follows from the assumption of constant land endowments per village. However, there may be land reallocation among households as households with higher urban productivity are more affected by the distance to the urban center. To see this, set $\varphi = 0$, which reduces the direct effects of labor allocation in response to the distance to the urban center to zero.

Technology: The direct effect of distance on technology is positive except for cases with labor-saving technology when it is ambiguous. The indirect effect is negative because of the concavity assumption and the Assumptions (10) and (11) ($f_{\theta a} > 0, f_{la} > 0$).

Production: The value of total production in village j is given by the sum of the values of individual production y_{ijk} i.e.

$$Y_j = \sum_i \sum_k y_{ijk}.$$

The aggregate production changes in response to distance according to

$$\begin{aligned} \frac{dY_j}{d\tau} &= \sum_i \sum_k \frac{dy_{ijk}}{d\tau} \qquad (17) \\ &= \sum_i \sum_k p_k \omega_{jk} \left(\frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \frac{\partial l_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} \frac{\partial a_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \tau} \right) \end{aligned}$$

The partial derivative of production with respect to labor, land, and technology is positive. The total effect is, therefore, determined by the adjustments of labor, land, and technology to distance.

Further, note that land markets equalize the marginal productivity of land across households and crops such that $p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} = \rho_j$. We can therefore write

$$\sum_i \sum_k p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} \frac{\partial a_{ijk}}{\partial \tau} = \rho_j \sum_i \sum_k \frac{\partial a_{ijk}}{\partial \tau} = 0$$

since the total village land, A_j, is constant and therefore $\sum_i \sum_k \frac{\partial a_{ijk}}{\partial \tau} = 0$.

Equation (17), therefore, simplifies to

$$\frac{dY_j}{d\tau} = \sum_i \sum_k p_k \omega_{jk} \left(\frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \frac{\partial l_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \tau} \right) \qquad (18)$$

For labor complementary technologies, we get

$$\frac{dY_j}{d\tau} = \sum_i \sum_k p_k \omega_{jk} \left(\frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \frac{\partial l_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \tau} \right) > 0$$

D.2 Proof of Prediction 1

Here, we show the effect of urban productivity shocks on agricultural labor. In the following, we use subscripts to denote partial derivatives and omit household, village, and crop subscripts (see Appendix D.1). First, we fully differentiate the first-order conditions with respect to the urban wage:

$$p_w \omega f_l + p\omega(f_{ll} l_w + f_{la} a_w + f_{l\theta}\theta_w) - \frac{\varphi}{\tau} = 0, \tag{19}$$

$$p_w \omega f_a + p\omega(f_{al} l_w + f_{aa} a_w + f_{a\theta}\theta_w) - \rho_w = 0, \tag{20}$$

$$p_w \omega f_\theta + p\omega(f_{\theta l} l_w + f_{\theta a} a_w + f_{\theta\theta}\theta_w) = 0. \tag{21}$$

We then solve these equations for l_w, a_w, and θ_w:

$$\begin{aligned}
l_w =& \{\frac{\varphi}{\tau}(f_{\theta a}^2 - f_{aa} f_{\theta\theta}) \\
&+ \rho_w(f_{la} f_{\theta\theta} - f_{l\theta} f_{\theta a}) \\
&- p_w\omega[f_l(f_{\theta a}^2 - f_{aa} f_{\theta\theta}) + f_a(f_{la} f_{\theta\theta} - f_{l\theta} f_{\theta a}) + f_\theta(f_{l\theta} f_{aa} - f_{la} f_{\theta a})]\}D^{-1} \\
\theta_w =& \{\frac{\varphi}{\tau}(f_{l\theta} f_{aa} - f_{la} f_{\theta a}) \\
&+ \rho_w(f_{ll} f_{\theta a} - f_{l\theta} f_{la}) \\
&+ p_w\omega(f_l f_{la} f_{\theta a} - f_l f_{l\theta} f_{aa} + f_a f_{la} f_{l\theta} - f_{ll} f_a f_{\theta a} - f_\theta f_{la}^2 + f_{ll} f_\theta f_{aa})\}D^{-1} \\
a_w =& -\{\frac{\varphi}{\tau}(-f_{la} f_{\theta\theta} + f_{l\theta} f_{\theta a}) \\
&+ p_w\omega(f_l f_{la} f_{\theta\theta} - f_l f_{l\theta} f_{\theta a} + f_a f_{l\theta}^2 - f_{ll} f_a f_{\theta\theta} - f_\theta f_{la} f_{l\theta} + f_{ll} f_\theta f_{\theta a}) \\
&+ \rho_w(f_{ll} f_{\theta\theta} - f_{l\theta}^2)\}D^{-1} \\
D :=& \omega p(f_{ll} f_{\theta a}^2 - f_{ll} f_{aa} f_{\theta\theta} + f_{la}^2 f_{\theta\theta} - 2 f_{la} f_{l\theta} f_{\theta a} + f_{l\theta} 2 f_{aa}) > 0
\end{aligned}$$

The denominator, D, is the determinant of the Hessian matrix times $(-\omega p)$. This is positive because of the concavity assumption. The numerator, therefore, determines the direction of the effect. The numerator comprises a direct and indirect effect through the product and land price channels.

Here, we focus on l_w and report the effects of urban productivity on land and technology for the remaining proofs. The direct effect of urban wages on agricultural labor is given by $\frac{\varphi}{\tau}(f_{\theta a}^2 - f_{aa} f_{\theta\theta}) < 0$. It is negative because the production function is concave. Partial differentiation shows that it increases with the urban productivity of the household and declines with the distance to the urban center. The results are for crop-household-level agricultural labor. Changes in household-level agricultural labor are the sum of changes in crop-household-level agricultural labor. The results directly apply to household-level agricultural labor because the direction of crop-specific labor responses to urban productivity shocks is homogeneous across crops.

D.3 Proof of Prediction 2

Here, we show how agricultural technology use responds to labor reallocation. To do so, we express the changes in technology and land as a function of labor adjustments. We solve the first and the last of the fully differentiated first-order condition (see Appendix D.2) for a_w and θ_w as a function of l_w :

$$\begin{aligned} a_w =&(l_w \omega p(f_{l\theta} f_{\theta a} - f_{la} f_{\theta\theta}) + p_w \omega(f_\theta f_{\theta a} - f_a f_{\theta\theta}) + \rho_w f_{\theta\theta}) D_2^{-1} \\ \theta_w =&(l_w \omega p(f_{la} f_{\theta a} - f_{l\theta} f_{aa}) + p_w \omega(f_a f_{\theta a} - f_\theta f_{aa}) - \rho_w f_{\theta a}) D_2^{-1} \\ D_2 :=&(f_{aa} f_{\theta\theta} - f_{\theta a}^2)\omega p > 0 \end{aligned}$$

D_2 is the determinant of the second-order minor. Here, we focus on θ_w and report the result for a_w only to support the proof in Appendix D.4. The direct effect is given by $l_w \omega p(f_{l\theta} f_{\theta a} - f_{la} f_{\theta\theta})$. The first part, $l_w \omega p$, is negative (see Prediction 1). The second part is ambiguous. To see this, note that $f_{la} f_{\theta\theta} < 0$ because of the concavity assumption and assumption (10). Further, note that $f_{l\theta} f_{\theta a} > 0$ for $f_{l\theta} > 0$ and $f_{l\theta} f_{\theta a} < 0$ for $f_{l\theta} < 0$ because of assumption (11).

Define household-level technology use as $\bar{\theta}_{ij} := \sum_{k=1}^K \theta_{ijk}$ and therefore $\frac{d\bar{\theta}_{ij}}{dw} = \sum_{k=1}^K \frac{d\theta_{ijk}}{dw}$. If $\frac{d\theta_{ijk}}{d_w} > (<)0$, then also $\frac{d\bar{\theta}_{ij}}{dw} > (<)0$ because of the assumption that the effect of technology on labor is homogeneous across crops i.e., the direction of the technology response to urban productivity shocks is the same for all crops. The same argument applies to the remainder of the proof.

The effect of village emigration on technology use is determined by $-\rho_w f_{\theta a}$. Aggregate village emigration reduces land prices by assumption (see Section C) such that $\rho_w < 0$. Assumption (11) completes the proof for the second statement of Prediction 2.

Aggregate crop-growing region emigration, which reduces aggregate production of crop k, increases the crop price by assumption (see Section C) such that $p_w > 0$. The direction of this indirect effect is therefore determined by $\omega(f_a f_{\theta a} - f_\theta f_{aa})$. $f_a f_{\theta a} > 0$ by assumption (positive marginal productivity and assumption (11)) while $f_\theta f_{aa} < 0$ by assumption (positive marginal productivity and concavity) which completes the proof for the last statement of Prediction 2.

D.4 Proof of Prediction 3

Here, we decompose the impact of urban productivity shocks on household crop production. We use subscripts to denote partial derivatives instead of crop, household, and village indices. Using the results and definitions from Appendix D.3 we can express the response of household crop production to urban productivity shocks as

$$\begin{aligned}
y_w =& f_l l_w + f_a a_w + f_\theta \theta_w \\
=& f_l l_w \\
&+ f_a (l_w \omega p (f_{l\theta} f_{\theta a} - f_{la} f_{\theta\theta}) + p_w \omega (f_\theta f_{\theta a} - f_a f_{\theta\theta}) + \rho_w f_{\theta\theta}) D_2^{-1} \\
&+ f_\theta (l_w \omega p (f_{la} f_{\theta a} - f_{l\theta} f_{aa}) + p_w \omega (f_a f_{\theta a} - f_\theta f_{aa}) - \rho_w f_{\theta a}) D_2^{-1} \\
=& l_w \left[f_l + \frac{\omega p (f_{l\theta} f_{\theta a} f_a - f_{la} f_{\theta\theta} f_a + f_{la} f_{\theta a} f_\theta - f_{l\theta} f_{aa} f_\theta)}{D_2} \right] \\
&+ \rho_w \frac{f_{\theta\theta} - f_{\theta a}}{D_2} \\
&+ p_w \frac{\omega (f_\theta f_{\theta a} - f_a f_{\theta\theta} + f_a f_{\theta a} - f_\theta f_{aa})}{D_2}
\end{aligned}$$

Define

$$\begin{aligned}
\phi_{1ijk} &:= \left[f_l + \frac{\omega p (f_{l\theta} f_{\theta a} f_a - f_{la} f_{\theta\theta} f_a + f_{la} f_{\theta a} f_\theta - f_{l\theta} f_{aa} f_\theta)}{D_2} \right] \\
\phi_{2ijk} &:= \frac{f_{\theta\theta} - f_{\theta a}}{D_2} < 0 \\
\phi_{3ijk} &:= \frac{\omega (f_\theta f_{\theta a} - f_a f_{\theta\theta} + f_a f_{\theta a} - f_\theta f_{aa})}{D_2} > 0.
\end{aligned}$$

These composite parameters are unaffected by marginal changes in urban productivity. We added subscripts to underline that they are household-crop-specific. For labor complementary technologies, ϕ_{1ijk} is positive, while it is ambiguous for labor-saving technologies.

The results are for crop-household combinations. To derive household-level results, we reintroduce household, village, and crop subscripts. Total household agricultural production is given by

$$\bar{y}_{ij} = \sum_{k=1}^{K} y_{ijk}$$

and the response of household production to urban productivity shocks by

$$\frac{d\bar{y}_{ij}}{dw} = \sum_{k=1}^{K} \frac{dy_{ijk}}{dw} = \sum_{k=1}^{K} \frac{dl_{ijk}}{dw} \phi_{1ijk} + \frac{d\rho_j}{dw} \sum_{k=1}^{K} \phi_{2ijk} + \sum_{k=1}^{K} \frac{dp_k}{dw} \phi_{3ijk}.$$