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

October 17, 2025

Abstract

How does urban growth and migration reshape rural agriculture during the process of structural transformation? Using household microdata from India, we show that agricultural households do not offset the loss of labor from emigration by increasing capital investment. Instead, they cultivate less land and lower their use of agricultural technology, reducing crop production. Resulting changes in land and crop prices induce non-migrant households to expand agricultural investments and production. In aggregate, market adaptation mitigates over three-fourths of the direct agricultural losses from urbanization. This results in a spatial reorganization of agriculture where food production moves from land near urban areas where emigration is high, toward remote areas where emigration is low.

Keywords: Internal migration, agriculture, structural transformation, India
JEL Codes: O13, O15, Q15, R11, J43

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

†Department of Applied Economics, University of Minnesota; rmadhok@umn.edu
‡Food and Resource Economics Group, University of British Columbia; frederik.noack@ubc.ca
§Yale University and Deakin University; ahmed.mobarak@yale.edu
¶Department of Economics, UC Santa Barbara; olivier@econ.ucsb.edu

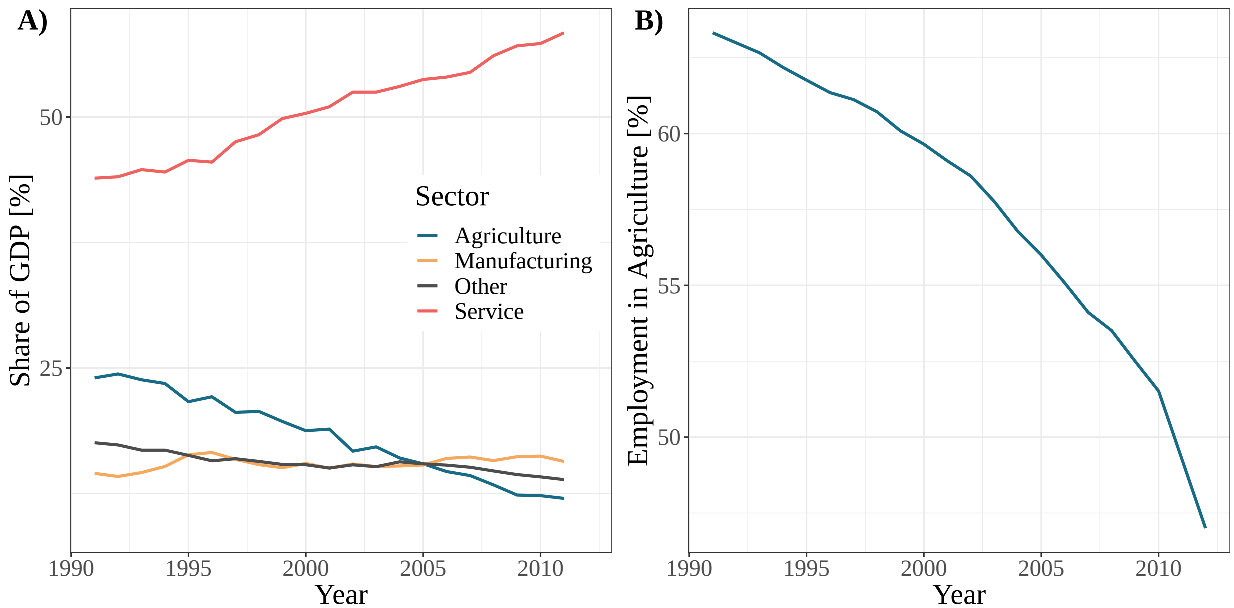
1 Introduction

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 saw rural workers moving *en masse* to urban factories (Gao et al., 2022). Likewise, India's 1991 economic liberalization was followed by rapid growth in the service sector (Figure 1A, red) and massive exodus from agriculture, with the agricultural workforce shrinking by more than 30% over the next two decades (Figure 1B). This raises the central question of this paper: how does urban economic growth reshape the rural agricultural sector in the face of such massive agricultural labor loss?

In laying the foundation for development economics, the Lewis (1954) "dual sector model" assumed that rural areas have unlimited surplus labor, enabling rural-urban migration and structural transformation without reducing agricultural output. In reality, emigration from rural areas reduces agricultural labor and crop output, as documented in subsequent empirical research (Rozelle et al., 1999; Mendola, 2008). Maintaining domestic food production as labor exits agriculture thus requires some form of agricultural transformation. County- and state-level analyses in the historic United States suggest that agricultural modernization occurs following rural labor loss as farmers adopt new technologies and mechanize (Manuelli and Seshadri, 2014; Hornbeck and Naidu, 2014; Alvarez-Cuadrado and Poschke, 2011). But these aggregated data are not fine enough to uncover the precise mechanism by which the transformation occurs. Do the families of farmers who emigrate directly substitute capital for labor? Or is it a more complex process in which factor markets readjust following their departure, and other non-migrant farming households in the same village or in other regions alter their farm operations in response to new factor prices and fill the agricultural production gaps?

While the literature has mainly focused on migrants and their destinations, we answer these questions by studying the dynamics of agricultural transformation in left-behind rural areas. We use household-level panel data where we observe the migration and cultivation decisions of 42,000 households in India as the country underwent large-scale transformation between 2005 and 2012. These data allow us to paint a richer, more nuanced picture of the agricultural investment decisions of rural farming households facing new urban migration opportunities. They also capture the subsequent readjustments of other households residing in the same village, or producing the same crops elsewhere in India, who are *indirectly* affected through general equilibrium (GE) channels. We design

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.

a plausibly exogenous shift-share instrument for migration and also develop a research design to track indirect land and crop market adjustments.

We offer three new insights on how agricultural development unfolds during structural transformation. First, contrary to the conventional wisdom developed in the historical US-based literature, farming households facing new migration opportunities in nearby cities *do not* mechanize; instead, they downsize. Second, we show that farm households without migrants respond to aggregate emigration and the resulting land and crop market adjustments by expanding production and increasing agricultural capital investment. Third, in aggregate, these direct and indirect effects spatially reallocate agricultural activity away from rural areas near cities (where migration opportunities are plentiful), and toward more remote locations (where migration opportunities are scant). India's structural transformation was thus accompanied by a *spatial reorganization* of agriculture. We quantify the net effect on food supplies and find that indirect GE effects through land and crop markets recover roughly 80% of the direct loss in agricultural output from households that emigrate and downsize their farms.[1]

The development literature has studied a wider variety of agricultural responses to new migration opportunities, including pollution (Garg et al., 2024), tree cover loss (Brewer

[1]By aggregating our data from the household to the district level, we can replicate the mechanization results in the existing literature (Section 6.1). But our household data shows how and where this happens.

et al., 2022), consumption volatility and crop choice (Kinnan et al., 2018), and technology choice (Mendola, 2008). Research in some developing contexts finds, like we do, that labor loss reduces cultivated area (Rozelle et al., 1999; Brewer et al., 2022), but others document land consolidation and mechanization with no loss to cultivation (De Janvry et al., 2015; Tian et al., 2023). Our contribution is develop a unified framework to estimate both direct and indirect general equilibrium effects on agriculture, and outline its spatial reorganization at the macro scale, for the case of India – the largest developing country – as it underwent a massive structural transformation.

The key challenge to identifying the effects of migration on agricultural development is that both processes are co-determined with origin-based push factors (Foster and Rosenzweig, 2007). We address this with a shift-share instrument where the "shift" consists of income shocks at potential migration destinations. The "share" comprises two variables: 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. We therefore measure household migration potential by the number of working-age males at home during the baseline period. Interacting 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.[2]

We find no evidence of labor-capital substitution within households that directly benefit from migration opportunities. Instead, migrant-sending households invest *less* in agricultural technologies, such as agrochemicals, irrigation water, and work animals. This, in turn, drives a reduction in crop profits and farm size. 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 documents indirect responses to labor migration by investigating equilibrium effects on agricultural investment and profits as markets react to labor reallocation. Guided by a spatial equilibrium model of migration and agricultural production, we hypothesize that declining output and farm size through the direct chan-

[2] We validate our empirical strategy using recent advances in the shift-share literature (Borusyak et al., 2025). The concern is that households with many working-aged men are systematically different than those with few. Yet Borusyak et al. (2022a) show that shift-share instruments remain valid as long as the shift is quasi-randomly assigned. We present a two-period orthogonality test and find that the second period distance-weighted income shock is orthogonal to first period outcomes, implying that systematic differences between "exposed" and "non-exposed" households will not be picked up by the instrument since the shift component is plausibly random with respect to baseline outcomes.

nel trigger higher crop prices and lower land prices. Remote households with fewer migration opportunities benefit from these market spillovers by expanding farm activity. A key insight from the spatial general equilibrium literature 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.[3] 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 prime-age males 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 that set of crops. The crop market adjustment is measured as 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.

We find that farmers from households that do not send migrants experience *increased* crop profits in response to crop market adjustments, partially offsetting the migration-induced losses from the direct channel. We build a district crop price index from ICRISAT to test mechanisms and find that mass emigration from the crop market increases prices of those same crops, 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.

On the land market side, we show that other farmers in the village who are less exposed to the migration shock *expand* their farm size in response to land market adjustments, again counteracting the downsizing of farms among migrant-sending households. We use land transaction data from IHDS to test mechanisms, and find that emigration from the land market depresses land prices. "Left-behind" farmers in the village exploit this opportunity to expand cultivation.

In this expanded framework, the direct effect remains instrumented with the shift-share variable, while indirect crop and land market adjustments are uninstrumented. This raises concerns because village-level shocks (e.g., drought, local agricultural policy) can

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

co-determine village emigration as well as household agricultural decisions. Similarly, crop-specific shocks (e.g., crop disease, trade disruptions) can affect both crop-region emigration and household agriculture. We thus deploy a variety of robustness checks to support the credibility of our indirect crop and land market estimates.

First, the IHDS village module records environmental disasters and government programs for all 1400 villages in our sample. Our estimates of indirect land market effects are robust when directly controlling for village-level shocks like droughts and agricultural extension activity. Second, we link each household's main crop to annual harvest prices and find highly stable estimates of crop market adjustments when controlling for crop price shocks. While this may appear to absorb the price mechanism itself, our identification strategy relies on deviations from *general* crop price trends induced by supply contractions due to aggregate emigration from the crop market. Third, we design placebo tests to validate our GE estimates: household farm size does not respond to aggregate emigration by *non-agricultural* households in the village, suggesting that our land market measure is not picking up other non-land factor market changes. Households also do not respond to aggregate emigration from outside the crop market, reinforcing that our crop market measure captures adjustments specific to the relevant crop's supply region.

Our GE results embed an important spatial component: the (negative) direct effect of migration on agricultural development dominates for households near cities that face low migration costs. The (positive) market-driven effects dominate for households further away who face higher migration costs. Although these remote households do not participate directly in migration, they still contribute to the process of structural transformation by producing more food. This leads to a spatial reorganization of agriculture from migrant-sending areas toward remote areas.

To test whether this spatial feature of our findings appear in alternative data, we use both geo-coded census data on economic growth, agricultural labor, as well as high-resolution satellite data on crop output from across India. At the district level, we show that agricultural labor declines near high-growth cities, while crop output rises in remote districts, mirroring our GE estimates of spatial reallocation through crop markets. Our high-resolution satellite maps show that yields decline in peri-urban zones of a city with high labor outflows, and rise in the outer fringes, consistent with our GE estimates of spatial reallocation through local land markets. The consistent patterns we observe across both our econometric estimates and descriptive maps lend strong support to our central claim that structural transformation reshapes the geography of agriculture.

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 necessarily 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 aggregate 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 losses due to emigration. These results do not necessarily 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.

Micro-Development Contributions For the micro-development literature that studies the effects of labor loss on agriculture, our contribution is to decompose agricultural responses to labor reallocation into direct (labor-driven) and indirect (market-driven) channels in a unified framework. Other 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 enables us to compare the two forces at the aggregate level, and then estimate how much of the agricultural losses are compensated through market adaptation.

In doing so, we also 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. 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 into the U.S. 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 separately quantifying direct and indirect effects, our micro data also allow us to paint a richer picture of the post-migration adjustment process.[4]

Macro-Development Contributions Our paper uses household microdata and applied microeconomic techniques to also elaborate long-standing issues in macro-development. 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 advance this work with household data to 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, where lower trade costs leads to fewer firms, and an indirect channel where cheaper agricultural imports releases agricultural labor leading to more firms. While we also find that direct and indirect effects move in opposite directions, we study impacts on farms rather than firms.

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.

The next section describes the data and presents stylized facts about migration in India. Section 3 outlines the shift-share 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 contextualizes our findings and conducts back-of-the-envelope calculations to estimate changes in aggregate 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 three stylized facts about migration in India. These facts motivate the empirical strategy in Section 3.

[4]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.

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[5]. 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, which is reported in both survey rounds.

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 members. The IHDS definition thus characterizes longer-term spells[6], 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 age window, Table A1 shows migrant types by age group. The share of male migrants who leave for education begins a persistent drop after age 14, the same age at which male migration for work jumps five-fold. This suggests that migrant males transition from school to work around age 15. Similarly, the share of male labor migrants drops after age 59, suggesting they stop migrating around this age.

[5]Except Andaman and Nicobar Islands and Lakshadweep, which contain $<$ 1% of the population.
[6]Short-term ($<$6 months) circular migration is only documented in Wave II.

2.2.2 Agricultural Activity

Farming households report agricultural capital, labor, and crop profits. 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 the total wage bill 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[7].

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 each one by its standard deviation for 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.

We use crop profits (revenue minus expenses, in 2005 prices) as a proxy for output since profits are reported in both surveys[8], whereas production volume is only reported in the second round. For subsistence crops, profits are based on prices that farmers report they *would have* received at market. Note that aggregate crop price changes are absorbed by year fixed effects (or state-year fixed effects in robustness checks) in our regressions, in which case impacts can be interpreted as output changes despite being measured in rupees. However, since crop-specific prices may trend differently across states, we are careful about using the term "crop output" and "crop profits" interchangeably.

2.2.3 Weather Covariates

Droughts, temperature, and rainfall are the key covariates in our analysis. Accounting for climate is crucial because it affects 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. Gridded annual temperature (°C) and rainfall (mm) are

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

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

Table 1: Summary Statistics: Migrant Profiles

	# Migrants	Share	SD
A: Gender			
Male	13994	0.79	0.41
Female	3656	0.21	0.41
B: Status			
Student	4755	0.27	0.44
Working	11928	0.68	0.47
Neither	967	0.05	0.23
C: Destination			
Within State	11426	0.65	0.48
Out of State	6224	0.35	0.48
D: Stream			
Rural-Rural	5395	0.31	0.46
Rural-Urban	7898	0.45	0.50
Urban-Rural	2061	0.12	0.32
Urban-Urban	1182	0.07	0.25

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.

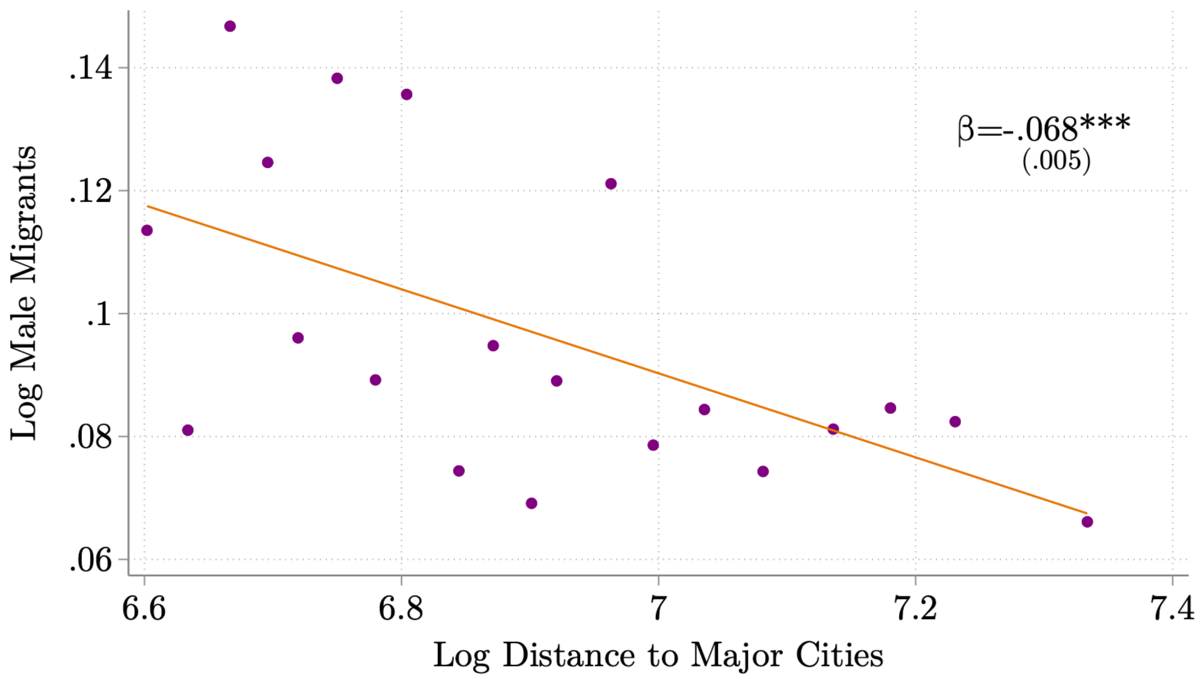
from the ERA5 product on a 0.125°× 0.125°grid (Hoffmann et al., 2019). For each covariate, we extract the mean over cells within districts and then compute annual averages for 2005 and 2012 to match with the IHDS.

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 1 profiles the typical migrant: 80% are male, and nearly 70% are labor migrants, supporting 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 1 Panel C). Interestingly, rural-rural migration accounts for a large share of migration. Figure B1 splits migration streams by inter- and intra-state travel, revealing an interesting pattern: 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.

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.

2.4 Three Stylized Facts About Migration in India

Next, we describe three stylized facts from the data that motivate our empirical strategy. The first fact is that entire-family migration is rare. Table A2 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: on a log-log scale, households twice as remote (100 percent increase) send 7% fewer migrants. The inverse correlation is robust when measuring remoteness as distance to the nearest large city. This pattern provides evidence that migration costs increase with distance. We incorporate this fact into our instrument for migration in Section 3.2, which leverages household exposure to destination income shocks where exposure declines with distance.

The third fact is that agricultural labor constraints increase in farm size. This fact is supported by previous work showing that small farms in India tend to have surplus labor (Foster and Rosenzweig, 2022). Table A3 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. The fact that large farms are more labor-constrained suggests that the impact of labor loss will be more salient on these farms (Section 4.2). Table A4 provides additional summary statistics on land, labor, and capital.

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 or weather. Household fixed effects, α_i, absorb time-invariant differences between households, such as distance to cities or land quality, both of 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 farming decisions. Improved agricultural technology can also release surplus labor, leading to reverse causality. Our shift-share instrument, which we discuss next, mitigates these threats to identification.

3.2 Shift-Share Instrument Design

3.2.1 Overview

We construct an instrument for household migration that combines plausibly exogenous urban productivity shocks with household-specific exposure to these shocks. This setup is guided by a conceptual model provided in Appendix C. Urban productivity shocks include shocks to the service and manufacturing sectors (e.g., the rollout of telecoms) which are plausibly orthogonal to agricultural production. Exposure to these shocks is measured by (i) distance to the affected urban center, which proxies for migration costs, and (ii) the baseline number of adult male household members, which captures migration capacity given that migration in India is predominantly male (see Section 5.3).

The interaction of these two components generates variation in migration incentives that is independent of local or household-level shocks. While household composition could, in principle, influence agricultural investment directly, we show that it is uncorrelated with exposure to urban productivity shocks, ensuring that the instrument does not capture differential investment trends. Together with household and time fixed effects, this design exploits variation in both migration potential and migration costs across households and villages to identify the causal effect of migration.

3.2.2 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 1).

Unlike previous studies, we use two "shares" to 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 $\frac{1}{\tau_{dd'}}$[9], where τ is distance. Since

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

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 1). We use a distance elasticity of one based on similar values from the literature (Bryan and Morten, 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. We measure households' urban productivity potential in this way becaue internal migration in India is overwhelmingly male-dominated (see Section 2.3), implying the existence of migration barriers that prevent women from fully benefiting from urban opportunities.

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

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

3.2.3 Instrument Validity

z_{idt} is a valid instrument if: 1) it strongly predicts labor migration, M^{labor}_{idt}, and, 2) it fulfills the exclusion restriction. We test the first criteria with the first stage equation in 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 discuss potential violations through each component of z_{idt}.

The first concern is that peri-urban households may face different input and output markets compared to remote households, opening "backdoor" channels through which z_{idt} can affect household agricultural decisions. Likewise, urban income shocks can increase aggregate food demand and, therefore, household crop production through output prices, especially for peri-urban households. We address these concerns in two ways: first, our fixed effects absorb time-constant differences between households close and far from urban centers. Second, and more importantly, we directly control for the inverse-distance weighted income shock, $s_{dt} := \sum_{d' \in \Theta / d} \frac{inc_{d't} \cdot pop_{d'}}{\tau_{dd'}}$, in both the first and second stage, such that identification only relies on heterogeneous household *exposure* to the shock through differences in their urban productivity potential, φ_i.

The second concern is that households with more working-age male members at baseline, our measure of φ_i, may be on a different agricultural development trajectory than those with few working-age male household members. However, this only threatens instrument validity if φ_i is correlated with the second part of the instrument, s_{dt}. Otherwise, the instrument would not pick up the differential trend. Next, we show evidence

Table 2: Orthogonality Between Second Period Shocks and First Period Outcomes

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Distance Wt. Income (t=2012)	0.079	0.063	-0.187	14.390	331.990
	(0.081)	(0.043)	(0.118)	(14.766)	(733.091)
Controls	Yes	Yes	Yes	Yes	Yes
State FEs	✓	✓	✓	✓	✓
Observations	16874	15812	18213	15208	36907
R^2	0.250	0.192	0.150	0.139	0.043

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Data are a cross section of households. Outcomes are for 2005 and explanatory variables for 2012. Column 1 and 2 are indices of agricultural technology expenses and machinery ownership. Column 3 is farm size in acres. Column 4 person-days of labor (hired + family). Column 5 is crop revenue (in Rs.) net of expenses. All regressions include state fixed effects and control for district income and number of household working-age males. Standard errors clustered by state.

that supports instrument validity despite potentially endogenous shares.

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, 2025). Figure B3A 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 differential trends of households with a large number of working-aged male household members.

Second, we formally test for whether the shock is as-good-as-randomly assigned. With our two period panel, we conduct a balance test by regressing a set of baseline agricultural outcomes, Y_{idt_1}, on the second-period shock, s_{dt_2}, in a pooled cross-section:

$$Y_{idt_1} = \phi \cdot s_{dt_2} + \Gamma X'_{dt_2} + \gamma_s + \varepsilon_{ijdt} \tag{2}$$

where ϕ is the balance coefficient of interest. If s_{dt_2} is as-good-as-randomly assigned, it should not predict Y_{idt_1} and thus $\phi = 0$. X'_{dt_2} is a covariate vector including the number of household males and district income. γ_s is a state fixed effect. Table 2 reports estimates of ϕ from regressions with agricultural technology, farm size, labor-days, and profits as outcomes. The second period distance-weighted income shock, s_{dt_2}, is uncorrelated with all baseline outcomes, implying that the shock does not systematically affect households with certain pre-existing outcomes. This supports the validity of our shift-share design as the shift appears quasi-random and thus unlikely to capture differential pre-existing trends among male-dominated households (Borusyak et al., 2022a, 2025).

3.3 Two-Stage Least Squares

Equipped with our shift-share instrument, z_{idt}, and having shown evidence supporting 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, as before, $s_{dt} := \sum_{d' \in \Theta / d} \frac{inc_{d't} \times pop_{d'}}{\tau_{dd'}}$ is 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 the 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. X'_{dt} is a vector of covariates including drought intensity, temperature, and rainfall. α_i and γ_t are household and year fixed effects, respectively. To the extent that z_{idt} is plausibly exogenous (Section 3.2.3) 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 outcomes are technology adoption, farm size, labor-days, and crop profits. 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 increasing production, in line with the idea of labor-capital substitution. $\beta_1 < 0$ implies that households downsize farming operations.

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 bilateral nature of migration decisions (Borusyak et al., 2022b) whereby migrants respond to *differences* between origin and potential destination incomes.

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

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (σ)	-1.076***	-0.733***	-1.237***	-1.902***	-1.378***
	(0.237)	(0.220)	(0.264)	(0.331)	(0.254)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes
Outcome SD	1.017	1.028	3.621	211.329	21059.452
Explanatory SD	0.524	0.521	0.533	0.518	0.513
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25928	24970	29346	20910	25854
F-Statistic	55.2	55.6	46.8	52.6	60.0

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Outcomes and explanatory variables are in standard deviations. Male Migrants 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. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Column 3 is farm size in acres. Labor (column 4) is total person-days of labor (family + hired workers). Profits (column 5) is crop income net of expenses. All specifications control for drought, temperature, and rainfall. Standard errors clustered by district.

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 5.2 extends the analysis to enable market adjustments.

4.1 Main Estimates

Table A5 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 has 1.76 male working-age residents. Therefore, a 1σ increase in destination incomes pulls 21% (=0.369/1.76) 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 compare coefficient magnitudes across outcomes. The coefficient of interest is negative and statistically significant for technology, farm size, farm labor, and crop profits, suggesting that labor loss prompts agricultural decline within the household, not modernization.[10] The point estimate in column 1 implies that a 1σ reduction in farm labor causes households to reduce agricultural expenses by 1.08σ. When decomposing the index by individual technologies (Table A8, 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 less investment in tubewells, water pumps, bull carts, and threshers (Table A8, columns 6-10).

Columns 3-5 of Table 3 explore additional response margins. Column 3 implies that labor loss causes a reduction in cultivated area by 1.24σ, in line with the idea that declining labor and technology reduce the marginal productivity of land, which in turn prompts farm contraction (Appendix C). As we show in Section 4.2, the large coefficient is driven by large farms downsizing, whereas small and medium farms remain of similar size. Column 4 shows that total farm labor days, including both family and hired labor, declines significantly in response to emigration. This finding is consistent with the presence of imperfect rural labor markets: households do not fully compensate for the loss of family labor by hiring additional workers. A 1σ increase in migration leads to a 1.9σ reduction in total farm labor, corresponding to roughly 400 person-days.

Lower investment in capital, land, and labor inputs in response to labor migration leads to lower profits (column 5), likely because of reduced output. The point estimate implies that a 1σ increase in male migration leads to a 1.38σ decline in agricultural profits over the past year, corresponding to about Rs. 29,000.

Finally, since labor (column 4) declines more than capital (columns 1 and 2), agriculture may become more capital-intensive. We test this using the capital-labor ratio (total agricultural expenses divided by total labor days) as an outcome. The results in Appendix Table A7 are consistent with capital intensification, although estimates are not statistically significant in the IV specification. We do, however, find evidence of capital uptake at the district-level (Section 6.1) when general equilibrium forces are at play.

[10]Corresponding OLS estimates are in Table A6. Estimates are positive, small, and relatively noisy. This is likely because migration is endogenous due to two counteracting forces: on one hand, richer households can send more migrants and also have access to better technology, leading to a positive correlation between migration and agricultural development. On the other hand, urban opportunities draw workers out of agriculture, leading to agricultural contraction (Table A5, Table A9). These push and pull factors draw the coefficient in opposing directions, leading to attenuated OLS estimates. Our IV estimates overcome this issue by isolating the pull stream of migration that is plausibly orthogonal to push factors at the origin.

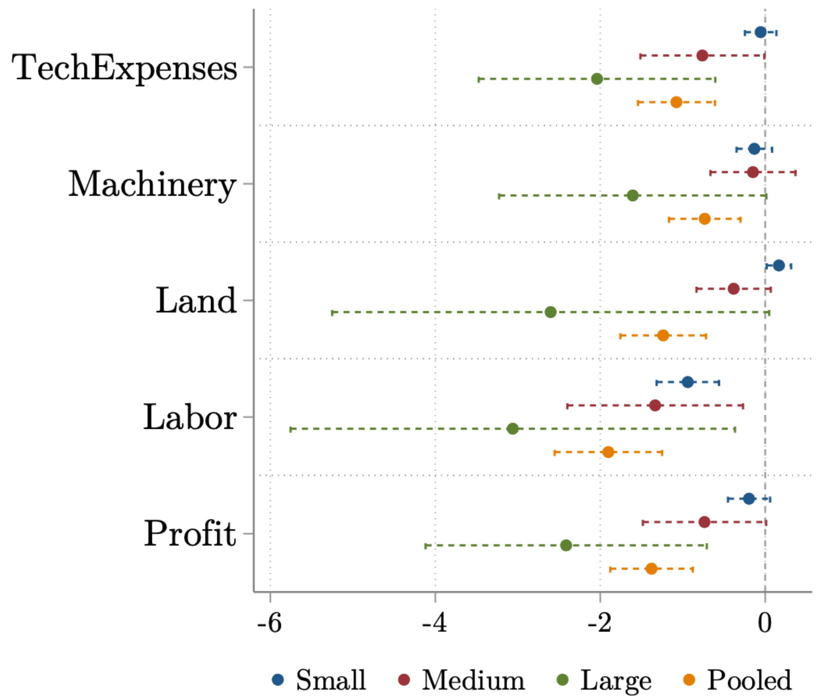


Figure 3: IV Estimates—Direct Effects of Migration on Agriculture by Farm Size

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. Outcomes on the y-axis are standardized. Blue, red, green, and orange coefficients are for small (0-2 ac.), medium (2-4 ac.), large (4+ ac.),and pooled farm sizes, respectively. Tech Expenses and Machinery are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Land is farm size in acres. Labor is person-days of labor (family + hired). Output is crop profits net of expenses. All specifications control for drought, temperature, and rainfall. Standard errors clustered by district.

4.2 Estimates by Farm Size

To test whether our results in Table 3 are a direct response to labor loss, we compare 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 impact. Our setting is well suited for this test since, as documented in Fact 3 (Section 2.4), labor constraints increase in farm size. We therefore expect farmer responses to emigration to increase in farm size as labor constraints become tighter. Another reason that larger farms likely drive our estimates is because these farms are more mechanized at baseline, enabling larger margins of adjustment compared to smaller farms.

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[11]. Although 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 and often statistically insignificant. 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-constrained farmers directly responding to labor loss, as opposed to another channel.

4.3 Additional Estimates

Our IV strategy leverages urban income shocks, weighted by rural households' ability to exploit those shocks, to study how male labor loss shapes agricultural development. This approach implicitly takes the view that pulling male labor off farms is the primary mechanism through which urban opportunities affect rural agriculture.

To justify this view, Table A9 presents reduced form estimates of Equation 4, which are agnostic about mechanisms. Any factor correlated with urban income shocks that also affects agricultural outcomes—whether migration or otherwise—will be reflected in these estimates. Yet despite capturing all channels, the reduced form continues to show that urban income shocks trigger agricultural disinvestment, with coefficients closely matching our second stage estimates that isolate only the migration mechanism. The similarity between the two supports our view that male labor migration is a key channel of interest.

Remittances provide a useful illustration of the above logic. Urban growth raises remittances, particularly for households with more males, which may be re-invested back into the family farm. If remittances were the dominant channel linking urban growth to rural agricultural outcomes, then this would be reflected by *positive* reduced form estimates. The fact that the reduced form is negative and similar to the second stage thus implies that our main estimates hold *in spite of* remittances. Table A10 confirms this directly by controlling for remittances in the second stage, where the estimates remain negative and slightly increase in magnitude. This is because once the upward pressure from remittances is removed, the negative effect of labor loss on agricultural development becomes more apparent.

4.4 Robustness Checks

We now probe the sensitivity of our estimates of direct effects. Panel A of Table A11 tests robustness to using state-by-year instead of year fixed effects. State-year fixed effects

[11] Figure B2 shows the distribution of household farm size. The vast majority of farms are < 5 acres.

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 tests robustness to including the extensive margin in the sample—households that left or joined agriculture between the two surveys. 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.

Panels C-E show that our estimates are robust to alternative shift-share instruments. First, we explore alternative shifts. Panel C redefines z_{idt} to isolate rural-urban migration by restricting the destination choice set (Θ in Equation 1) to urban destinations only, and restricting the sample of out-migrants to rural households only. 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 D, we use inverse-distance weighted nightlights as the shift[12]. 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 E 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 1).

Lastly, Table A12 reports estimates from alternative methods of statistical inference. Our baseline estimates report standard errors clustered by district, since households in the same district are exposed to the same shock. We also test robustness to clustering at the Population Sampling Unit level, a fundamental geographic unit defined by IHDS for the initial stage of sampling, and within which households may share unobserved characteristics. 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[13]. Overall, the statistical significance of our estimates of the direct effect of migration on agricultural outcomes is robust to these alternative methods of accounting for spatial and serial correlation.

[12] 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.

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

4.5 Summary of Direct Effects

We have shown that rural agricultural households in India do not substitute labor with capital when facing labor loss. Instead, they reduce agricultural technology adoption, downsize their farms, and grow less food. India's fast urbanization (Figure 1) thus raises a concern about domestic food supply. But the net effect on food supply also depends on general equilibrium adjustments through land and food markets: how do other agricultural households who reside in the same village, or other households in the state who grow the same crops, react to migrant households' divestment from agriculture? Our rich household panel data allows us to decompose these partial and general equilibrium responses and better reconcile our findings with existing work (e.g. Hornbeck and Naidu (2014)), which has only managed to conduct such analysis at aggregated levels. We turn to this decomposition next and, in doing so, extend the literature and deepen our understanding about the spatial reorganization of agriculture in response to structural change.

5 Equilibrium Crop and Land Market Adjustments

To build intuition about the nature of indirect effects, first consider crop markets. If declining food supply through the direct channel (Table 3) raises crop prices, then other households may exploit this opportunity by scaling up farming. This indirect crop market channel may be especially pronounced among non-migrant households, who avoid negative direct effects, but still benefit from higher crop prices. Since non-migrant households are more remote (Fact 2, Section 2.4), we expect production booms in remote areas. Similarly, if declines in migrant households' farm size depress land prices, we may also see farm expansion among other village residents through a land market channel.

We formalize these spatial patterns into a spatial equilibrium 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 channels. We use this feature to extend our 2SLS setup to incorporate both of these channels.

5.1 Measurement

5.1.1 The Land Market Channel

To measure land market adjustments, we 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 5.5, we show that their response to M_{ijdt}^{land} materializes through changes in village land prices (Table 5).

5.1.2 The Crop Market Channel

To measure crop market adjustments, we describe crop markets as the combination of all households in the state producing similar crops. One concern is that crop choices are endogenous because they depend on crop prices, labor supply, and land supply, which all respond to migration and urban productivity shocks. We, therefore, define two households as part of the same crop market if: (i) they live in areas with similar crop-specific crop suitability (e.g., both located in rice-growing areas), and (ii) in a robustness test, if they produce similar crop portfolios at baseline based on IHDS data. Both measures are exogenous to urban productivity shocks during the study period.

To quantify crop similarity, consider two vectors $\mathbf{x}_i = (x_{i1}, ..., x_{iK})$ and $\mathbf{x}_{i'} = (x_{i'1}, ..., x_{i'K})$ that list the K crops available to households i and i' based on eco-climatic suitability. 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 suitability vectors[14]:

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

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

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

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

We assume state-level crop markets[15] since cross-state trade restrictions in India preventing crop prices from equilibrating across states (Chatterjee, 2023). N_s/i thus denotes all households in state s excluding household i. Since emigration from other households growing equally suitable crops to i receives more weight in the aggregation, M_{ijdt}^{crop} effectively measures aggregate emigration from the same crop market as household i. Thus, household i's response to changes in M_{ijdt}^{crop} reflects indirect effects through crop markets. In Section 5.5, we show that household responses to crop market adjustments captures an underlying relationship between M_{ijdt}^{crop} and crop prices (Table 5). Intuitively, if urban shocks draw agricultural labor away from a rice-growing area, thereby affecting aggregate rice supply and rice prices, then households elsewhere will respond more to the price change if they also live in a rice-suitable part of the state.

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[16]. District suitability for each crop is computed by extracting means over all cells within a district.

5.2 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:

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

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

where M_{ijdt}^{land} is the land market adjustment and M_{ijdt}^{crop} is the crop market adjustment. Other terms are the same as Equation 4. The total response of farmers to emigration

[15]We ignore crop trade cost for simplicity. A possible extension could weight emigration of other households by their inverse distance to the focal household. However, since the crop market is heavily regulated within states, we assume the government absorbs inter-state trade costs.

[16]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.

comprises the direct effect through household labor loss, β_1, as well as the land and crop markets adjustments, captured by β_2 and β_3, respectively. Only the direct labor channel, M_{ijdt}^{labor}, is instrumented with z_{ijdt}, whereas the market responses 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 is that households i and i' may respond to village-level shocks (e.g., a local drought or agricultural programme) that co-determine village emigration and household agriculture. We therefore control for droughts, temperature, and rainfall in all specifications, as well as for village agricultural extension programs in a robustness check. To the extent that these shocks manifest as pull factors, these are captured by distance-weighted destination income shocks, s_{dt}. Another threat is that emigration from the land market may affect non-land factor markets. We rule this out by controlling directly for non-land factor market changes in a placebo test (Section 5.4.2).

For crop markets, the concern is that crop-specific shocks, such as crop disease or trade disruption, can simultaneously affect emigration from the crop market and household agriculture via prices. We therefore control for crop prices in a robustness check such that identification of β_3 is off of deviations from general crop price trends driven by aggregate supply contractions induced by aggregate emigration from the crop region. We validate this measure of the crop market channel, M_{ijdt}^{crop}, by showing that households do not respond to aggregate emigration from outside the crop market.

5.3 Results: Direct and Indirect Effects

Table 4 presents estimates of β_1, β_2, and β_3 from Equation 8. Estimates 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. However, column 1 shows that migration-induced declines in agricultural technology due to labor loss are partially offset by household responses to land (row 2) and crop (row 3) market adjustments.

When decomposing by technology type, Table A13 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 5.5 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 par-

Table 4: Direct and Indirect Effects of Internal Migration

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants	-1.073***	-0.741***	-1.245***	-1.952***	-1.359***
(direct labor channel)	(0.228)	(0.218)	(0.257)	(0.316)	(0.245)
Village emigration	0.235***	0.188***	0.260***	0.405***	0.245***
(indirect land channel)	(0.052)	(0.048)	(0.055)	(0.077)	(0.057)
Crop region emigration	0.206***	0.031	0.263***	0.450***	0.180***
(indirect crop channel)	(0.057)	(0.057)	(0.058)	(0.084)	(0.050)
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25924	24966	29342	20906	25852
F-Stat on Direct Effect	64.5	59.7	52.7	56.5	67.4

Note: * $p < .1$, ** $p < .05$, *** $p <.01$. All variables are standardized. *Male Migrants* is 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 5.1). Columns 1 and 2 are indices of agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired). Profit (column 5) is crop revenue net of expenses. All columns control for origin income, the uninteracted shift, drought, temperature, and rain. Standard errors clustered by district.

tially 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 5.5. Column 4 reports labor market responses and indicates that total labor days declines directly in response to emigration. The indirect effects, however, operate in the opposite direction, which is counterintuitive since wage rates possibly increase as labor supply declines. We interpret this pattern as reflecting households' increased own-family labor supply in response to higher input use and rising crop prices.

Lastly, the estimates for crop profits in column 5 show the same directionality through each channel. Point estimates are precise and economically significant. Whereas households experience lower crop profits by 1.36σ in response to their own labor loss, agricultural expansion through crop market and land market adjustments offset these effects. In Section 6.2, we conduct an accounting exercise that uses these coefficients to quantify the extent to which aggregate agricultural losses are compensated by market spillovers.

Note that negative direct effects primarily impact households sending many migrants, whereas households sending zero migrants indirectly benefit from agricultural expansion. Since remote households send fewer migrants (Fact 2, Section 2.4), we expect a spatial reorganization of agriculture away from peri-urban areas with high emigration toward remote areas with low migration. In Section 5.6, we show that this spatial pattern is indeed borne out in aggregate and high-resolution data, supporting our empirical findings and corroborating our story about the spatial reorganization of agriculture.

5.4 Robustness and Placebo Tests

We test the sensitivity of our estimates of crop and land market adjustments with a variety of robustness and placebo tests. Our estimates are remarkably stable, and our market measures appear robust and well-defined.

5.4.1 Robustness Tests

Table A14 shows that estimates of land market adjustments are robust to controlling for village level environmental shocks (drought, flood, cyclone, or earthquake) as well as the presence of agricultural extension programs. This specification is important since M_{ijdt}^{land} is uninstrumented in Equation 8 and village level shocks can influence both village-level emigration as well as household agricultural decisions. Yet the coefficient on M_{ijdt}^{land} remains highly stable, suggesting that our main estimate of land market adjustments is unbiased by correlated village-level shocks.

Table A15 shows that estimates of crop market adjustments also remain stable when controlling for annual prices of households' highest-value crop[17]. This absorbs crop-level shocks, such as crop disease or trade disruptions, that affect both crop-region emigration and household agriculture via crop prices. While prices are the proposed mechanism in our model, making this control seem redundant, it in fact only captures general crop price trends. β_3 is thus identified off of deviations from general crop price trends driven specifically by supply contractions induced by aggregate emigration from the crop region. We provide empirical support for this mechanism in Section 5.5.

Table A16 tests robustness to an alternative definition of the crop market. For this test, crop similarity weights are defined in crop production space instead of 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, estimates of β_3 are virtually identical to Table 4 when we use this alternative measure of M_{ijdt}^{crop}.

[17] Estimates are also robust to controlling for linear crop-time trends.

Table A17 drops inverse distance-weighted income as a covariate. Recall that this was partialled out in the main shift-share specification to improve identification of direct effects. Yet, in doing so, variation in migration from pull factors becomes underexploited for identifying indirect effects. We find remarkably stable estimates when adding back this pull factor variation by not controlling for destination incomes.

Table A18 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.

Table A19 test robustness to including the extensive margin in the sample. Recall that Table 4 describes the intensive margin since the sample comprises households that owned land in both periods. Households who left or entered agriculture between surveys have missing outcomes when they are landless. We add them to the sample by zero-filling their agricultural technology expenses and machinery. 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.

5.4.2 Placebo Tests

In addition to the robustness checks, we design placebo tests to validate our measures of crop and land market adjustments and provide additional support for causal interpretation of our estimates of indirect effects. 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 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 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 A20, we control for the placebo in Equation 8, and its coefficient turns negative and converges to zero. The robust positive coefficient on M_{ijdt}^{crop} confirms 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)[18]. Taking the inverse implies that households growing unrelated crops receive more weight in the aggregation. Column 2 of Table A20 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 our land market measure. One concern is that emigration by other agricultural households in the village might also shift non-land factor prices, like wages or input costs, in addition to land prices. We therefore control for emigration by landless households in the village in Equation 8, which affects only non-land GE channels and therefore disentangles the effect of land and non-land GE channels on household agriculture. We are well-powered for this test since over half of the households in our sample are non-agricultural. The land market placebo does not affect household farm size, whereas our main land market coefficient remains stable. We conclude that our land market measure, M_{ijdt}^{land}, is indeed picking up farm size adjustments through land markets and not other indirect channels.

5.5 Mechanisms: Land and Crop Prices

We have argued throughout the paper 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 market responses. We provide evidence 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, with weights equal to the suitability index for each crop in the district (Section 5.1 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:

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

[18]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	(3) Land Price
Crop region emigration (indirect crop channel)	0.137*** (0.052)		
Village emigration (indirect land channel)		-0.052** (0.025)	-0.055** (0.025)
Vil. Environmental Shocks	No	No	Yes
Vil. Agricultural Extension	No	No	Yes
Aggregation	District-year	Village-year	Village-year
District FEs	✓		
Village FEs		✓	✓
Year FEs	✓	✓	✓
Observations	586	1222	1100

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Coefficients are OLS estimates. All outcomes are standardized. In columns 2 and 3, regressions are weighted by the number of households in a village 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. Errors clustered by district in column 1 and by village in columns 2-3.

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

where d, j, and t index districts, villages, and time, respectively. $CropPrice$ is the crop price index in district d and $LandPrice$ is land rental price in village j . M^{crop}_{dt} and M^{land}_{jt} are measured as before, but aggregated at the district and village level, respectively. The coefficients of interest are Φ and ψ. These are estimated with OLS since M^{labor}_{ijdt} is not part of this specification and M^{land}_{jt} and M^{crop}_{dt} are uninstrumented. 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, significant and implies that a 1σ increase in 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 direct labor channel raises crop prices, leading non-migrant farmers in the crop market to increase production.

Column 2 documents the impact of 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 cultivated land through the direct labor channel increases land availability, thereby reducing land prices and leading to farm expansion among non-migrant households in the land market (village). The land price effect remains stable when controlling for the presence of village environmental shocks and agricultural extension programs (column 3).

5.6 The Spatial Reorganization of Agriculture in Geo-coded Data

We have uncovered two distinct responses to migration: (i) negative direct effects, in the form of *agricultural contraction* among migrant-sending families, and (ii) positive indirect effects, in the form of *farm expansion* by non-migrant households in shared land and crop markets. Our empirical model suggests that these responses unfold along a spatial gradient: agricultural contraction (Table 4) materializes among peri-urban households, who face low migration costs and send many migrants, whereas agriculture expands among remote households, for whom migration costs are high and emigration is uncommon. This implies a spatial reorganization of agriculture that can be tested with geocoded data. We therefore map changes in economic growth, agricultural labor, and crop output across India using both aggregated census data as well as high-resolution satellite imagery.

Maps presented in Figure 4 validate our empirical findings and help us visualize how structural transformation reshapes the geography of agriculture. Panels A and B map changes in agricultural labor and output from 2001-2011, respectively. The ten fastest-growing cities by IHDS income—the "shift" in our shift-share design—are labeled. Two spatial patterns emerge, each consistent with our empirical findings. First, near the cluster of high-growth cities in South India, there are sharp declines in agricultural labor (Panel A, red) in surrounding districts where the migration pull-factor is strongest. These same labor-sending districts also show declining crop output (Panel B, red), echoing our estimates of direct effects (Table 3): as males migrate to cities, left-behind families scale back farming. Second, moving away from these urban clusters into remote regions of Central India, where there are no high-growth cities, we observe a compensating *expansion* of agriculture (Panel B, blue). This aligns with our finding that indirect effects are more pronounced for remote households further away from the pull of economic growth.

Table A21 systematically explores the spatial correlations evident in Panels A and B by regressing the changes in agricultural laborers per hectare and crop yields in a district between 2001 and 2011 on the district's distance to the nearest of the ten high-growth cities. Column 1 shows that agricultural labor increases with distance, although the coefficient is imprecise. Column 2 reveals a more precise association for crop output: districts 100

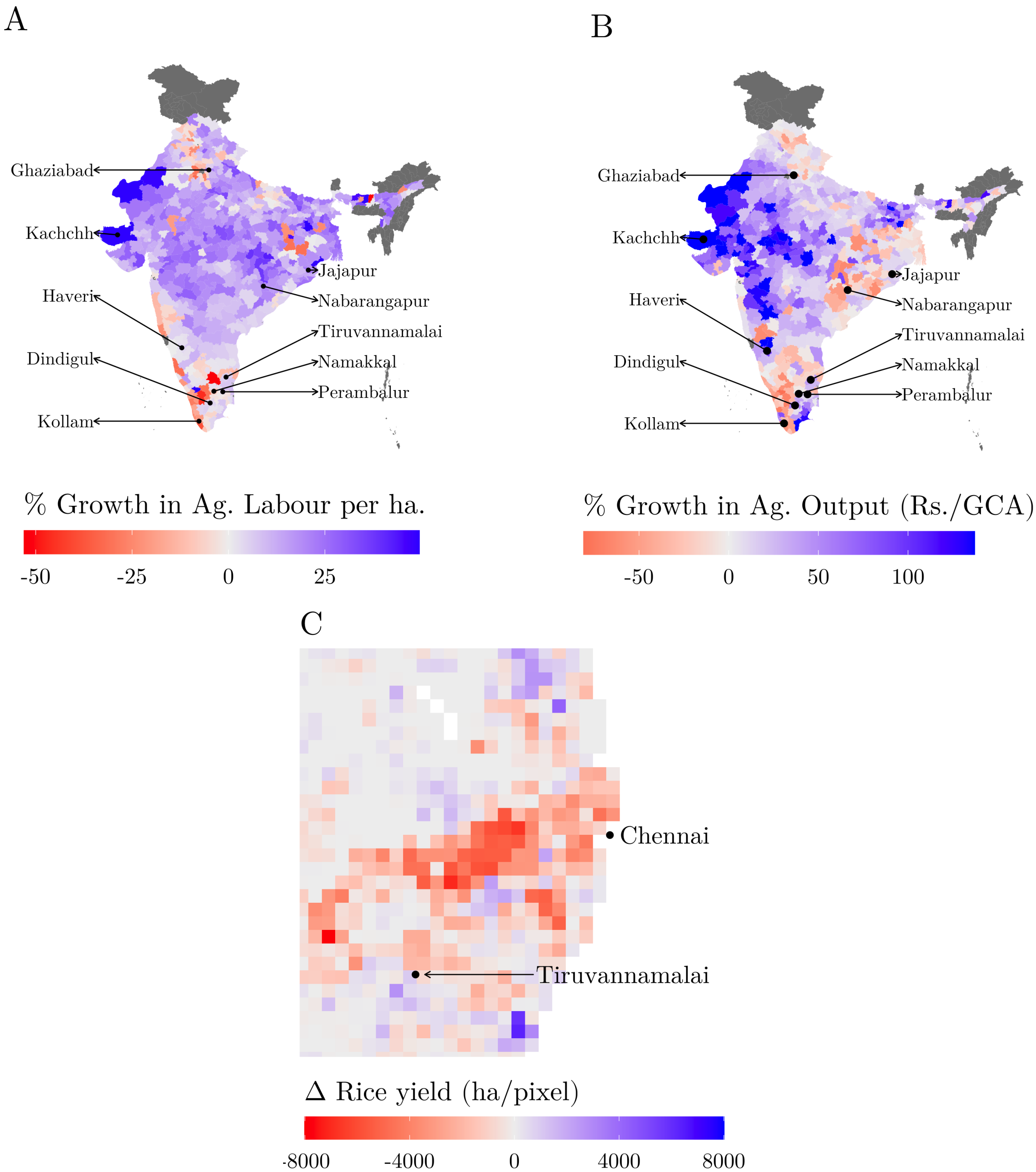


Figure 4: Economic Growth, Labor Exit, and Agricultural Growth (2001-2011)

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 the change in harvested rice area per pixel from the GGCP10 satellite product.

km farther from growth hubs exhibit roughly 3 percent higher yields.

Having established the spatial reorganization of agriculture across districts, Figure 4C zooms in on even more localized patterns using high-resolution, satellite-based measures of rice yields in the fast-growing regions of South India. Panel C centers on Tiruvannamalai, a city near Chennai and one of the highest-growth cities during this period. Yields decline to the northeast and northwest of the city—areas with significant labor outflows (Panel A). Yet, yields begin to rise beyond these zones as we move into more suburban areas. This pattern mirrors land market spillovers: as migrant-sending households scale back farming, non-migrant neighbors in the land market (village) expand cultivation. In some cases, red and blue pixels even appear side-by-side, vividly capturing this hyper-localized spatial reallocation of agriculture. These visual patterns help rationalize our empirical findings, consistent with the claim that a geographic redistribution of agriculture accompanies structural transformation.

6 Discussion

This section contextualizes our findings, first, by connecting to the existing literature and, second, by conducting an accounting exercise to quantify the extent to which aggregate agricultural losses are compensated by crop and land market spillovers.

6.1 District-Level Analysis

A key contribution of our paper is the use of household data to disentangle direct and indirect effects of emigration on agriculture. This decomposition is only possible with household data: direct effects unfold within households, whereas indirect effects unfold spatially across households linked through shared markets. Given that the literature studies structural transformation and agriculture at higher geographic scales (see Literature Review), we can aggregate our household data and estimates up to coarser scales to reconcile our findings with existing work.

Related studies find that labor loss leads to agricultural mechanization using county-level (Hornbeck and Naidu, 2014), district-level (Aggarwal, 2018), or even state-level (Clemens et al., 2018) data. On these broader geographic scales, the effect of labor loss on agriculture reflects both labor-capital substitution *and* broader market adjustments to labor loss. To bridge our findings with this literature, Table 6 presents district-level estimates of Equation 4, where all variables are aggregated to district means. Market responses are not separately modeled in this equation, so aggregation to the district level

Table 6: IV Estimates—District Level Aggregate Impacts of Migration on Agriculture

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (σ)	0.970**	1.357**	-0.807*	-1.638	-1.253**
	(0.459)	(0.675)	(0.486)	(1.126)	(0.553)
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes
Outcome SD	0.623	0.666	1.941	124.462	11914.408
Explanatory SD	0.160	0.160	0.158	0.161	0.159
District FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	614	618	656	608	640
F-Statistic	5.3	5.7	4.7	5.2	6.1

Note: * $p < .1$, ** $p < .05$, *** $p <.01$. Data are aggregated to district means and converted to standard deviations. Regressions are weighted by number of households over which the mean is taken. Male Migrants is the number of working age male migrants sent from the household, instrumented with inverse-distance population weighted income. All columns control for drought, temperature, and rainfall. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired). Profits (column 5) is crop revenue net of expenses. Standard errors clustered by district.

implies that the coefficient entangles both the response of households to emigration of their own household members as well as responses to broader market adjustments within districts from aggregate emigration. District emigration is instrumented with inverse distance-weighted income shocks (s_{dt}) while also controlling for district income (inc_{dt}), in line with Borusyak et al. (2022b). Note that our instrument is not as strong at the district level, but these estimates are intended only to contextualize our findings.

Our aggregate estimates are consistent with the main patterns in the literature. Labor migration leads to increased *technology adoption* at the district level (Table 6, columns 1 and 2), implying that positive indirect effects outweigh negative direct effects in aggregate. This is not true for farm size and profits (columns 3 and 5), which decline in response to aggregate emigration. The negative effect on labor days is statistically insignificant. Overall, these findings suggest that technology adoption by non-migrant households does not fully compensate for direct labor losses in terms of production impacts.

These aggregate estimates in Table 6, and those in the literature, capture only the subset of broader market responses that manifest *within* districts. Yet, recall that crop markets can stretch beyond districts (Section 5.1). Crop market spillovers are therefore insufficiently unaccounted for in these district-level estimates. This may explain why district crop profits decline despite greater technology adoption: agricultural expansion unfolds

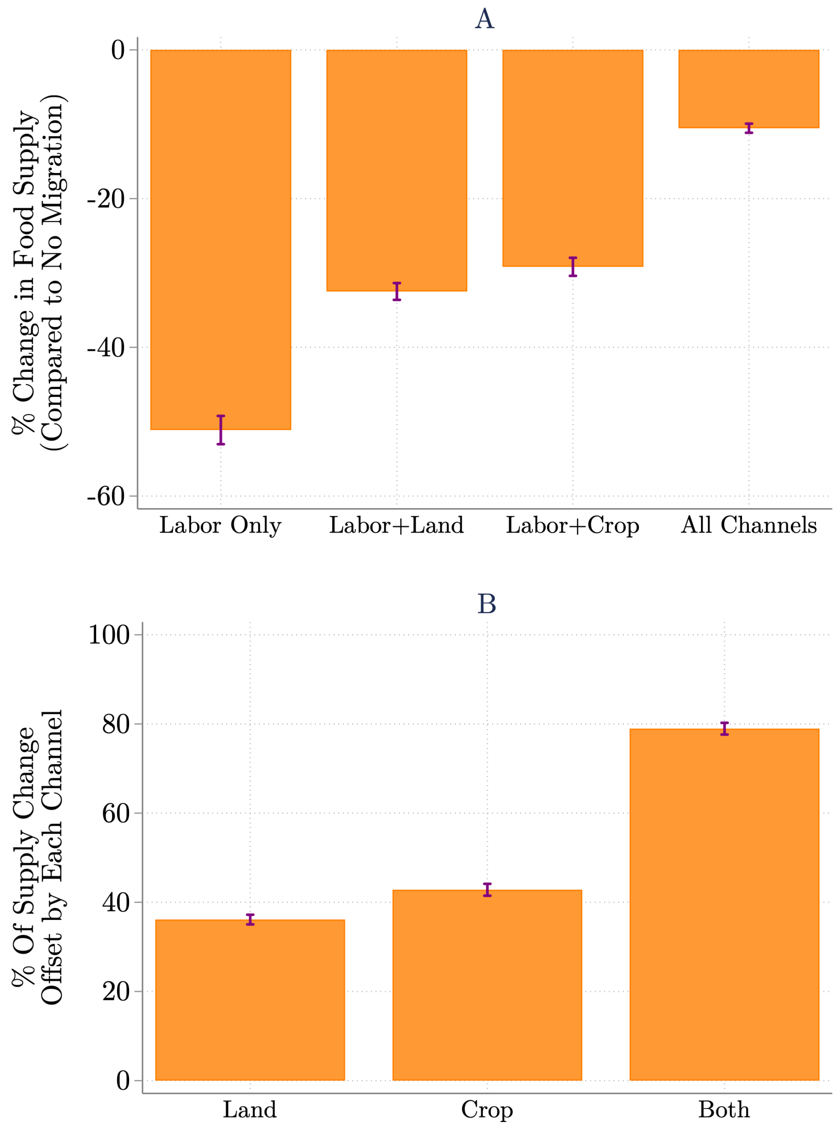


Figure 5: Aggregate Extent of Indirect Effects

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.

through spatial market spillovers in ways that Equation 4 overlooks when aggregated at higher levels (across states, or across countries, through international trade).

6.2 How much agricultural loss does land and crop markets offset?

In aggregate, how much of the crop losses from agricultural divestment among migrant families are compensated by market adjustments? We use our household-level econometric estimates from Equation 8 to calculate below that 80% of the migration-induced decline in food supply is compensated by domestic market adjustments. These indirect effects materialize in remote areas, which explains the spatial reorganization of agriculture away from peri-urban areas and toward remote regions evident in Figure 4.

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 crop market adjustment is fixed at t_1, as well as a *Labor and Crop* (LC) scenario where the 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.

Figure 5A shows estimates of the change in aggregate crop output with and without 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 both market channels shut off, aggregate migration would have caused a 51% 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 five times smaller.

Panel B shows how much of the migration-induced output decline is mitigated by market adjustments (Equation 9). 36% of agricultural losses through the labor channel are mitigated through land markets and 43% through crop markets. Both indirect channels together mitigate 79% 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.

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 in rural areas left behind 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.

When analyzing structural transformation at an aggregated district level in India, we find evidence consistent with historical US literature that farms mechanize in response to emigration. But our analysis using detailed farm-level data indicate that Indian farmers do not replace labor with capital when facing agricultural labor loss. Instead, they downsize farms and reduce crop output. 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. Aggregate effects in the literature thus represent a combination of these direct and indirect general equilibrium effects.

Our analysis uncovers an interesting spatial pattern in the reorganization of agriculture in response to migration. 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 remote households increase production, expand farmland, and adopt technology. In aggregate, our estimates suggest that market spillovers mitigate 80% of the migration-induced food shortage between 2005-2012. The spatial redistribution of agriculture through markets is, therefore, economically significant.

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. How these shifts have played out in China and other developing countries that have experienced massive urbanization and structural transformation in recent decades is an important under-explored area that should be the subject of 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.

Agarwal, Bina and Malvika Mahesh, "Does the landowner's gender affect self-cultivation and farm productivity? An analysis for India," *The Journal of Development Studies*, 2023, *59* (5), 758–777.

Aggarwal, Shilpa, "Do rural roads create pathways out of poverty? Evidence from India," *Journal of Development Economics*, 2018, *133*, 375–395.

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.

__ , __ , and __ , "A practical guide to shift-share instruments," *Journal of Economic Perspectives*, 2025, *39* (1), 181–204.

__ , 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.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "Labor rationing," *American Economic Review*, 2021, *111* (10), 3184–3224.

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.

Dinkelman, Taryn and Martine Mariotti, "The long-run effects of labor migration on human capital formation in communities of origin," *American Economic Journal: Applied Economics*, 2016, *8* (4), 1–35.

_ , Grace Kumchulesi, and Martine Mariotti, "Labor migration, capital accumulation, and the structure of rural labor markets," *Review of Economics and Statistics*, 2024, pp. 1–46.

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.

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, "Economic development and the decline of agricultural employment," *Handbook of development economics*, 2007, *4*, 3051–3083.

_ and _ , "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.

Garg, Teevrat, Maulik Jagnani, and Hemant K Pullabhotla, "Rural roads, farm labor exits, and crop fires," *American Economic Journal: Economic Policy*, 2024, *16* (3), 420–450.

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.

Hoffmann, Lars, Gebhard Günther, Dan Li, Olaf Stein, Xue Wu, Sabine Griessbach, Yi Heng, Paul Konopka, Rolf Müller, Bärbel Vogel et al., "From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations," *Atmospheric Chemistry and Physics*, 2019, *19* (5), 3097–3124.

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.

Janvry, Alain De, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet, "Delinking land rights from land use: Certification and migration in Mexico," *American Economic Review*, 2015, *105* (10), 3125–3149.

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.

Khanna, Gaurav, Emir Murathanoglu, Caroline B Theoharides, and Dean Yang, "Abundance from abroad: Migrant income and long-run economic development," Technical Report, National Bureau of Economic Research 2022.

Kinnan, Cynthia, Shing-Yi Wang, and Yongxiang Wang, "Access to migration for rural households," *American Economic Journal: Applied Economics*, 2018, *10* (4), 79–119.

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.

Tian, Yuan, Junjie Xia, and Rudai Yang, "Trade-induced urbanization and the making of modern agriculture," *Working Paper*, 2023.

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

	Male Migrated for School	Male Migrated for Work
0-4	0.29	0.20
5-9	0.55	0.03
10-14	0.55	0.06
15-19	0.41	0.31
20-24	0.24	0.55
25-29	0.06	0.75
30-34	0.01	0.81
35-39	0.01	0.84
40-44	0.00	0.83
45-49	0.00	0.83
50-54	0.00	0.82
55-59	0.00	0.79
60+	0.00	0.57
Total	0.18	0.58

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: 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 A3: 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 A4: Summary Statistics of Land, Labor, and Capital

	Obs.	Mean	SD
A: Land			
Area cultivated (ac.)	30425	3.33	3.61
Area rented in (ac.)	27024	0.28	0.79
Rental price (Rs./ac./yr)	2980	3346.86	13541.57
Yield (Rs./ac.)	24730	7738.27	25546.84
B: Expenses/Acre			
Seeds	27450	668.57	2355.20
Fertilizer	27238	930.41	3449.98
Pesticides	26497	255.38	1524.74
C: Machinery (Num./Acre)			
Pumps	26002	0.14	0.93
Tractors	26005	0.02	0.13
Bullock Carts	26091	0.05	0.24
D: Labour (past yr.)			
Family person-days/acre	27674	223.20	614.61
Hired person-days/acre	24721	12.96	62.57
Wages Paid/Acre	26214	770.60	4526.35

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.

Table A5: First Stage Results

	(1)	(2)	(3)
z_{idt} : Wt. Income $\times$ Working Age Males	0.321***	0.367***	0.369***
	(0.051)	(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	39.59	58.94	59.96
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, temperature, and rain. Standard errors clustered by district.

Table A6: OLS Estimates—Direct Effect of Migration on Agricultural Development

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (σ)	0.022***	0.002	0.018***	0.003	0.005
	(0.007)	(0.007)	(0.006)	(0.011)	(0.008)
Weather Controls	Yes	Yes	Yes	Yes	Yes
Outcome SD	1.017	1.028	3.621	211.329	21059.452
Explanatory SD	0.524	0.521	0.533	0.518	0.513
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25928	24970	29346	20910	25854
R^2	0.742	0.739	0.771	0.645	0.741

Note: * p < .1, ** p < .05, *** p <.01. Outcomes and explanatory variables are in standard deviations. Male Migrants is the number of working age male migrants in the household. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profits (column 5) is crop income net of expenses. All specifications control for drought, temperature, and rainfall. Standard errors clustered by district.

Table A7: Direct Effect of Migration of Labor-Capital Ratio

	OLS	IV
	(1)	(2)
	Log(K/L)	Log(K/L)
Male Migrants (σ)	0.030**	0.067
	(0.013)	(0.177)
Wt. Income (s_{dt})	No	Yes
Origin Income (inc_{dt})	No	Yes
HH FEs	✓	✓
Year FEs	✓	✓
Observations	19092	19092
F-Statistic		47.6

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The outcome is total agricultural expenses (K) divided by person-days of labor (L). We transform this by log(x+0.01) to avoid divide-by-zero errors. Male Migrants is the number of working age male migrants in the household. Column 1 is the OLS estimate. In column 2, migration is instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. All specifications control for drought, temperature, and rainfall. Standard errors clustered by district.

Table A8: IV Estimates—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)	-1.039***	-1.452***	-0.760***	-0.580**	-0.798***	-0.413***	-0.465*	-0.973***	-0.357	-0.600***
	(0.227)	(0.275)	(0.189)	(0.252)	(0.188)	(0.152)	(0.246)	(0.217)	(0.244)	(0.208)
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-Statistic	55.1	51.4	47.8	46.5	53.4	54.1	56.8	53.6	53.3	54.1

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Outcomes and explanatory 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 the number of baseline working age males in the household. 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, temperature, and rainfall. Standard errors clustered by district.

Table A9: Reduced Form Estimates of Direct Effects

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
z_{idt}: Wt. Income × Working Age Males	-0.819***	-0.553***	-0.886***	-1.568***	-1.134***
	(0.127)	(0.144)	(0.117)	(0.146)	(0.123)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes
Outcome SD	1.017	1.028	3.621	211.329	21059.452
Explanatory SD	0.524	0.521	0.533	0.518	0.513
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25928	24970	29346	20910	25854
R^2	0.020	0.021	0.024	0.050	0.046

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Outcomes and explanatory variables are in standard deviations. "Wt. Income" is inverse-distance, population-weighted income (s_{dt}). "Working Age Males" is number of working age male household residents (φ_i). "Origin Income", inc_{dt}, is mean per capita district income. z_{idt} is in standard deviations. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profits (column 5) is crop income net of expenses. All specifications control for drought, temperature, and rain. Standard errors clustered by district.

Table A10: IV Estimates—Direct Effects of Migration on Agricultural Development Controlling for Remittances

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (σ)	-1.659***	-1.142***	-2.011***	-2.842***	-2.577***
	(0.401)	(0.354)	(0.486)	(0.534)	(0.578)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes
Remittances	Yes	Yes	Yes	Yes	Yes
Outcome SD	1.017	1.028	3.621	211.329	21059.452
Explanatory SD	0.524	0.521	0.533	0.518	0.513
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25928	24970	29344	20910	25854
F-Statistic	36.4	34.6	28.5	36.8	29.6

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Outcomes and explanatory variables are in standard deviations. Male Migrants 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. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 5.1). Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profits (column 5) is crop income net of expenses. All specifications control for remittances, drought, temperature, and rainfall. Standard errors clustered by district.

Table A11: Robustness Checks: Direct Effects (IV Estimates)

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Panel A: State-Year FEs					
Male Migrants (σ)	-0.857***	-0.582***	-1.245***	-1.805***	-1.480***
	(0.199)	(0.185)	(0.247)	(0.300)	(0.270)
Observations	25928	24970	29346	20910	25854
Panel B: Extensive Margin					
Male Migrants (σ)	-1.332***	-0.965***	-1.208***	-1.976***	-1.292***
	(0.265)	(0.235)	(0.219)	(0.300)	(0.233)
Observations	69756	68698	75334	63970	66040
Panel C: Rural-Urban Migration					
Male Migrants (σ)	-1.117***	-0.752***	-1.289***	-2.055***	-1.425***
	(0.254)	(0.247)	(0.278)	(0.354)	(0.271)
Observations	25148	24288	28176	20256	24836
Panel D: Nightlights Shift-Share					
Male Migrants (σ)	-1.162***	-0.694***	-1.396***	-2.204***	-1.598***
	(0.276)	(0.246)	(0.308)	(0.407)	(0.297)
Observations	25928	24970	29346	20910	25854
Panel E: All Gender Migration					
Migrants (σ)	-2.083***	-1.532***	-2.557***	-3.431***	-2.449***
	(0.589)	(0.503)	(0.696)	(0.843)	(0.551)
Observations	25928	24970	29346	20910	25854

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. Outcomes and explanatory variables are standardized. Male Migrants is the number of working age male migrants in the household. See Table 3 footnote for details on the migration IV). Panel A includes state-year FEs and remaining panels only include year FEs. Panel B adds households who left and joined agriculture between survey waves to the estimation sample. Panel C restricts to rural-urban migration only. Panel D uses distance-weighted nightlights as the shift. In Panel E, the explanatory variable and the IV "share" includes both males and females. All specifications control for drought, temperature, and rain. Standard errors clustered by district.

Table A12: Robustness: Alternative Standard Error Clustering

	Tech. Exp.	Machinery	Land	Labor	Profit
Coefficient on Migration	-1.076208	-.7329422	-1.236715	-1.902099	-1.377935
SE: District	.2370647	.2203796	.2641385	.3312023	.2542923
SE: PSU	.215519	.1894054	.2386777	.3246026	.2353193
SE: State	.311985	.2482762	.3394252	.4121896	.328847
SE: Conley (200km radius)	.2408701	.1659192	.2011565	.3118904	.2327848
SE: Conley (500km radius)	.1806315	.1066075	.2116086	.2888249	.2637012

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 A13: 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	-1.033***	-1.445***	-0.761***	-0.585**	-0.799***	-0.417***	-0.476*	-0.987***	-0.369	-0.611***
(direct labour channel)	(0.215)	(0.258)	(0.183)	(0.253)	(0.185)	(0.151)	(0.251)	(0.219)	(0.248)	(0.209)
Village emigration	0.222***	0.309***	0.146***	0.139***	0.195***	0.131***	0.112**	0.237***	0.084	0.122**
(indirect land channel)	(0.048)	(0.059)	(0.044)	(0.054)	(0.045)	(0.040)	(0.053)	(0.052)	(0.056)	(0.053)
Crop region	0.205***	0.268***	0.208***	0.118**	0.124**	-0.076	0.046	0.047	0.084*	0.024
emigration (indirect crop channel)	(0.052)	(0.063)	(0.051)	(0.057)	(0.052)	(0.055)	(0.062)	(0.054)	(0.043)	(0.044)
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	64.3	59.3	54.0	51.7	60.6	58.1	58.7	56.3	55.1	56.1

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 in the state weighted by crop suitability (see section 5.1 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, drought, temperature, and rain. Standard errors clustered by district.

Table A14: Robustness: Village Level Shocks

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (direct labour channel)	-1.115*** (0.238)	-0.719*** (0.230)	-1.295*** (0.271)	-2.062*** (0.341)	-1.418*** (0.256)
Village emigration (indirect land channel)	0.251*** (0.057)	0.187*** (0.051)	0.270*** (0.059)	0.449*** (0.083)	0.260*** (0.060)
Crop region emigration (indirect crop channel)	0.211*** (0.062)	0.029 (0.058)	0.251*** (0.060)	0.468*** (0.091)	0.188*** (0.054)
Vil. Enviro Shocks	Yes	Yes	Yes	Yes	Yes
Vil. Ag. Extension	Yes	Yes	Yes	Yes	Yes
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	24192	23356	27090	19384	23922
F-Stat on Direct Effect	63.4	58.4	51.7	54.3	66.4

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 agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, the uninteracted shift, a village-level environmental shock dummy, and a village-level dummy for agricultural extension programmes. Standard errors clustered by district.

Table A15: Robustness: Crop Prices

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants (direct labour channel)	-1.061***	-0.748***	-1.283***	-1.894***	-1.386***
	(0.231)	(0.223)	(0.281)	(0.321)	(0.286)
Village emigration (indirect land channel)	0.237***	0.188***	0.276***	0.402***	0.254***
	(0.053)	(0.048)	(0.059)	(0.078)	(0.066)
Crop region emigration (indirect crop channel)	0.211***	0.028	0.277***	0.442***	0.238***
	(0.058)	(0.057)	(0.062)	(0.083)	(0.053)
Crop Price	Yes	Yes	Yes	Yes	Yes
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25054	24044	26462	20090	20444
F-Stat on Direct Effect	60.0	57.5	47.1	51.3	45.8

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 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 agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, the uninteracted shift, crop prices, drought, temperature, and rain. Standard errors clustered by district.

Table A16: Robustness: Crop Markets Measured in Crop Production Space

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants	-1.115***	-0.805***	-1.433***	-2.242***	-1.727***
(direct labour channel)	(0.290)	(0.278)	(0.353)	(0.417)	(0.361)
Village emigration	0.246***	0.214***	0.305***	0.480***	0.322***
(indirect land channel)	(0.067)	(0.062)	(0.075)	(0.107)	(0.085)
Crop region	0.292***	0.055	0.291***	0.573***	0.302***
emigration (indirect crop channel)	(0.075)	(0.079)	(0.079)	(0.116)	(0.074)
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	19366	18346	21066	15226	17430
F-Stat on Direct Effect	42.8	38.7	36.0	36.5	36.0

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 agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, the uninteracted shift, drought, temperature, and rain. Standard errors clustered by district.

Table A17: Robustness: Dropping Distance-Weighted Incomes as a Control

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Profit
Male Migrants	-1.320***	-0.971***	-1.252***	-1.818***	-1.371***
(direct labour channel)	(0.291)	(0.280)	(0.295)	(0.330)	(0.268)
Village emigration	0.288***	0.239***	0.261***	0.376***	0.248***
(indirect land channel)	(0.065)	(0.062)	(0.062)	(0.078)	(0.062)
Crop region	0.243***	0.065	0.264***	0.433***	0.181***
emigration (indirect crop channel)	(0.071)	(0.069)	(0.063)	(0.082)	(0.052)
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25924	24966	29342	20906	25852
F-Stat on Direct Effect	50.9	47.7	39.1	45.5	56.4

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. Distance-weighted income on its own is *not* included as a covariate. *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 5.1 for details). Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, drought, temperature, and rain. Standard errors clustered by district.

Table A18: Robustness: Direct and Indirect Effects (Nationwide Crop Markets)

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Output
Male Migrants	-1.089***	-0.748***	-1.262***	-1.966***	-1.385***
(direct labour channel)	(0.238)	(0.220)	(0.264)	(0.332)	(0.251)
Village emigration	0.277***	0.193***	0.310***	0.487***	0.283***
(indirect land channel)	(0.062)	(0.056)	(0.063)	(0.086)	(0.065)
Crop region	0.650**	0.348	1.187***	0.974**	0.857***
emigration (indirect crop channel)	(0.263)	(0.241)	(0.374)	(0.483)	(0.209)
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	25924	24966	29342	20906	25852
F-Stat on Direct Effect	61.1	57.8	50.6	55.9	64.4

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 India weighted by crop suitability (see Section 5.1 for details). Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, the uninteracted shift, drought, temperature, and rain. Standard errors clustered by district.

Table A19: Robustness: Direct and Indirect Effects (Extensive Margin)

	(1) Tech. Exp.	(2) Machinery	(3) Land	(4) Labor	(5) Output
Male Migrants	-1.406***	-1.011***	-1.291***	-2.125***	-1.350***
(direct labour channel)	(0.278)	(0.249)	(0.233)	(0.314)	(0.245)
Village emigration	0.270***	0.200***	0.254***	0.382***	0.230***
(indirect land channel)	(0.053)	(0.047)	(0.043)	(0.062)	(0.046)
Crop region	0.182***	0.039	0.197***	0.303***	0.155***
emigration (indirect crop channel)	(0.050)	(0.048)	(0.041)	(0.056)	(0.036)
HH FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓
Observations	69744	68686	75320	63958	66030
F-Stat on Direct Effect	71.5	69.9	66.0	65.0	63.4

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 stock of machinery, respectively. Column 3 is farm size in acres. Labor (column 4) is person-days of labor (household + hired workers). Profit (column 5) is crop revenue net of expenses. Standard errors clustered by district. All columns control for origin income, the uninteracted shift, drought, temperature, and rain. Standard errors clustered by district.

Table A20: Placebo Tests of Indirect Effects

	Crop Market Placebo		Land Market Placebo
	(1)	(2)	(3)
	Profit	Profit	Farm Size
Male Migrants	-1.351***	-1.321***	-0.820***
(direct labour channel)	(0.243)	(0.270)	(0.211)
Village emigration	0.248***	0.232***	0.140**
(indirect land channel)	(0.057)	(0.062)	(0.059)
Crop region	0.250***		0.228***
emigration (indirect crop channel)	(0.062)		(0.051)
Crop region	-0.099**		
emigration (placebo crop channel)	(0.040)		
Crop region		0.227***	
emigration (elasticity wt.)		(0.048)	
Crop region		0.059	
emigration (placebo: inv. elasticity wt.)		(0.038)	
Village emigration			0.064
(placebo land channel)			(0.053)
HH FEs	✓	✓	✓
Year FEs	✓	✓	✓
Observations	25852	20886	16116
F-Stat on Direct Effect	68.0	48.3	31.6

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The outcome is 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 non-agricultural households in the village. All columns control for origin income, the uninteracted shift, drought, temperature, and rain. Standard errors clustered by district.

Table A21: OLS Estimates—Distance to Nearest City, Agricultural Laborers, and Yields

	(1) % Δ Labor	(2) % Δ Yields
Distance to Nearest City (km)	0.007	0.030*
	(0.006)	(0.016)
State FEs	✓	✓
Observations	481	483
R^2	0.290	0.309

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. This table is the empirical analog of the map in Figure 4A and 4B. Distance to the nearest city is from the centroid of each district to the nearest of the 10 high-growth cities shown in Figure 4. In column 1, the outcome is percent change in the number of agricultural laborers per ha. between 2001 and 2011. In column 2, it is percent change in yields, measured as the value of all crops divided by cropped area. Standard errors are heteroskedasticity robust.

B Appendix Figures

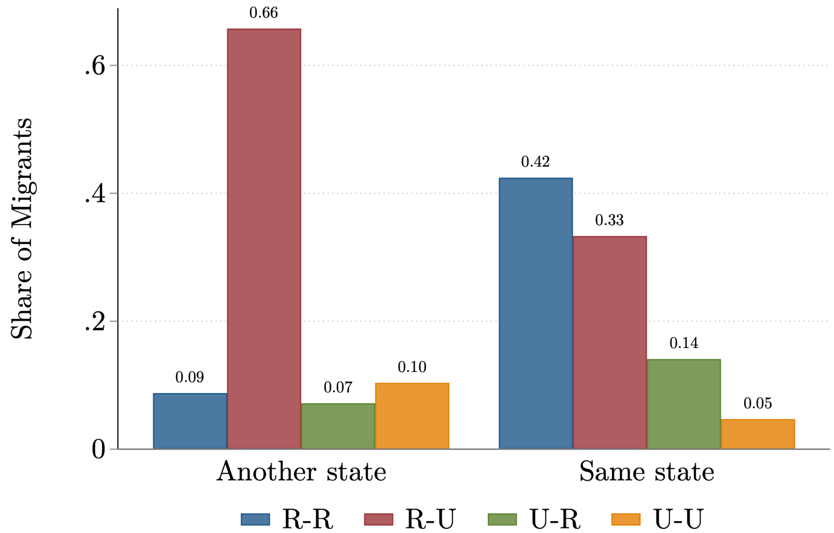


Figure B1: Migration Streams by Distance

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.

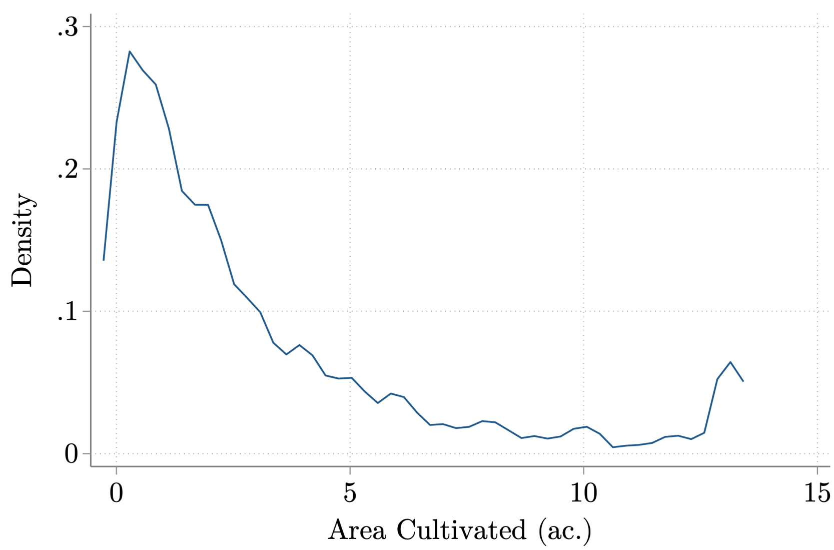


Figure B2: Distribution of Household Farm Size

Note: Kernel density plot of household farm size (cultivated area).

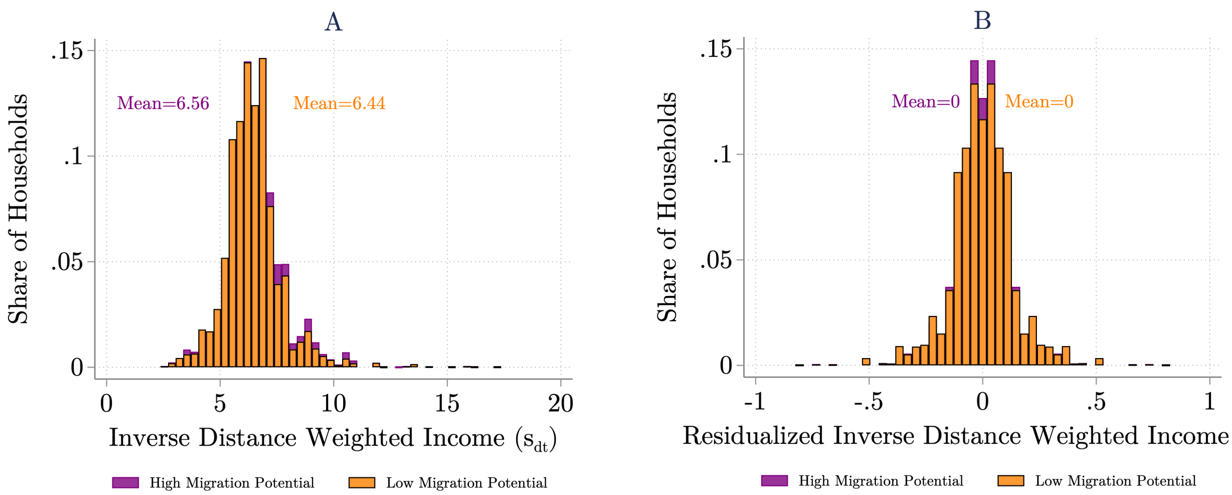


Figure B3: Distribution of Shocks

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 Theory 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 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; Breza et al., 2021). 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).[19] 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

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

such as irrigation water or immobile capital. Land prices, ρ_j, are endogenously determined 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.

Households differ in both labor productivity and the rural–urban productivity gap. Departing from Lagakos and Waugh (2013), where heterogeneity arises in both sectors, we assume it originates solely in the urban sector. Hence, variation in the rural-urban productivity gap reflects only differences in urban productivity. This assumption yields clear predictions about heterogeneous migration responses, which underlie our empirical strategy. Migration affects agricultural productivity; however, not through differential labor productivity in agriculture, as in Lagakos and Waugh (2013), but through differential local crop suitability and spatially heterogeneous migration responses driven by differential migration costs. Concretely, agricultural labor productivity is homogeneous across households, while urban labor productivity, φ_{ij}, varies. This generates heterogeneous responses to urban productivity shocks consistent with 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 the 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.[20]

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. Throughout this study, we assume that the urban productivity shock is exogenous to the rural sector, such as the effect of India's telecom and internet rollout on productivity in the service sector. 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:

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. □

[20] 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.

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 directly affect land and technology. 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.[21]

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.[22] 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.

C.3 Testing the Model's Predictions

In the main part of this paper, we empirically test these predictions, though we do not attempt a structural estimation of the model. Exploiting the panel structure of our data,

[21] 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. Similarly, migration could also affect rural households through investment in skills and education (Dinkelman and Mariotti, 2016; Khanna et al., 2022; Dinkelman et al., 2024).

[22] 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.

identification relies on changes in key variables rather than their levels. For instance, the agricultural labor loss from rural-to-urban migration is proxied by $-\frac{dl_{ijk}}{dw}$. More broadly, our empirical strategy absorbs aggregate changes over time and compares outcomes under migration with a counterfactual world absent migration but otherwise subject to the same dynamics (e.g., technological progress). For example, our estimates of migration's impact on food production are silent on whether food production is generally increasing or decreasing, since time fixed effects capture common temporal trends.

Testing Prediction 1 constitutes the first stage of our analysis. The insights from this prediction guide the construction of our migration instrument. To translate the model assumptions into empirical tests, we use Euclidean distance between the household and urban centers as a proxy for iceberg migration costs and the number of working-age male household members in the baseline period as a proxy for the household's urban productivity. The latter is justified by two observations: (i) productivity differences in agriculture between men and women are small (Agarwal and Mahesh, 2023), and (ii) men are roughly four times more likely to migrate (see Section 2.3), likely reflecting gender-specific migration barriers.

Testing Predictions 2 and 3 builds on the decomposition of direct and indirect effects underlying Prediction 3. Whether particular technologies are labor-saving or labor-complementary is often ambiguous. We therefore present results both for individual technologies and for a composite technology index. A complication in testing Prediction 3 is that the survey does not report agricultural output directly. Instead, we use agricultural profits as the dependent variable. By the Envelope Theorem, however, these estimates are theoretically equivalent. We discuss the details of this empirical strategy in Section 3.

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.[23] 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 conditions with respect to the distance to the urban center, τ:

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

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

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

The denominator, D is the determinant of the Hessian matrix times $(-\omega p)$. 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

[23] 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.

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

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

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:

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

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 $\frac{d\bar{\theta}_{ij}}{dw} = \sum_{k=1}^{K} \frac{d\theta_{ijk}}{d_w}$. 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

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