Internal Migration and the 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

January 30, 2024

Abstract

This paper studies how food production responds to agricultural labor loss during the process of urbanization. Building on a two-sector spatial equilibrium model with crop and land markets, we develop an estimation framework to track and quantify the ensuing spatial reorganization of agriculture. Using household microdata from India and exogenous variation in migration induced by urban income shocks, we document *direct effects* in response to labor loss and *indirect effects* as crop and land markets adjust. The direct effects are characterized by sharp declines in crop production, farm size, and technology investment among migrant-sending households predominantly located near cities. The indirect effects are characterized by greater production, technology adoption and farm size expansion among remote households with no migrants. Counterfactual simulations show that the indirect market spillovers mitigate over three-fourths of the direct agricultural losses driven by urbanization. This process leads to a spatial reorganization in which agricultural production moves away from land near urban areas with high emigration towards remote areas with low emigration.

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

*An earlier version of this paper is available as a *Yale EGC Discussion Paper (No. 1095)*. We thank Werner Antweiler, Marc Bellemare, Francisco Costa, Bruno Esposito, and Carol McAusland for their insightful comments. We also thank seminar and conference participants at the UC San Diego 21CIC Seminar, Canadian Resource and Environmental Economics Association Annual Conference, 30th Occasional Workshop in Environmental and Resource Economics, Agricultural Economics Workshop at Berkeley, 2022 AERE Summer Conference, Aix-Marseille School of Economics, OSWEET, PacDev, and Paris Dauphine University. Noack also thanks the Social Sciences and Humanities Research Council of Canada for their financial support (NFRFE-2018-01353). All errors are our own.

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

1 Introduction

Internal labor migration is perhaps the most pervasive feature of structural transformation. America's rise in the early 20th century was powered by dramatic worker migrations from farms to cities in search of manufacturing and service jobs (Alvarez-Cuadrado and Poschke, 2011). China's 20th-century economic boom witnessed rural workers moving *en masse* to urban factories: 147 million migrants were counted in China's 2005 census (Gao et al., 2022). In 1991, India liberalized its economy to stimulate the service sector, and GDP per capita tripled thereafter (Figure 1A). Labor shifted away from agriculture during that period: the urban population share grew by almost 30% while employment in agriculture dropped by more than 30% (Figure 1B).

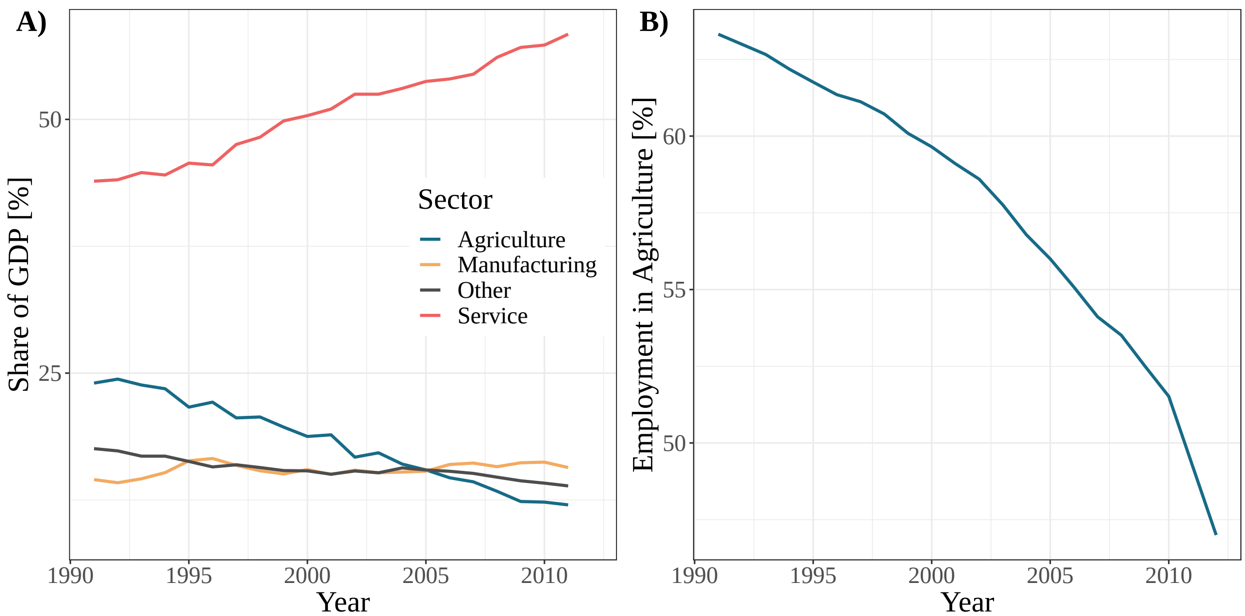
This paper documents the consequences of this structural transformation for India's agricultural sector and the spatial organization of the economy. How is food grown after workers leave farms for cities? Do farms replace labor with capital and modernize in response to agricultural labor loss? If not, then does food production shift to regions with less emigration? Beyond deepening our understanding of structural transformation, answering these questions is crucial for designing appropriate policies for populations remaining behind in rural areas.

The recent development economics literature argues that rural areas with high emigration rates become centers of agricultural modernization (Manuelli and Seshadri, 2014; Hornbeck and Naidu, 2014). This defies traditional views of factor endowments and comparative advantage, whereby labor-abundant agricultural regions specialize and become agricultural hubs (Matsuyama, 1992). The main contribution of our paper is to reconcile these views by documenting the direct effects (labor channel) and indirect or general equilibrium effects (market channel) of internal migration on agriculture in a single theoretical and empirical framework. We show that emigration induces a contraction of farm output in migrant-sending households through direct labor loss. We find no evidence that Indian farmers adopt labor-saving technologies in response to rural labor loss, as in Brewer et al. (2022) in Uganda. The contraction of agricultural production and reduced demand for farmland due to aggregate emigration causes non-migrant-sending households to increase production indirectly as land and crop markets adjust. The result is an internal spatial reallocation of agriculture in response to urban growth, towards non-migrant-sending households, and towards more remote rural areas with low emigration rates.

We start by developing a spatial household production model with rural regions and an urban center. Locations vary in terms of crop suitability and their migration distance to the city. Crop and land prices are determined by aggregate supply and demand. This setup allows migration to affect crop production directly through labor loss and indirectly through crop and land markets. Households across and within regions differ in their opportunity costs of agricultural labor, comprised of their migration costs and their urban productivity.

The model yields three predictions that we later take to the data. First, positive urban income shocks prompt rural-to-urban migration and a decline in crop production for migrant-sending households (hereafter "migrant households"). This agricultural contraction increases with proximity to the urban center and the urban productivity of the household. Rural labor and production

Figure 1: Economic Growth and Sectoral Re-allocation 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 Databook. Panel B reports the percent of total employment in agriculture using data from ILOSTAT database obtained through the World Bank Open Data Portal.

loss leads to higher crop prices and lower land prices. Second, the model predicts that households react directly to labor loss with reduced investments in labor-complementary technology and increased investment in labor-saving technologies. Third, the model predicts that the overall response of households to urban income shocks can be decomposed into a direct labor reallocation and investment effect and an indirect response to declining land prices and increasing crop prices. Which of those effects dominates depends on the urban productivity of the household, the aggregate labor losses within the same land market, and the aggregate production decline within the crop-producing area.

The model yields a parsimonious empirical equation for identifying the direct and indirect impacts of internal migration on agriculture. We derive a shift-share design (Goldsmith-Pinkham et al., 2020) from our theoretical predictions to recover the direct response of agricultural investments to the emigration-induced labor loss. The "shift" is given by an urban income shock, while the "shares" comprise household distance to the city and their urban productivity. We use the number of resident working-age males at baseline as a proxy for urban productivity of a household, since most migrant work opportunities in India are for young males. Although these shares are endogenous, we show that they are orthogonal to the shift, i.e., households with more males are not more exposed to urban income shocks (Borusyak et al., 2022a). The interaction of shift and share thus yields an instrument that exploits the exogenous variation in opportunity elsewhere induced by urban income shocks.

We expand this empirical framework to estimate the indirect crop and land market effects. We conceive a land market at the village level: when some households from the village emigrate in response to urban opportunities and divest from agriculture, agricultural land prices fall, which allows others in the village to expand production. We conceive a crop market operating at a more aggregated crop-growing region: when many others in the state who grow the same crop-mix emigrate to urban areas, that would induce each remaining household to expand production through market price effects. This captures the fact that households react more to changes in the aggregate supply of a crop if they also grow that same crop. We apply our empirical framework to the Indian Human Development Survey (IHDS), a detailed panel survey of 42,000 households conducted in 2005 and 2012 during a period of rapid urbanization.

Our analysis yields two main findings about factor reallocation in response to internal migration. First, we find an overall contraction in agricultural activity through the direct channel. There are sharp *declines* in technology adoption among migrant households, driven by reduced investment in labor-complementary technologies such as agrochemicals, irrigation water, and work animals. This leads to a reduction in crop output and a contraction of farm size.

Second, we find that these direct effects are counteracted by the indirect or general equilibrium effects that materialize through land and crop markets. Increased production through the indirect land-market channel offsets part of the migration-induced output loss from the direct labor channel. The mechanism is through farm expansion among non-migrant households, which aligns with our theoretical predictions. Similarly, output increases through the crop-market channel, again offsetting part of the migration-induced output loss from the direct channel. The mechanism is through additional investment in yield-enhancing technology among households with fewer migrants.

Overall, our results highlight the spatial incidence of direct and indirect agricultural effects of rural-to-urban migration. The (negative) direct effect dominates for households who live near cities and face low migration costs. The (positive) general equilibrium effects dominate for (non-migrant) households who live further away and face high migration costs. Although these remote households do not participate in structural transformation directly, they still experience production and technology benefits triggered by indirect market spillovers from those who do migrate.

We conduct a series of placebo exercises to show that the indirect crop market effect indeed captures price spillovers. The main specification measures indirect crop-market effects as aggregate emigration across households growing similar crop portfolios in the baseline period. As a placebo test, we show that households do not respond to aggregate emigration by households growing *different* crop mixes. They also do not respond to aggregate emigration from regions growing unrelated crops, as measured by cross-price elasticities. These findings help build confidence in a causal interpretation of our estimates.

Our findings imply that structural transformation triggers a spatial and inter-household reallocation of agricultural activities toward more remote areas and households without migrants. Put differently, labor reallocation *does* drive agricultural development, but not through capital substitution or other direct responses to labor loss. Instead, agricultural development spatially shifts through market forces, in the form of increased crop production, farm expansion, and technology adoption among remote, non-migrant households.

The paper concludes by investigating whether remote households fully pick up the slack left by urban migrant-sending households who reduce food production. This helps determine whether India has a "missing food problem" (Tombe, 2015). We implement a simulation in which special cases of the empirical specification aggregate to counterfactual scenarios. These include total crop value with no migration, with migration but no market spillovers, with migration and land markets but no crop markets, and so on. We find that the spatial reorganization of agriculture mitigated 84% of the aggregate agricultural production losses due to rural emigration. Quantitatively, the food "saved" by domestic reallocation amounts to Rs. 147 million, or 50% of total crop value. Note that the counterfactual is not the baseline period but a counterfactual world without migration. In other words, the simulation does not state that food production declined in absolute terms but rather that food production is reduced compared to a world without migration.

Literature Contributions: This paper contributes to at least three bodies of work. First, we extend research on labor reallocation and technology adoption in partial equilibrium. Rozelle et al. (1999) use cross-sectional survey data from China to show that migration reduces maize yields. Taylor et al. (2003) use the same survey to show that farm revenue falls among migrant-sending households. Similarly, Mendola (2008) use cross-sectional data from Bangladesh to show that migration reduces the adoption of high-yielding seed varieties (HYV). An exception to the cross-sectional literature is Brewer et al. (2022), who study migration and land use change in Uganda with panel data and a shift-share instrument. We advance this work by considering spatial spillovers.

Our second contribution is to decompose agricultural responses to structural transformation into direct (labor-driven) and indirect (market-driven) channels. In doing so, we advance the literature on structural transformation that uses data aggregated at higher 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 like tractors. Similarly, Clemens et al. (2018) find that reduced agricultural immigration increases the adoption of labor-saving technologies, although this does not compensate for the direct effects of labor loss, thus leading to lower crop output. Technology adoption may be a direct response to labor loss or an indirect response to changing sectoral compositions across the broader economy. In addition to including general equilibrium effects, our micro data also allow us to paint a richer picture of the post-migration adjustment process.

A related literature studies the effect of agricultural innovation on labor migration, the reverse of our question. Caprettini and Voth (2020) show with parish-level data that the introduction of threshing machines in England "released" agricultural labor and led to structural transformation (and riots). Bustos et al. (2016) use municipality-level data to show that HYV adoption in Brazil was labor-saving and prompted industrial growth. Caunedo and Kala (2021) study the adoption of labor-saving technologies and the implications for household labor reallocation in India.

Another related literature studies the impacts of agricultural productivity shocks across sectors. Emerick (2018) using district-yearly rainfall shocks in India to instrument crop productivity and find that positive shocks increase the *non-agricultural* labor share. The explanation is that rising district incomes increase demand for non-agricultural goods. Similarly, Moscona (2019) study the cross-sectoral impacts of India's green revolution. We extend this literature not only by studying the reverse scenario but also by documenting market spillovers within the same sector.

The third contribution of this paper is the characterization of direct and indirect effects from structural transformation in a single empirical framework. Previous studies have done this at small spatial scales (Blakeslee et al., 2023) or studied spillovers at different geographies separately (Asher et al., 2023). Our approach unifies this literature by identifying the spatial incidence of direct and indirect effects of structural change together in one framework. This enables comparisons of the two forces and counterfactual simulations of how much direct production losses are "saved" by spatial market-driven spillovers.

Methodologically, this paper joins the new literature on spatial general equilibrium estimation. We build our estimation framework on Adao et al. (2019), which develops a reduced form framework for estimating spatial equilibrium effects of economic shocks. Using the "China shock" from Autor et al. (2013) as a case study, they show that overlooking general equilibrium effects leads to severely biased estimates. Our results corroborate this finding since partial and general equilibrium effects draw in opposite directions in our theory and regression results.

The next section describes the data and four stylized facts about agriculture and migration in India. Based on these facts, we develop a spatial model of internal migration and agriculture in Section 3. We then use the model to construct our empirical strategy in Section 4. Section 5 presents our main results on the direct and indirect effects of labor reallocation on agriculture. Section 6 shows counterfactual simulations under different scenarios, and Section 7 concludes.

2 Data and Stylized Facts

This section describes the data we use to study how internal migration reorganizes agricultural activity. We define key variables and present four stylized facts about migration and agriculture that motivate the theoretical model in Section 3 and empirical framework in Section 4.

2.1 IHDS Survey

Individual-level migration data are sparse in India. An exception is the two-wave IHDS household panel, a nationally representative survey covering 384 districts (out of 594 at the time of the survey) across all states[1]. Wave I (2004-05) surveyed 41,554 households, of which 83% were located again in Wave II (2011-12). These households span 65% of districts across all states.

There are at least three advantages of using IHDS. First, it is among the few Indian surveys documenting emigration and agricultural production together. Second, households are interviewed

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

twice, enabling us to include household fixed effects in the analysis to control for time-invariant unobserved heterogeneity across households, such as caste or baseline poverty. Lastly, IHDS provides high-quality income measures. Survey staff disaggregate income into eight categories (e.g., crop income, business income, etc.) 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 here since 80% of dropouts did not own land in Wave I and would have been excluded from our analysis anyway. Second, seasonal migration (<6 months) is not reported in Wave II, restricting our analysis to medium-term migration.

2.2 Main Variables

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, destination, and other data about these migrant family members. The IHDS definition thus characterizes longer-term spells[2], such as a son who has been living away for the past five years. If he instead returned home after five months, then he would be a household resident and not a migrant. We exclude international migrants, which describes $< 10\%$ of migrants, as our paper focuses on internal migration.

The key migrant demographic for our analysis is working-age males, defined as age 15-60. To justify this window, Table B1 shows migrant types by cohort. The student migrant share sharply drops after age 14, at which time the share of employed sons jumps five-fold. This suggests that migrant males transition from school to work around age 15. Similarly, the share of husband migrants sharply drops after age 59, suggesting they return home around this time.

Agriculture: IHDS respondents engaged in farming report capital, labor, and crop income. The survey divides capital into input expenses, which include seeds, fertilizer, pesticides, irrigation water, and hired animals, as well as machinery, which includes tubewells, electric/diesel pumps, bullock carts, tractors, and threshers. There is also a labor expense sheet that documents wages paid for hired labor and person-days of unpaid family labor in the past year. Expenses are deflated to 2005 prices using the rural or urban Consumer Price Index, depending on household location[3].

Since capital and machinery are each described by five variables, we collapse them into indices to allay concerns of multiple hypothesis testing. We follow Anderson (2008) whereby each variable influences its index proportional to the information it adds. Intuitively, if seed expenses are highly correlated with other input expenses, it adds little to the expense index. We first demean the five variables making up capital expenses and machinery and divide them by the standard deviation of non-migrant households. Each index is computed as a weighted sum of these standardized values with weights equal to the row sum of the inverse covariance matrix. We report results for individual variables in the appendix.

[2]Short-term (<6 months) circular migration is only documented in Wave II.
[3]The price deflator is a pre-constructed variable distributed by IHDS.

We measure production with crop income in 2005 prices[4]. Crop income equals crop revenue minus expenses, where revenue is the product of quantity and price. About half of crop-growing households report sales price. The other half, who farm for personal consumption, report the crop price they *would* have received at market.

Crop Suitability: We measure gridded crop suitability (1km $\times$ 1km) as an index over the 1980-2010 period from the GAEZ FAO portal (Fischer et al., 2021). We obtain separate suitability rasters for each major crop in India[5] and normalize values to one. District suitability for each crop is computed by extracting means over all cells within a district.

Droughts: We measure droughts using the gridded (0.5-degree resolution) Standardized Precipitation-Evapotranspiration Index (SPEI) from the SPEIbase. Accounting for droughts in the empirical analysis is crucial because they affect labor movement and agricultural decisions. SPEI measures the difference between potential evapotranspiration and precipitation over a 12-month timescale. We extract the mean over cells within districts and then compute annual averages for 2005 and 2012 to match with the IHDS.

2.3 Four Stylized Facts About Migration and Agriculture

After summarizing our data, we describe four stylized facts from our data to motivate our theoretical and empirical framework.

Summary Statistics: Our sample frame consists of 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 B2 profiles the typical migrant. 80% are male, and nearly 60% are labor migrants in the first period. The labor migration rate jumps to 71% in the second period. This supports our focus on working-age males as the key demographic incentivized by urban work opportunities. The bulk of remaining migrants leave for education. Our working age window of 15-60 excludes the majority of these student migrants (Table B1).

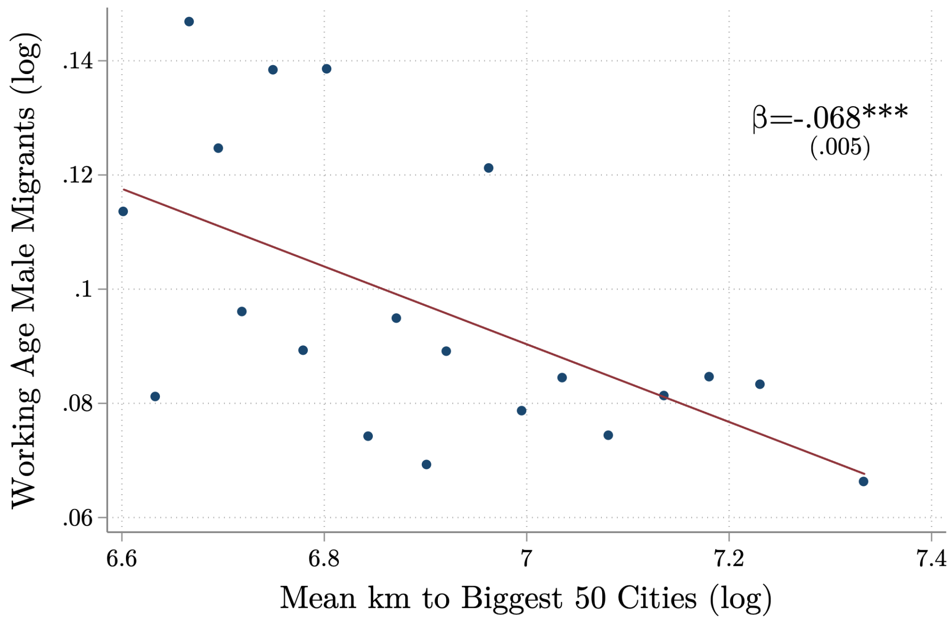
Most migration is within-state (Table B2 Panel C). Interestingly, rural-rural migration is important, accounting for 25% and 32% of migrants in Waves I and II, respectively. Figure C1 splits migration streams by inter- and intra-state travel, revealing a novel fact: among inter-state migrants, rural-urban migrants dominate rural-rural by 7-to-1, whereas rural-rural migration dominates among intra-state migrants. We therefore include all streams of migration in our analysis.

Fact I: Entire-Family Migration is Rare: Table B3 documents households that migrated *as an entire family* between surveys. These are defined as households surveyed in Wave II (2012) that reported having moved after Wave I (2005). Only 1% of households moved as a family. Among them, 85%

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

[5]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.

Figure 2: Binscatter Plot of Migration and Remoteness



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

migrated within the district. We, therefore, treat migration as a continuous (number of migrants to send) rather than a binary decision (move or stay) throughout the analysis.

Fact II: Remote Households Send Fewer Migrants: Figure 2 shows a binscatter plot of district remoteness (mean distance to the 50 biggest cities) against number of household migrants, residualized on year fixed effects. The inverse association implies that remote households send fewer migrants. The slope is the regression coefficient: households twice as remote send 7% fewer migrants, providing simple evidence of high migration costs to distant destinations.

Fact III: Land Markets are Active: Table 1 shows land, labor, and capital among farming households. Farming is small-scale, with the typical household cultivating three to four acres (Table 1 Panel A). The share of rented land almost doubles from 12% to 23% between survey waves, implying that land markets are active. Cultivation yields annual farm income of approximately 8000 Rs./acre across the two surveys. Fertilizer is the largest expense (Panel B), amounting to 10% of farm income on average. Water pumps are the main machinery, while large equipment like tractors are rarely owned (Panel C).

Fact IV: Rural Labour Markets are Malfunctioning: Family contributes over 15 times more person-days of farm labor than hired workers each year (Panel D). In fact, only 6 % of total farm labor is

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

	IHDS-I (2004-05)		IHDS-II (2011-12)	
	Mean	SD	Mean	SD
A: Land				
Area cultivated (ac.)	3.16	4.10	3.84	4.82
Area rented in (ac.)	0.39	1.87	0.87	7.80
Yield (Rs./ac.)	9213.11	31386.59	7381.06	28087.75
B: Expenses/Acre				
Seeds	748.83	1342.19	657.80	3184.63
Fertilizer	1078.16	2493.72	894.15	4757.34
Pesticides	336.49	1164.81	277.97	2491.50
C: Machinery (Num./Acre)				
Pumps	0.17	1.06	0.11	0.82
Tractors	0.02	0.14	0.01	0.13
Bullock Carts	0.07	0.31	0.04	0.22
D: Labour (past yr.)				
Family person-days/acre	267.67	611.42	204.59	664.34
Hired person-days/acre	17.97	53.86	12.82	121.28
Wages Paid/Acre	815.72	2276.32	871.86	6307.49

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

hired. This is evidence of malfunctioning rural labor markets in India, in line with the literature (Fernando, 2022; Foster and Rosenzweig, 2022). It also helps justify our assumption of absent rural labor markets in the model (Section 3).

3 Model

This section develops a spatial equilibrium model of household production in the face of urban growth motivated by the empirical facts in the previous section. We use the model to decompose household responses into direct effects through labor reallocation and indirect market effects. The model yields estimable equations that quantify these response channels in the data. Although the model is adapted to our current context, it builds on general spatial trade models.[6]

3.1 Set-up

Environment: The economy 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. Multiple villages

[6] See Allen and Arkolakis (2023) for a recent overview of the literature.

can be equally suitable for the same crops. The urban center uses labor from surrounding rural areas to produce the service goods. Space is characterized by distance to the urban center, which determines the opportunity cost of agricultural labor. Residents of remote villages face lower opportunity costs of agriculture than those in villages close to the urban center.

Markets: Rural labor markets are absent, while land, product, and technology markets are frictionless. The labor market assumption draws on the evidence of malfunctioning rural labor markets in India (Foster and Rosenzweig, 2022). The simplifying assumption of frictionless land, product, and technology markets allows us to focus on the impact of rural-to-urban migration on agriculture while abstracting from other market imperfections. Frictionless product markets equalize the crop price, p_k, across villages similar to the assumptions in Borusyak et al. (2022b).[7] Crop prices are endogenously determined by aggregate supply within the economy.

We use the term land in a wider sense to also include other immobile production factors 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 (see Section 2.3).

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

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

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

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

where Y is aggregate income and Y_k is aggregate supply of crop k. Unlike agricultural goods,

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

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.

$$\frac{\partial^2 f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk} \partial a_{ijk}} > 0 \tag{1}$$

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

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:

$$\max_{a_{ijk}, \theta_{ijk}, l_{ijk}} \sum_{k=1}^{K} \left[p_k \omega_{jk} f(l_{ijk}, a_{ijk}, \theta_{ijk}) - \rho_j a_{ijk} - \nu \theta_{ijk} \right] + \frac{w \varphi_{ij}}{\tau_j} l_{ijw} \quad \text{s.t.} \quad \sum_{k=1}^{K} l_{ijk} + l_{ijw} = 1 \tag{3}$$

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 $l_{ijk'}, a_{ijk'}, \theta_{ijk'} = 0$).

3.2 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 A.

3.2.1 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 A.1.[8]

3.2.2 Urban Productivity Shocks and the Reorganization of Agriculture

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

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.*

[8] 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.

Proof. The proof is given in Appendix A.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 4.2.

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 A.3. □

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

Urban productivity does not affect land and technology directly. It does so only through labor reallocation and corresponding changes in crop and land prices. We, therefore, describe the impact of urban productivity shocks on crop production through labor and price adjustments alone.[9]

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

$$\frac{dy_{ijk}}{dw} = \frac{dl_{ijk}}{dw}\phi_{1ijk} + \frac{d\rho_j}{dw}\phi_{2ijk} + \frac{dp_k}{dw}\phi_{3ijk} \tag{4}$$

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

Proof. The proof is given in Appendix A.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.[10] 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

[9] 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.

[10] In Appendix A.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.

A.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 becomes an empirical question.

3.3 From Theory to Empirics

The previous section showed that the effect of labor reallocation on agriculture comprises opposing partial (direct) and general equilibrium (indirect) forces. Next, we show how to map Prediction 3 into an empirical framework to test the net effect of urbanization on agriculture. The mapping rests on the assumption that the theoretical parameters ϕ_{1ijk}, ϕ_{1ijk} and ϕ_{1ijk} are independently distributed with mean β. This enables us to estimate corresponding empirical parameters β_1, β_2, and β_3 as constant terms despite their theoretical counterparts varying across crops and households. We test effect heterogeneity in several robustness tests (Section 5.2.1 and 5.3.1).

Direct Effect of Migration on Agriculture: ϕ_{1ijk} of Prediction 3 describes the direct effect of migration on agricultural production. This parameter can be directly estimated with household data on labor migration and crop output. Since migration is endogenous, we use Prediction 1 to build an instrument for the migration decision in Section 4.2. The instrument incorporates heterogeneous responses by distance to the center and the urban productivity of the household. Prediction 1 describes two dimensions of exposure to urban shocks. The first relates to distance and states that households are more exposed to nearby shocks. The second relates to urban productivity and states that households with greater urban productivity are more exposed to wage shocks conditional on distance. We interact these two terms to build a shift-share instrument for migration.

Indirect Effect of Migration on Agriculture through Land Markets: ϕ_{2ijk} of Prediction 3 describes the indirect effect of migration on agriculture through land markets. Since land markets typically operate at the village level, aggregate village emigration captures changes in land prices in response to urban shocks. The link is through scarcity: the marginal productivity of land declines with village agricultural labor. Unlike the direct effect, we do not instrument aggregate village emigration[11] under the assumption that migration choices of others in the village are beyond household i's control.

Indirect Effect of Migration on Agriculture through Crop Prices: ϕ_{3ijk} of Prediction 3 describes the indirect effect of migration on agriculture operating through crop markets. To measure this, we first observe that crop prices are determined by aggregate demand and supply of crop k.[12] We assume that demand and supply effects are additively separable and uncorrelated. The supply-side effect is determined by output shifts in villages with similar crop suitability and specialization,

[11] Leaving the index household's own migration decision out of the emigration rate construction

[12] Appendix A.4 shows that that the response of household crop production is a linear combination of crop-household-level production.

while year fixed effects absorb aggregate demand effects. We conceive crop markets as state-level objects because of India's cross-state agricultural trade restrictions (Chatterjee, 2023).

Aggregate emigration of households from similar crop-growing regions thus identifies the supply side effect. The exact construction of this variable is defined in the next subsection. Similar to the indirect land market channel, we do not instrument aggregate migration from the crop-growing region because that is beyond the control of any individual household. Instead, we use placebo tests to build confidence in our identification strategy, showing that aggregate emigration of other households growing *unrelated* crops has no effect on each household's agricultural decisions.

3.4 Measuring Indirect Effects of Migration on Agriculture

Land Market Channel: Since land markets are at the village level, we measure the impact of emigration on land markets with aggregate emigration by all other households i' of village j, in district d, excluding household i, as follows:

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

where M_{ijdt}^{land} measures mean migration from households $i' \neq i$ in the same village. $Migrants_{i'jdt}$ is the number of working-age male migrants sent from household i' in village j at time t. N_{jdt}/i is the set of households in village j excluding household i and $|N_{jdt}/i|$ is the number of elements (cardinality) 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 responses to changes in M_{ijdt}^{land} capture responses through the indirect land channel.

Crop Market Channel: We measure the crop channel as aggregate emigration from the crop-growing region of household i of village j in district d:

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

where N_s/i denotes all households in state s excluding household i. We define crop markets at the state level because agricultural markets are heavily regulated in India, and cross-state trade is restricted, preventing crop prices from equilibrating across states (Chatterjee, 2023). We let the market span the whole country in a robustness test.

The crop-growing region of household i consists of all other households in state s that grow similar crops at baseline. To quantify the idea of similar crop portfolios, consider two vectors $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{iK})$ and $\mathbf{x}_{i'} = (x_{i'1}, x_{i'2}, ..., x_{i'M})$ that list all possible K crops that households i and i' can grow. The number of crops and their order is identical for both households. x_{ik} is crop k's share in total production of household i. The share equals zero if the household does not grow crop k. The crop similarity between household i and i' is the inverse Euclidean distance between their crop portfolios[13]:

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

These pairwise crop similarities reveal other households growing the same crops in similar proportions at baseline. We interpret this similarity as an indicator of being in the same crop-growing region. The value $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ enters as a weight when aggregating labor migration among other households (Equation 6). Since migration by working-age men from other households growing similar crops to i receives more weight in the aggregation, M^{crop}_{ijdt} effectively measures aggregate emigration from the same crop-growing region as i. Household responses to changes in M^{crop}_{ijdt} thus capture adjustments through crop markets. Intuitively, if urban shocks trigger widespread labor migration from a rice-growing region, affecting aggregate rice supply, then household i elsewhere in the state will respond to the price change if they also grow rice.

As a robustness test, we measure responses through the crop channel using pairwise similarities in crop-specific agricultural suitability space (Section 5.3.1). This helps account for the potential endogeneity of crop choices. For the test, $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ is constructed in the same way, except K defines the set of India's major crops, and $x_{ik} \in \mathbf{x}_i$ is the suitability index for crop k in the district of household i.

4 Empirical Framework

This section develops an empirical framework for estimating the direct and indirect effects of internal migration on agricultural development. First, we introduce our baseline specification. Second, we propose a shift-share instrument for internal migration based on Prediction 1 from the model. Lastly, we test instrument validity through various validation exercises.

4.1 Baseline Equation

We established that the direct and indirect effects of migration on agriculture can be described by an additively separable function of household migration, village migration, and migration from the crop-growing area (Prediction 3 and Section 3.3). The regression equation analog is:

$$Y_{ijdt} = \beta_1 M^{labor}_{ijdt} + \beta_2 M^{land}_{ijdt} + \beta_3 M^{crop}_{ijdt} + X_{dt} + \alpha_i + \gamma_t + \varepsilon_{ijdt}$$

where Y_{ijdt} are agricultural outcomes for household i of village j in district d at time t. M^{labor}_{ijdt} is the number of working-age male migrants sent from household i. β_1 captures household responses to its own labor migration through the direct labor channel. M^{land}_{ijdt} is aggregate emigration among other households in the village (Equation 5). M^{crop}_{ijdt} is aggregate emigration among other households in the crop-growing region of household i (Equation 6). β_2 and β_3 thus capture household

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

adjustments through the indirect land and crop market channel, respectively. X_{dt} is a vector of covariates that jointly influence labor and agricultural decisions, such as drought conditions. Household fixed effects, α_i, absorb time-invariant differences between households, such as distance to urban centers, crop suitability, and location, which jointly affect labor and agricultural decisions. Demand-side effects from increased urban productivity are captured by year fixed effects, γ_t.

Endogeneity Concerns: The endogeneity of M^{labor}_{ijdt} is our main concern. Although α_i absorbs factors such as baseline wealth and household composition, changes in these characteristics could influence migration and agricultural decisions. Furthermore, improvements in agricultural technology can release surplus labor, leading to reverse causality. The next section introduces a shift-share instrument to address the endogeneity of M^{labor}_{ijdt}. In contrast, we do not instrument M^{land}_{ijdt} and M^{crop}_{ijdt} since aggregate emigration is largely outside household i's control.

4.2 Shift-Share Instrument Design

We construct a shift-share instrument for migration based on Prediction 1 from the model. The instrument combines income shocks at each destination (the shift) with measures of household exposure to these shocks (the share). The combination of shift and share yields an instrument that isolates the pull stream of migration that is plausibly orthogonal to push factors at the origin.

The Shift: Income shocks, $inc_{d't}$, form the "shift" of our shift-share design. This is measured as mean IHDS-reported income across surveyed 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 ensure that urban destinations are more attractive than rural ones (Table B2).

The Shares: We use two shares based on Prediction 1, which states that the strength of the urban pull is a weighted combination of household distance to the city, τ, and their urban productivity, φ. The first share is computed as the inverse distance from each origin district d to every potential destination district $d' \in \Theta/d$, denoted as $\frac{1}{\tau_{dd'}}$[14]. Since Θ spans all districts, rural-rural migration is allowed in the underlying structure (Section 2.3). Inverse distance weights build a gravity structure where potential migrants are attracted more to urban destinations but less so the further away they are. We use a distance elasticity of one based on similar values from the literature (Bryan and Morten, 2019; Schwartz, 1973) and our theoretical framework.

The second share is urban productivity of the household, φ. 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. Although female household members may work on the farm, males dominate labor migration in India (Table B2).

Combining Shift and Share: Finally, we combine the shift and the share to form our shift-share

[14] Distances are between district centroids from the official 2001 District Census shapefile. All distances are measured in kilometers using the WGS84 projection.

instrument for labor migration, z_{ijdt}, as follows:

$$z_{ijdt} = \varphi_{ijd} \sum_{d' \in \Theta / d} \frac{inc_{d't} \times pop_{d'}}{\tau_{dd'}} \tag{7}$$

where φ_{ijd} are the number of baseline working-age males in household i in district d, $inc_{d't}$ is mean household income in other districts $d' \neq d$, $pop_{d'}$ is other-district population from the 2001 Census, and $\tau_{dd'}$ is the distance between district d and d'. Section 4.4 tests the instrument validity.

4.3 Two-Stage Least Squares

Equipped with the shift-share instrument, z_{ijdt}, we now specify the effect of labor reallocation on agricultural outcomes in a two-stage least squares framework:

$$M^{labor}_{ijdt} = \mu_1 z_{ijdt} + \mu_2 s_{dt} + \mu_3 inc_{dt} + \mu_4 M^{land}_{ijdt} + \mu_5 M^{crop}_{ijdt} + X_{dt} + \alpha_i + \gamma_t + \varepsilon_{ijdt} \tag{8}$$

$$Y_{ijdt} = \beta_1 \widehat{M^{labor}_{ijdt}} + \beta_2 M^{land}_{ijdt} + \beta_3 M^{crop}_{ijdt} + \beta_2 s_{dt} + \beta_3 inc_{dt} + X_{dt} + \alpha_i + \gamma_t + \eta_{ijdt} \tag{9}$$

where $s_{dt} := \sum_{d' \in \Theta / d} \frac{inc_{d't} \times pop_{d'}}{\tau_{dd'}}$, the distance-weighted income shock.

First Stage: Equation (8) is the first stage, which relates labor outflows from the origin, M^{labor}_{ijdt}, to the instrument, z_{ijdt}, controlling for the independent effect of the income shock, s_{dt}. We control for mean origin household income, inc_{dt}, to account for spatial correlation between destination and origin income shocks. α_i and γ_t are household and year fixed effects, respectively. To the extent that z_{ijdt} is plausibly exogenous (see Section 4.4) conditional on fixed effects and controls, predicted migration, $\widehat{M^{labor}_{ijdt}}$, from this equation represents the pull stream of migration that is orthogonal to push incentives at the origin. μ_1 thus represents the labor response to urban income shocks and its coefficient describes the curvature of agricultural production (Appendix A.2).

Second Stage: Equation (9) estimates the second stage impact of labor migration on agricultural activity, Y_{ijdt}, of household i in village j and district d at time t. The response comprises the direct effect of migration through the labor loss channel, β_1, as well as the indirect effects through land and crop markets, captured by β_2 and β_3, respectively. These three coefficients are the empirical analogs of ϕ_1, ϕ_2, and ϕ_3 in Prediction 3 of the model.

The indirect channels are not instrumented with the z_{ijdt}. This specification choice assumes that migration of other households in the village and state is unrelated to the decision of household i, conditional on controls and fixed effects. The main threat to this assumption is that village shocks (e.g., a local drought) may co-determine agricultural outcomes and village emigration. We address this by including origin income shocks and local drought conditions as controls. Another threat is that demand shocks in urban centers may determine agricultural development for nearby regions through channels other than emigration. We address this by controlling for the distance-weighted shock, s_{dt}, directly in both stages of the regression.

4.4 Instrument Validity

This section explores instrument validity in detail. Equation (9) yields a consistent estimate of β_1 if migration is the only channel through which z_{ijdt} affects Y_{ijdt}. Recent literature shows that consistency in shift-share designs can be achieved if either the shift (Borusyak et al., 2022a) or the share (Goldsmith-Pinkham et al., 2020) is exogenous.

4.4.1 Endogenous Shares

Our shift-share design is unique in that we have two shares: distance to potential destinations, $\tau_{dd'}$, and urban productivity, φ_{ijd} (Section 4.2). Household location is potentially endogenous since households near cities face different incentives and opportunities than their remote counterparts. Time-constant differences between locations are absorbed by household fixed effects. However, agricultural outcomes may also develop differently in remote areas, which generates a mechanical correlation between the distance-weighted income shock, s_{dt}, and outcomes of interest. Therefore, we include s_{dt} directly as a control, leaving identification to rely on differences in household exposure to the urban income shocks, conditional on the shock itself.

Addressing the endogeneity of migration potential, φ_{ijd}, is less straightforward. Households with many working-age males may differ from those with few in many ways. The time-invariant component of these differences is absorbed by household fixed effects. However, household composition may also predict *changes* in agricultural outcomes *even in the absence of migration*. In the next section, we address this with visual evidence and a series of falsification tests.

4.4.2 Exogenous Shifts

Given that φ_{ijd} is likely endogenous, the validity of our shift-share design hinges on the assumption that shifts—the destination income shocks—are as-good-as-randomly assigned conditional on controls and fixed effects (Borusyak et al., 2022a). If this holds, we can rule out the possibility that s_{dt} picks up differential time paths of households with high and low migration potential.

Visual Evidence: Distribution of Shocks: We characterize the distribution of s_{dt} across households with high (above-median) and low (below-median) migration potential, which reveals whether either type differentially experiences the shock (Borusyak et al., 2022a). Figure (3)A shows that the distributions are almost identical. The similarity is even more evident after residualizing on household and year fixed effects (Panel B). Overall, households in both groups are exposed to similar levels of the shock, which provides evidence *against* the possibility that the shock picks up correlated characteristics of more and less exposed households.

Formal Evidence: Falsification Tests: While Figure (3) provides supporting evidence for shift-share validity, it is still possible that z_{ijdt} is correlated with other unobserved household characteristics that affect outcomes through non-migration channels. For example, if districts more exposed

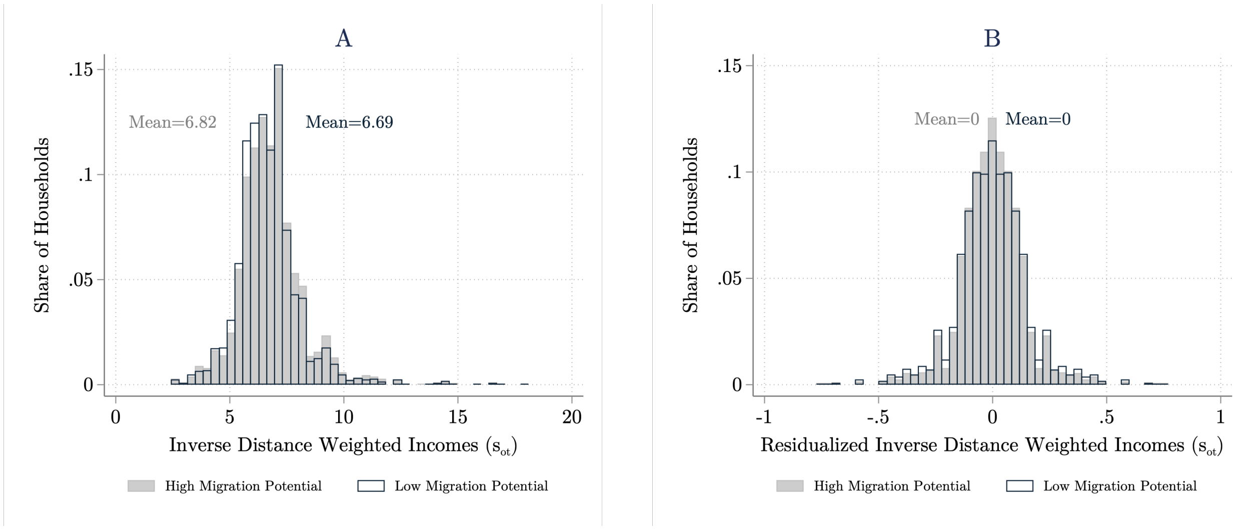


Figure 3: Distribution of Shocks

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.

Table 2: Shock Balance Tests

	(1) Males	(2) Educated	(3) Farm Size	(4) Ag. HH	(5) Landowner
Δ Wt. Income	0.078	0.235***	-0.943	0.044	-0.070
	(0.068)	(0.079)	(0.575)	(0.037)	(0.047)
Δ Origin Income	Yes	Yes	Yes	Yes	Yes
State FEs	✓	✓	✓	✓	✓
Observations	38589	38589	12355	38589	38588
R^2	0.021	0.020	0.103	0.051	0.088

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

to destination income shocks have more land-owning households, then these households may face different agricultural growth trends even in the absence of migration.

We follow Borusyak et al. (2022a) and Xu (2022) and formally rule out such violations of the exclusion restriction through a set of falsification tests. We choose a set of baseline confounders, X_{ijd}, and regress them on the *change* in the shock, $\Delta s_d := s_{dt} - s_{dt-1}$, in a pooled cross-section:

$$X_{ijd} = \beta_1 \Delta s_d + \beta_2 \Delta inc_d + \gamma_s + \varepsilon_{ijd}. \tag{10}$$

Table 3: First Stage: Distance-Weighted Productivity Shocks and Internal Migration

	(1)	(2)	(3)
Wt. Income × Working Age Males (z_{ijdt})	0.039***	0.046***	0.046***
	(0.005)	(0.005)	(0.005)
Wt. Income (s_{dt})	No	Yes	Yes
Origin Income (inc_{dt})	No	No	Yes
Outcome Mean	0.201	0.201	0.201
HH FEs	✓	✓	✓
Year FEs	✓	✓	✓
KP (2006) F-Stat	58.29	74.45	74.82
Observations	25032	25032	25032

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 at the district level (s_{dt} in Equation 7). "Working Age Males" is the number of male household residents aged 15-60 (φ_{ijd} in Equation 7). "Origin Income" is mean per capita district income, inc_{dt}. All specifications control for drought conditions. Standard errors clustered by PSU.

β_1 is the balance coefficient. If Δs_d is as-good-as-randomly assigned to households, it should not predict X_{ijd}. Δinc_d is the change in the origin district income, which accounts for demand effects (Section 4.4.1), and γ_s is a state fixed effect. Table 2 reports balance test results for the number of household working age males (φ_{ijd}), number of educated household members, farm size, earning income primarily from agriculture, and land ownership.

The first column presents the key result: the shock is uncorrelated with the share. We control for household education in the robustness checks below since it appears that the shock is disproportionately felt by more educated households. Apart from education, all other estimates in Table 2 are consistent with the as-good-as-randomly assigned hypothesis.

5 Results

We now present estimates of the direct and indirect effects of labour loss on agriculture using the framework in Section 4. We first show first-stage estimates before turning to estimates of the direct and indirect effects.

5.1 First Stage

Table 3 reports first stage estimates and Kleibergen-Paap F-statistics (Kleibergen and Paap, 2006) for instrument relevance. Column 1 does not include controls. Column 2 adds controls for the direct shock, s_{dt}. Column 3 is the preferred specification and controls for both s_{dt} and origin income, inc_{dt}. The shift-share instrument strongly predicts labor outflows, consistent with Prediction 1.

Table 4: Second Stage—Direct Effect of Migration on Agricultural Development

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Person-days	(6) Crops
Male Migrants (M^{labor}_{ijdt})	-2.513***	-2.110***	-2.269***	0.253	-4.890***	-2.913***
	(0.585)	(0.532)	(0.561)	(0.271)	(0.863)	(0.561)
Wt. Income (s_{dt})	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income (inc_{dt})	Yes	Yes	Yes	Yes	Yes	Yes
Outcome SD	-	-	4.392	547.124	344.519	30291.547
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25928	24970	28748	20910	20910	25032
F-Stat	58.9	55.2	53.9	55.6	55.6	66.5

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The explanatory variable is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. All outcomes are in standard deviations. All specifications control for district drought conditions. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (Section 3.4). Wage bill (column 4) is total wages paid to all workers in the past year. Person-days (column 5) includes both household and hired labor. Crop profits (column 6) are net of expenses. Standard errors clustered at village level.

The average household has 1.76 resident working-age males. Column 3 thus implies that a one-unit increase in destination incomes pulls 2.6% (=0.046/1.76) of them away to join the destination labor force. F-statistics in all specifications are well above rule-of-thumb levels.

5.2 Second Stage: Direct Effects (Partial Equilibrium)

Table 4 shows estimates of Equation (9) *without* the M^{land}_{ijdt} and M^{crop}_{ijdt} terms. Outcomes are in standard deviations to facilitate comparisons across outcomes. The results show that households respond to labor loss by divesting from agriculture. Column 1 shows that losing a worker causes households to reduce agricultural expenses by roughly 2.5 standard deviations. This is driven by reduced spending on seeds, agrochemicals, irrigation water, and rented equipment (Table B4, columns 1-5). Column 2 of Table 4 shows that labor loss also causes households to reduce their stock of machinery. This is driven by lower ownership of tubewells, water pumps, bullock carts, and threshers (Table B4, columns 6-10). At least eight of the technologies in Table B4 are labor complementary, corroborating Prediction 2 from the model.

Columns 3-6 of Table 4 explore additional margins of response. Declining labor and technology use reduces the marginal productivity of land, which may prompt farm contraction. Column 3 corroborates this logic: the point estimate implies that labor loss causes households to reduce cultivated area by 2 standard deviations. In contrast to land markets, labor markets respond imperfectly. There is no impact on wages paid for hired labor (column 4), suggesting that households do not replace household labor with hired labor. Column 5 shows a decline in total person-days worked. If the migrant were replaced with hired or family labor, there would be no effect. Thus, the estimate in column 5 is likely driven by the migrant himself. Overall, the (lack of) labor response is evidence of India's malfunctioning rural labor markets (Fact IV, Section 2).

The output contraction in column 6 is expected, given the downward adjustment of technology and land (columns 1-3). WE find that agricultural profits in the past year declined by 3 standard deviations, corresponding to about Rs. 90,000.

5.2.1 Robustness and Heterogeneity of Direct Effects

Robustness of Direct Effects. These estimates are remarkably robust. Online Appendix D provides a thorough investigation of robustness to a variety of alternative specifications and measures of the shift-share instrument. Panel A of Table B14 tests robustness to state-year fixed effects. Panel B adds the extensive margin—households that left or joined agriculture between 2005-12—to the sample. Panels C and D control for household education and pre-period trends, respectively. Panels E-G test robustness to isolating rural-urban migration, measuring the shock with nightlights instead of income, and expanding the migrant definition to include all genders. Estimates are remarkably stable across all specifications. Table B15 implements alternative clustering methods to account for spatial correlation across shifts and shares.

Heterogeneity of Direct Effects. Prediction 3 stated that the direct response of households to labor loss, ϕ_1, is household-specific. In the model, households differ only by their village's distance to the urban center (τ) and their urban productivity (φ). We split the sample along these two dimensions to estimate heterogeneity of direct effects. Table B5 splits the sample by above- and below-median distance to the nearest city, which is reported in the village module of the IHDS survey. Directionality and statistical significance of coefficients among urban (Panel A) and remote (Panel B) households are comparable to the baseline estimates. Urban households reduce technology and farm size more than remote households (columns 1-3), in line with Prediction 1.

Table B6 splits the sample by above- and below-median number of resident working age males at baseline, our measure of φ. We are underpowered in the below-median case since the median is two and the majority of below-median values are zero. In the above-median sample, directionality and significance mirror the main estimates, whereas coefficient magnitudes are larger.

5.3 Second Stage: Direct and Indirect Effects (General Equilibrium)

At first glance, our findings in the previous section suggest that India is failing to modernize the agricultural sector, which prompts the question of how the food needs of 1.5 billion people will be met as urbanization unfolds. However, we have thus far painted only half the picture. This section completes the analysis by presenting estimates of the *indirect* effects of migration on agriculture. We document a counterbalancing pattern of agricultural modernization in remote regions.

Main Estimates: Table 5 presents estimates of Equation (9). As predicted by the model, the direct

Table 5: Second Stage—Direct and Indirect Effects of Migration on Agricultural Development

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Person-days	(6) Crops
Male Migrants	-1.147***	-1.007***	-1.064***	0.124	-2.332***	-1.441***
(direct labour channel)	(0.259)	(0.252)	(0.275)	(0.130)	(0.413)	(0.318)
Village emigration	0.235***	0.244***	0.231***	-0.030	0.492***	0.279***
(indirect land channel)	(0.058)	(0.057)	(0.060)	(0.029)	(0.099)	(0.075)
Crop region emigration	0.287***	0.098*	0.184***	-0.003	0.515***	0.246***
(indirect crop channel)	(0.060)	(0.056)	(0.054)	(0.015)	(0.089)	(0.061)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25662	24632	26524	20644	20644	20150
F-Stat on Direct Effect	63.9	56.3	53.3	54.9	54.9	50.1

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All variables are standardized. *Male Migrants* is the number of working-age male migrants sent from the household, instrumented with inverse-distance population weighted income interacted with 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 portfolios (Section 3.4). Columns 1 and 2 are indices for expenses and machinery, respectively. Wage bill (column 4) is wages paid to hired labor in the past year. Person-days (column 5) includes family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

and indirect effects draw in opposing directions. The direct effect remains negative and statistically significant in all columns (row 1). However, the indirect effects counterbalance the agricultural decline. Columns 1 and 2 show that declines in agricultural technology caused by migration are partially offset by the indirect crop channel (row 3). The model states that rising crop prices induced by the migration-induced supply contraction prompt non-migrant farmers to adopt technology. Indeed, Table B7 columns 1-5 show that technology uptake through the crop channel materializes through spending more on seeds, agrochemicals, irrigation water, and rental equipment. There is little change in machinery ownership (columns 6-10). We corroborate the crop price hypothesis through a placebo test further below.

Column 3 of Table 5 shows that downsizing farms through the direct channel is partially offset by the indirect land channel (row 2). This is in line with the model prediction that falling land prices induced by the direct effect prompt non-migrant households to expand farm size. Columns 4 and 5 describe the labor market responses. Wages paid to hired workers are unaffected by aggregate emigration at the level of the crop-growing region.

Since technology and land are the main factors of production, changes in output[15] (column 6) show the same directionality through each channel. The point estimates are precise and eco-

[15]Since indirect (price) effects are observed in the specification, and nationwide price changes are absorbed by year fixed effects, coefficients in Column 1 can be interpreted as output changes even though the measure is profits.

nomically significant. Whereas households reduce expenses on agricultural technology by 1.15 standard deviations in response to their own labor loss (column 1), mechanization through the indirect crop channel offsets this by 25% ($= 0.287/|1.147|$). The counterbalancing pressures are similar for farm size changes (column 3) and total production (column 6).

Placebo Tests: We provide two placebo tests that both validate our measure of the indirect crop channel, M_{ijdt}^{crop}, and support a causal interpretation of the results in Table 5. First, we compute aggregate emigration from parts of the state growing *different* crop mixes from household i. We do this by replacing $d(\mathbf{x}_i, \mathbf{x}_{i'})^{-1}$ in Equation (6) with $d(\mathbf{x}_i, \mathbf{x}_{i'})$. Since these euclidean distance weights are not inverted, households growing *different* crops receive *more* weight in the placebo aggregation. If our estimates of the indirect crop channel are driven by household i reacting to aggregate supply shifts induced by emigration from similar crop-growing regions, then emigration of households i' growing *different* crops should have no effect on household i's output.

Table B8 shows the placebo results with crop profits as the outcome. Column 1 replicates the baseline estimate (Table 5 column 6) for reference, and column 2 adds a control for the placebo. The placebo coefficient is near-zero and statistically insignificant, whereas the main crop market effect remains virtually unchanged. This supports our claim that the positive coefficient on M_{ijdt}^{crop} captures household i's genuine reaction to changes through crop markets.

Next, we compute a second placebo that accounts for substitutability between crops. Even if household i and i' grow different crops, their production may still be linked through prices if their crops are substitutes or complements. However, if this were the case, we would have failed the first placebo test. Nevertheless, we compute another placebo measure where $d(\mathbf{x}_i, \mathbf{x}_{i'})$ is the inverse cross-price elasticity between the main crop grown by household i and i' obtained from Anand et al. (2016)[16]. Taking the inverse means that households growing non-substitutes (i.e., unrelated crops) receive more weight in the aggregation. For this test, we measure the main crop market channel with elasticity weights as well (not inverted).

Column 3 of Table B8 controls for this placebo measure, and, once again, the coefficient is near-zero. The main crop market effect remains positive, significant, and similar in magnitude. This indicates that households do not only react to crop market effects driven by emigration from regions growing the same crops but also regions growing substitutes and complements.

5.3.1 Robustness and Heterogeneity of Indirect Effects

Crop Suitability Weights: Table B9 shows results when aggregate emigration from the crop-growing region, M_{ijdt}^{crop}, is measured in crop suitability space. The concern is that crop similarity weights Table 5 are endogenous if the crop mix of households i and j is jointly determined by an unobserved correlate of migration and production. We redefine the elements of vector x in

[16]Compensated elasticities for India are provided for four relevant categories of crops: cereals, pulses, veggies/fruits, and other. To construct the placebo, 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.

Equation (3.4) to be suitability indices for crop k at the district level (see Section 2 for data details). Reassuredly, the estimates are virtually identical.

Agricultural Regulation: Table B10 reports estimates when agricultural markets are allowed to span the whole of India. This allows household i to be affected by the emigration of households growing similar crops anywhere in India rather than only within the state. In Panel A, when crop similarity weights are between actual crop portfolios, coefficients on the crop channel remain positive for profits and farm size and turn negative when technology is the outcome. When the weights pertain to crop suitability (Panel B), coefficients on the crop channel remain positive across most specifications, but the magnitudes grow substantially. The sensitivity of these estimates is likely driven by the fact that agricultural markets in India function at the state level, not the national level, in which case this robustness test is misspecified.

Extensive Margin: Our main estimates describe intensive margin responses since the sample comprises households that owned land in both periods. Non-land-owning households have missing values for agricultural outcomes and are dropped. To study extensive margin responses, we zero-fill agricultural expenses and machinery for these households. Table B11 shows that our direct and indirect estimates remain remarkably stable. These findings, therefore, underscore that including the extensive margin of agricultural exit and entry does not change our main conclusions.

Heterogeneity of Direct and Indirect Effects: As with the direct effects (section 5.2.1), we estimate heterogeneity of the indirect effects by household remoteness and urban productivity. These are the dimensions by which households differ in the model. Table B12 splits the sample by above- and below-median distance to the nearest city. The sign and significance of coefficients among urban (Panel A) and remote (Panel B) households are similar to the main estimates. The magnitude of the direct and indirect effects is larger among urban households for the majority of outcomes.

Table B13 splits the sample by above- and below-median number of resident working age males at baseline, our measure of urban productivity. Once again, we remain underpowered in the below-median subsample since the median is two and the majority of below-median values are zero. In the above-median sample, however, directionality and significance mirror the main estimates, whereas coefficient magnitudes are larger.

6 Counterfactual Simulations: Aggregate Extent of Spatial Spillovers

The previous results quantify the redistribution of agricultural production across households. We now aggregate our estimates to study total changes in food supply following Moscona and Sastry (2022). We find that 84% of the migration-driven decline in food supply near urban areas is compensated by indirect market-driven spillovers. We use a hypothetical counterfactual since national trends are captured by year fixed effects in our regressions.

Methods: We define a set of counterfactuals that allow us to disentangle the direct and indirect channels. For each household i in year t, we use the coefficients from Equation (9) 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:

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

Next, we define a *Labor Only* (LO) scenario in which both indirect effects are fixed at t_1:

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

In the same way, we define a *Labor and Land* (LL) scenario where the indirect crop channel is fixed at t_1, as well as a *Labor and Crop* (LC) scenario where the indirect land effect is fixed at t_1. We then sum these household-level predicted values for crop production 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 spatial general equilibrium effects, with migration plus land market effects, and with migration plus crop market effects, 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) spatial general equilibrium effects relative to the total possible value absent migration:

$$PctChange^{GE} = 100 \cdot \left(\frac{TotY_{t_2} - TotY^{NM}_{t_2}}{TotY^{NM}_{t_2}} \right) \qquad PctChange^{PE} = 100 \cdot \left(\frac{TotY^{LO}_{t_2} - TotY^{NM}_{t_2}}{TotY^{NM}_{t_2}} \right) \tag{11}$$

The third is the amount of agricultural decline through the labor channel that is offset by indirect effects as a percentage of the counterfactual change absent spatial spillovers:

$$PctOffset = 100 \cdot \left(\frac{PctChange^{PE} - PctChange^{GE}}{PctChange^{PE}} \right) \tag{12}$$

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

Simulation Results: Figure 4A shows simulation estimates of the change in aggregate crop output with and without indirect market spillovers (Equation 11). To account for uncertainty in the underlying point estimates, we repeat the simulation procedure on 1000 bootstrapped samples and plot the mean (red) and confidence intervals (black error bars). Under the *Labor Only* scenario, with indirect channels shut off, aggregate migration would have caused a 66% reduction in agricultural output compared to the *No Migration* counterfactual. This amounts to Rs. 175 million worth of food. When all channels operate, the supply contraction becomes six times smaller.

Panel B shows how much of the migration-induced food shortages is mitigated by market spillovers (Equation 12). 43% of the shortage is mitigated through land markets and 41% through crop markets. Both indirect channels together mitigate 84% of the direct effect of migration. The recovered crops are worth Rs. 147 million, or 50% of the in-sample total crop value in 2012.

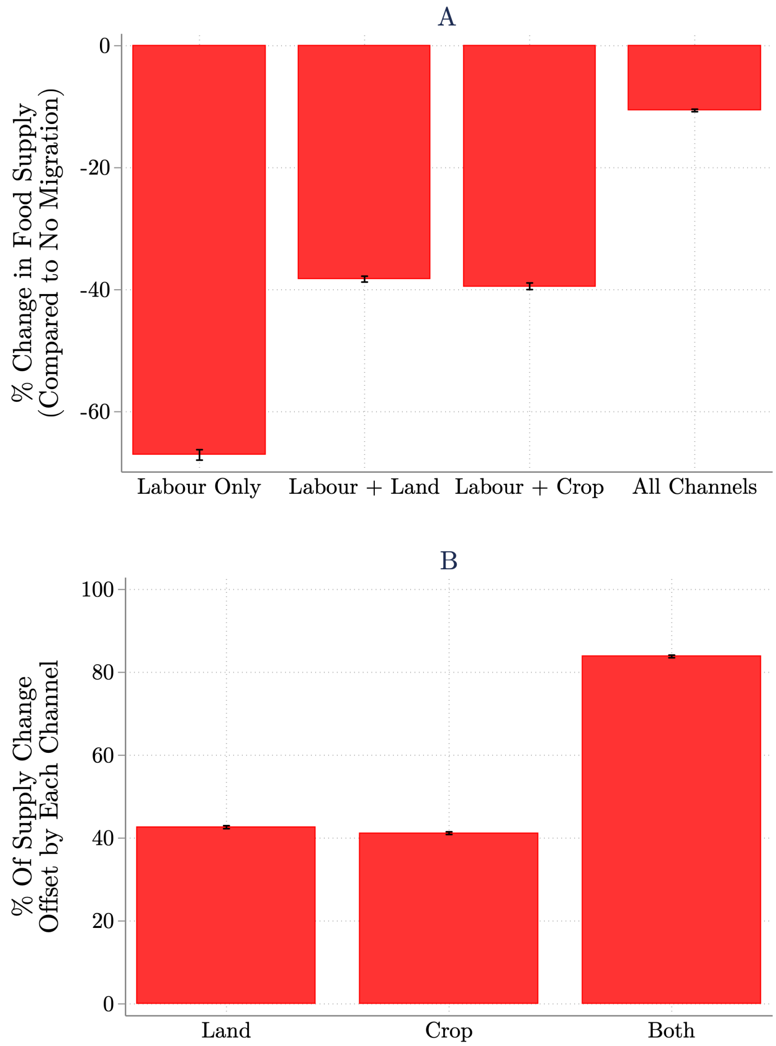


Figure 4: Aggregate Influence of General Equilibrium Effects

Note: Panel A displays the aggregate change in agricultural supply from migration under the four scenarios (Equation 11). *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 12). Confidence intervals computed from 1000 bootstrap draws.

7 Discussion and Conclusion

7.1 The Spatial Reorganization of Agriculture

The empirical estimates (Table 5) and simulation (Figure 4) have important implications for the spatial reorganization of agriculture in response to structural change. The direct effect predominantly impacts households sending many migrants, for whom large output declines are not offset by the indirect effects. In contrast, households sending zero migrants exclusively experience the indirect benefits in the form of greater technology adoption and production. Since remote house-

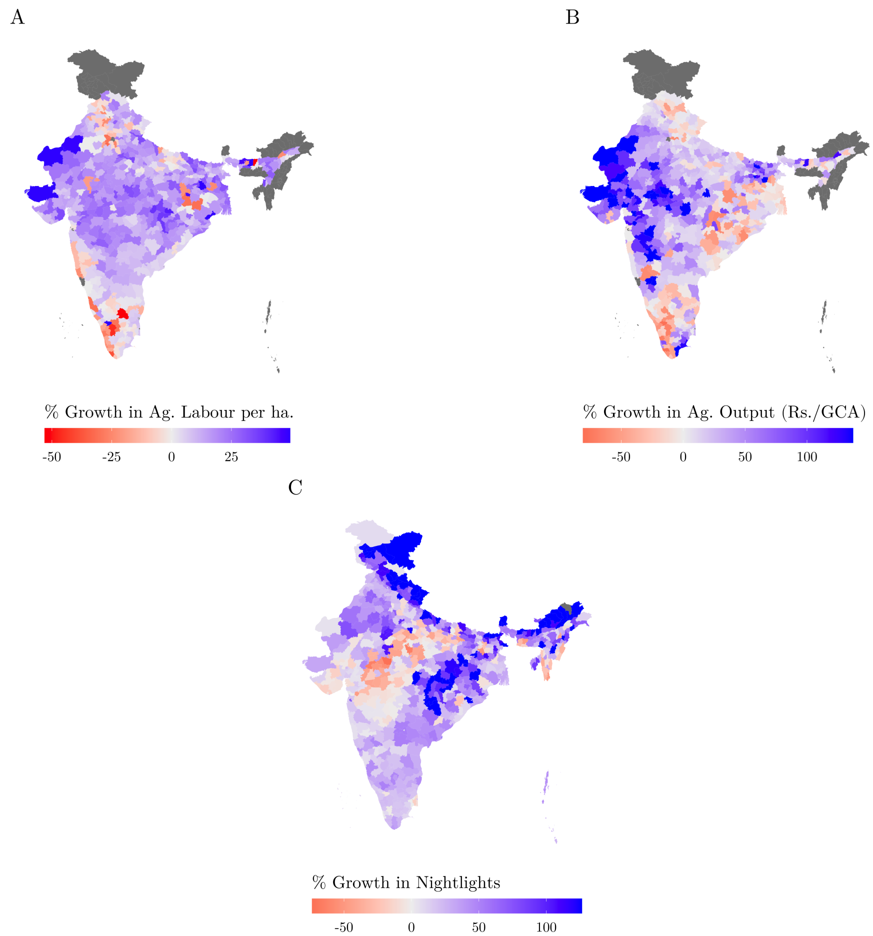


Figure 5: Labor Reallocation and Output Growth Across Districts (2001-2011)

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

holds send fewer migrants (Fact II, Section 2.3), we expect a spatial reorganization of agriculture away from urban areas (high-migration) and toward remote areas (low-migration).

We investigate this phenomenon using district data from the Census of India, the Agricultural Census, and nightlight intensity[17], a proxy for economic activity (Henderson et al., 2012). Figure 5 maps 2001-2011 growth rates of agricultural labor (Panel A), crop output (Panel B) and nightlight intensity (Panel C). Panels A and B show that Southern, Eastern, and Northern districts experiencing labor loss (Panel A, red) also experience output contraction (Panel B, red). A local analysis would thus imply that emigration leads to agricultural decline, as we saw in Section 5.2 (Table 4).

An aggregate view reveals compensating production increases in Central and Western India (Panel B, blue). Economic activity is a key driver of this spatial redistribution. Whereas labor exit

[17] Gridded data are obtained from the VIIRS satellite product at 15 arc-second resolution

and output contraction (Panels B, C; red) are concentrated in high-growth areas (Panel C, blue), output increases are concentrated in low-growth areas (Panel C, red). In general, agriculture declines in peri-urban, migrant-sending districts, but farmers in stagnant districts pick up the slack. This pattern is consistent with our spatial model, empirical estimates, and simulation results.

7.2 Conclusion

Reallocation of labor from farms to cities is an emblematic feature of economic development. While this process is well-studied, its effects on agriculture are relatively unexplored. This paper studies the impact of labor migration on agriculture taking spatial equilibrium into account. In doing so, we provide new evidence on structural change and the reorganization of agriculture.

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. Migration causes a contraction of agricultural technology, farm size, and crop output in direct response to labor loss.

This result does not mean that we should expect food scarcity as India urbanizes. In aggregate, declining household food production drives up crop prices. Similarly, increased land availability reduces land prices. Guided by a spatial general equilibrium model, we measure these two indirect channels and quantify partial and general equilibrium effects in a single empirical framework. We find that the direct labor effect and indirect market effects draw in opposing directions. Importantly, households with no migrants living in remote areas increase crop production and adopt technology. Documenting the spatial incidence of general equilibrium effects allows us to characterize the spatial reorganization of agriculture toward remote areas where emigration is low.

Our estimates show that spatial equilibrium effects mitigate 80% of the migration-induced food shortage between 2005-2012. The spatial redistribution of agriculture through market-based forces is, therefore, economically significant but not a panacea. These findings contrast previous studies–mainly from developed countries—on labor shocks and agricultural development. Many find the opposite result: that agriculture modernized more quickly in areas experiencing labor emigration (Hornbeck and Naidu, 2014; Manuelli and Seshadri, 2014). One reason for the discrepancy is that we study a developing country where food production differs structurally from developed countries. Another is that previous literature uses aggregate data at the county or district level. Our household level results suggest that non-migrant households in high-migration areas invest more in agriculture than non-migrant households in low-migration areas. This interpretation reconciles our findings with the literature.

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

References

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

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

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

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

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

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

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

Autor, David H, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American economic review*, 2013, *103* (6), 2121–2168.

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

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

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

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

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

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

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

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

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

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

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with arbitrary clustering," 2019.

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

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

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

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

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

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

Gao, Xuwen, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song, "Restrictions on Migration Create Gender Inequality: The Story of China's Left-Behind Children," 2022.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, August 2020, *110* (8), 2586–2624.

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

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

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

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

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

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

Matsuyama, Kiminori, "Agricultural productivity, comparative advantage, and economic growth," *Journal of economic theory*, 1992, *58* (2), 317–334.

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.

_ **and Karthik A Sastry**, "Does Directed Innovation Mitigate Climate Damage? Evidence from U.S. Agriculture," *The Quarterly Journal of Economics*, 10 2022, *138* (2), 637–701.

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.

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

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

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

Tombe, Trevor, "The missing food problem: Trade, agriculture, and international productivity differences," *American Economic Journal: Macroeconomics*, 2015, *7* (3), 226–58.

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

Xu, Chenzi, "Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures*," *The Quarterly Journal of Economics*, 04 2022. qjac016.

A Proofs

A.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

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} - \frac{w\varphi_{ij}}{\tau_j} = 0, \tag{13}$$

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} - \rho_j = 0 \quad \text{and} \tag{14}$$

$$p_k \omega_{jk} \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} - v = 0. \tag{15}$$

To simplify notation, we use subscripts in the following to denote partial derivatives.[18] 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

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

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 the response to the direct effect.

[18] 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.

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 (1) and (2) ($f_{\theta a} > 0, f_{la} > 0$).

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

$$\frac{dY_j}{d\tau} = \sum_i \sum_k \frac{dy_{ijk}}{d\tau} = \sum_i \sum_k p_k \omega_{jk} \left(\frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \frac{\partial l_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial a_{ijk}} \frac{\partial a_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \tau} \right) \tag{16}$$

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 (16), therefore, simplifies to

$$\frac{dY_j}{d\tau} = \sum_i \sum_k p_k \omega_{jk} \left(\frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial l_{ijk}} \frac{\partial l_{ijk}}{\partial \tau} + \frac{\partial f(l_{ijk}, a_{ijk}, \theta_{ijk})}{\partial \theta_{ijk}} \frac{\partial \theta_{ijk}}{\partial \tau} \right) \tag{17}$$

For labor complementary technologies, we get

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

A.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 A.1). First, we fully differentiate the first-order conditions with respect to the urban wage:

$$p_w \omega f_l + p\omega(f_{ll} l_w + f_{la} a_w + f_{l\theta}\theta_w) - \frac{\varphi}{\tau} = 0, \tag{18}$$

$$p_w \omega f_a + p\omega(f_{al} l_w + f_{aa} a_w + f_{a\theta}\theta_w) - \rho_w = 0, \tag{19}$$

$$p_w \omega f_\theta + p\omega(f_{\theta l} l_w + f_{\theta a} a_w + f_{\theta\theta}\theta_w) = 0. \tag{20}$$

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

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

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

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

A.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 A.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 A.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 (1). 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 (2).

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 3) such that $\rho_w < 0$. Assumption (2) 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 3) 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 (2)) while $f_\theta f_{aa} < 0$ by assumption (positive marginal productivity and concavity) which completes the proof for the last statement of Prediction 2.

A.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 A.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

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

B Appendix Tables

Table B1: Migrant Type by Age Cohort

Age Cohort	Student (%)	Employed Son (%)	Husband (%)
0-4	68.97	9.20	5.75
5-9	95.70	1.20	0.17
10-14	91.24	5.30	0.22
15-19	61.34	29.40	1.83
20-24	32.87	44.45	9.63
25-29	7.07	44.35	25.28
30-34	0.74	36.58	33.09
35-39	0.77	32.11	36.36
40-44	0.28	30.16	40.06
45-49	0.12	22.26	46.49
50-54	0.20	19.45	41.26
55-59	0.00	10.64	47.34
60+	0.00	1.80	18.28
Total	26.83	29.53	20.20

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

Table B2: Summary Statistics: Migrant Profiles

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

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

Table B3: 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 B4: Second Stage—Direct Effect of Migration on Technology Adoption

	Agricultural Expenses (Rs.)					Machinery (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants (M_{ijdt}^{labor})	-1.459***	-2.022***	-2.032***	-0.929**	-2.305***	-1.279***	-1.578***	-2.104***	-0.923	-1.262***
	(0.443)	(0.464)	(0.766)	(0.415)	(0.589)	(0.428)	(0.581)	(0.479)	(0.612)	(0.424)
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	5165.314	8303.526	4259.474	2289.573	3227.644	0.507	0.657	0.391	0.275	0.220
HH FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	25852	25454	24054	23496	24602	24692	23862	24042	23882	23746
F-Stat	58.7	56.9	53.1	52.0	56.3	54.4	55.4	53.9	53.1	52.9

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

Table B5: Direct Effect: Heterogeneity by Distance to Nearest City

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Panel A: Below Median						
Male Migrants (M_{ijdt}^{labor})	-3.635***	-2.558**	-1.947**	-0.152	-5.357***	-2.502***
	(1.358)	(1.094)	(0.796)	(0.102)	(1.672)	(0.903)
Outcome SD	-	-	4.442	96.350	346.054	30725.909
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	8046	7932	8946	6482	6482	7776
F-Stat	19.2	18.7	21.9	20.1	20.1	25.8
Panel B: Above Median						
Male Migrants (M_{ijdt}^{labor})	-2.414***	-1.914**	-3.824***	0.736	-4.843***	-3.468***
	(0.902)	(0.782)	(1.378)	(0.723)	(1.324)	(1.042)
Outcome SD	-	-	4.515	828.846	348.190	30101.007
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	11036	10442	12346	8848	8848	10520
F-Stat	20.9	17.2	16.6	17.7	17.7	23.6

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. In Panels A and B, median refers to distance from the household village to the nearest town. The explanatory variable is number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. All outcomes are in standard deviations. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (see Section 3.4 for index construction). Wage bill (column 4) is wages paid to all workers in the past year. Person-days (column 5) includes both household and hired labor. Crop profits (column 6) are net of expenses. All specifications control for origin income, the uninteracted shift, and district drought conditions. Standard errors clustered by PSU.

Table B6: Direct Effect: Heterogeneity by Urban Productivity

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Panel A: Below Median						
Male Migrants (M_{ijdt}^{labor})	-4.398	20.668	-15.994	1.183	4.962	5.579
	(18.659)	(588.239)	(44.408)	(4.273)	(12.777)	(21.233)
Outcome SD	-	-	3.730	826.748	273.260	25785.809
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	10754	10306	12514	8848	8848	11160
F-Stat	0.1	0.0	0.1	0.2	0.2	0.1
Panel B: Above Median						
Male Migrants (M_{ijdt}^{labor})	-6.658**	-4.881*	-6.642**	-0.405	-12.368**	-6.654***
	(3.164)	(2.568)	(2.905)	(0.462)	(6.029)	(2.540)
Outcome SD	-	-	4.770	132.550	378.231	33087.801
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	15174	14664	16234	12062	12062	13872
F-Stat	7.0	5.8	7.6	4.6	4.6	9.9

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. In Panels A and B, median refers to number of working age males living in the households at baseline. The explanatory variable is the number of working age male migrants in the household, instrumented with inverse-distance population weighted income interacted with the number of baseline working age males in the household. All outcomes are in standard deviations. Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively (see Section 3.4 for index construction). Wage bill (column 4) is wages paid to all workers in the past year. Person-days (column 5) includes both household and hired labor. Crop profits (column 6) are net of expenses. All specifications control for origin income, the uninteracted shift, and district drought conditions. Standard errors clustered by PSU.

Table B7: 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	-0.664***	-0.927***	-0.945***	-0.428**	-1.055***	-0.606***	-0.763***	-0.987***	-0.450	-0.597***
(direct labour channel)	(0.199)	(0.209)	(0.348)	(0.194)	(0.269)	(0.206)	(0.283)	(0.230)	(0.295)	(0.205)
Village emigration	0.155***	0.183***	0.185**	0.082**	0.217***	0.178***	0.164***	0.224***	0.111	0.120**
(indirect land channel)	(0.042)	(0.043)	(0.078)	(0.040)	(0.059)	(0.047)	(0.059)	(0.054)	(0.068)	(0.048)
Crop region emigration	0.191***	0.219***	0.319***	0.153***	0.199***	-0.030	0.120*	0.063	0.090*	0.040
(indirect crop channel)	(0.047)	(0.053)	(0.083)	(0.047)	(0.060)	(0.050)	(0.069)	(0.049)	(0.053)	(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	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	25600	25202	23798	23238	24340	24364	23544	23726	23564	23428
F-Stat on Direct Effect	63.8	61.2	56.2	54.2	59.6	55.3	55.1	54.1	52.6	52.1

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 similarity (see section 3.4 for measurement). Columns 1-5 are measured in Rs. spent in the past year. Water refers to purchases of irrigation water. Rentals include both hired equipment and animals. Column 6-10 are measured in quantities. Pumps include both electric and diesel pumps. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B8: Placebo Tests of Crop-Market Spillovers

	(1) Profits	(2) Profits	(3) Profits
Male Migrants	-1.441***	-1.440***	-1.435***
(direct labour channel)	(0.318)	(0.318)	(0.316)
Village emigration	0.279***	0.279***	0.276***
(indirect land channel)	(0.075)	(0.075)	(0.075)
Crop region emigration	0.246***	0.269***	
(indirect crop channel)	(0.061)	(0.077)	
Crop region emigration		-0.026	
(placebo: euclidean distance weights)		(0.054)	
Crop region emigration			0.216***
(elasticity weights)			(0.054)
Crop region emigration			0.071*
(placebo: inv. elasticity weights)			(0.039)
HH FEs	✓	✓	✓
Year FEs	✓	✓	✓
Observations	20150	20150	20150
F-Stat on Direct Effect	50.1	50.2	50.5

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. The outcome is crop profits. 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 replicates the main table 5 (column 6) where weights on *Crop region emigration* are inverse euclidean distance between crop portfolios of household i and j. Column 2 adds a placebo where weights on *Crop region emigration* are not inverted. In column 3, weights on *Crop region emigration* are cross-price elasticities between the main crop grown by household i and j and the placebo measure is the inverse elasticity. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B9: Robustness: Redefining Crop-Growing Region in Crop Suitability Space

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Male Migrants	-1.160***	-0.983***	-1.042***	0.119	-2.335***	-1.328***
(direct labour channel)	(0.260)	(0.248)	(0.254)	(0.127)	(0.408)	(0.255)
Village emigration	0.242***	0.242***	0.216***	-0.028	0.497***	0.259***
(indirect land channel)	(0.058)	(0.056)	(0.055)	(0.028)	(0.098)	(0.061)
Crop region	0.249***	0.088*	0.198***	-0.005	0.456***	0.178***
emigration (indirect crop channel)	(0.054)	(0.050)	(0.047)	(0.012)	(0.081)	(0.046)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25924	24966	28744	20906	20906	25030
F-Stat on Direct Effect	64.1	57.1	58.7	56.3	56.3	68.3

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 suitability indices of the India's major crops (see Section 3.4 for details). Columns 1 and 2 are indices for expenses and machinery, respectively. Wage bill (column 4) is wages paid to hired labor in the past year. Person-days (column 5) includes family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B10: Robustness: Direct and Indirect Effects (Nationwide Crop Markets)

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Male Migrants	-1.112***	-0.952***	-1.088***	0.139	-2.302***	-1.448***
(direct labour channel)	(0.259)	(0.242)	(0.279)	(0.140)	(0.416)	(0.319)
Village emigration	0.297***	0.258***	0.274***	-0.030	0.596***	0.334***
(indirect land channel)	(0.068)	(0.064)	(0.071)	(0.030)	(0.113)	(0.085)
Crop region	-0.460***	-0.612***	0.458***	-0.316	-0.074	0.198
emigration (indirect crop channel)	(0.162)	(0.139)	(0.134)	(0.246)	(0.264)	(0.206)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	25662	24632	26524	20644	20644	20150
F-Stat on Direct Effect	63.3	57.2	53.4	54.6	54.6	50.7

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 similarity (see Section 3.4 for details). Columns 1 and 2 are indices for agricultural expenses and stock of machinery, respectively. Wage bill (column 4) is wages paid to hired labour in the past year. Person-days (column 5) includes family and hired labor. Crop profits (column 6) are net of expenses. All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B11: Robustness: Direct and Indirect Effects (Extensive Margin)

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Person-days	(6) Crops
Male Migrants	-2.686***	-2.126***	-2.375***	0.156	-4.626***	-2.610***
(direct labour channel)	(0.679)	(0.542)	(0.585)	(0.231)	(0.952)	(0.639)
Village emigration	0.563***	0.505***	0.542***	-0.035	1.036***	0.562***
(indirect land channel)	(0.150)	(0.124)	(0.132)	(0.053)	(0.225)	(0.149)
Crop region	0.521***	0.197*	0.378***	0.014	0.853***	0.428***
emigration (indirect crop channel)	(0.126)	(0.102)	(0.098)	(0.024)	(0.160)	(0.099)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Observations	31130	29770	31776	25630	25630	24990
F-Stat on Direct Effect	40.5	35.1	33.0	32.3	32.3	31.0

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 similarity. Columns 1 and 2 are indices for agricultural expenses and machinery, respectively (see Section 3.4 for index construction). Wage bill (column 4) is in Rupees and describes total expenses on hired labor in past year. Column 5 is total person-days of family and hired labor in past year. Crop profits are net of expenses (Column 6). All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B12: Direct and Indirect Effects: Heterogeneity by Distance to Nearest City

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1) Expenses	(2) Machinery	(3) Cultivated	(4) Wage Bill	(5) Person-days	(6) Crops
Panel A: Below Median						
Male Migrants	-1.495***	-1.103**	-0.855**	-0.064	-2.276***	-1.118***
(direct labour channel)	(0.512)	(0.455)	(0.353)	(0.044)	(0.646)	(0.416)
Village emigration	0.369***	0.329***	0.253***	0.013	0.582***	0.277**
(indirect land channel)	(0.138)	(0.111)	(0.092)	(0.014)	(0.187)	(0.123)
Crop region emigration	0.345***	0.150	0.106	0.005	0.536***	0.095
(indirect crop channel)	(0.112)	(0.102)	(0.066)	(0.012)	(0.136)	(0.064)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
N	7966	7830	8306	6402	6402	6244
F-Stat on Direct Effect	25.0	23.2	25.1	24.1	24.1	25.7
Panel B: Above Median						
Male Migrants	-1.129***	-0.929**	-1.885**	0.358	-2.343***	-1.759***
(direct labour channel)	(0.425)	(0.390)	(0.744)	(0.352)	(0.700)	(0.624)
Village emigration	0.168**	0.162**	0.261**	-0.049	0.369***	0.247**
(indirect land channel)	(0.070)	(0.069)	(0.120)	(0.056)	(0.128)	(0.115)
Crop region emigration	0.311***	0.092	0.409***	-0.010	0.421***	0.312**
(indirect crop channel)	(0.101)	(0.081)	(0.152)	(0.031)	(0.145)	(0.121)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
N	10918	10266	11268	8732	8732	8456
F-Stat on Direct Effect	21.2	15.5	13.6	15.8	15.8	17.3

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All outcome variables are standardized. In Panels A and B, the median refers to distance from the household's village to the nearest town. *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 similarity. Columns 1 and 2 are indices for agricultural expenses and machinery, respectively (see Section 3.4 for index construction). Wage bill (column 4) is in Rupees and describes total expenses on hired labor in past year. Column 5 is total person-days of family and hired labor in past year. Crop profits are net of expenses (Column 6). All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

Table B13: Direct and Indirect Effects: Heterogeneity by Urban Productivity

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Panel A: Below Median						
Male Migrants	-3.625	1.839	-3.805	0.483	1.747	-3.663
(direct labour channel)	(28.958)	(12.949)	(4.713)	(1.622)	(4.038)	(10.766)
Village emigration	0.488	-0.245	0.540	-0.072	-0.234	0.533
(indirect land channel)	(3.871)	(1.832)	(0.657)	(0.228)	(0.564)	(1.667)
Crop region emigration	0.532	-0.262	0.513	-0.011	-0.089	0.406
(indirect crop channel)	(3.777)	(1.455)	(0.619)	(0.152)	(0.448)	(1.084)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
N	10582	10104	11200	8678	8678	8310
F-Stat on Direct Effect	0.0	0.0	0.6	0.2	0.2	0.1
Panel B: Above Median						
Male Migrants	-2.803**	-2.219*	-2.948**	-0.183	-5.618**	-3.292**
(direct labour channel)	(1.206)	(1.141)	(1.209)	(0.208)	(2.717)	(1.438)
Village emigration	0.724**	0.631**	0.791**	0.044	1.481**	0.828**
(indirect land channel)	(0.321)	(0.307)	(0.329)	(0.056)	(0.740)	(0.386)
Crop region emigration	0.682***	0.366	0.561**	0.032	1.261**	0.616**
(indirect crop channel)	(0.261)	(0.239)	(0.239)	(0.045)	(0.530)	(0.262)
HH FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
N	15080	14528	15324	11966	11966	11840
F-Stat on Direct Effect	8.9	6.3	9.1	4.9	4.9	7.4

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All outcome variables are standardized. In Panels A and B, median refers to the number of resident baseline working age males at baseline. *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 similarity. Columns 1 and 2 are indices for agricultural expenses and machinery, respectively (see Section 3.4 for index construction). Wage bill (column 4) is in Rupees and describes total expenses on hired labor in past year. Column 5 is total person-days of family and hired labor in past year. Crop profits are net of expenses (Column 6). All columns control for origin income, the uninteracted shift, and drought conditions. Standard errors clustered by PSU.

C Appendix Figures

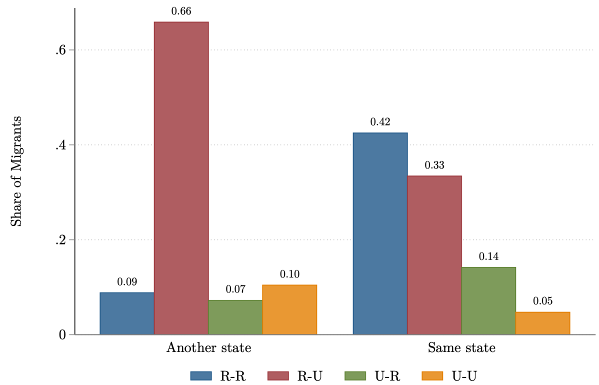


Figure C1: Migration Streams by Distance

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.

D Robustness of Direct Effects

The following paragraphs explore threats to our partial equilibrium estimates. Robustness tests are reported in Table B14 and sensitivity of statistical inference in Table B15.

Controlling for Observables. Panel A of Table B14 includes state-year fixed effects instead of year fixed effects. This accounts 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. Panel B includes the extensive margin by zero-filling outcome variables for non-land-owning households. This includes non-land-owning 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. Panel C controls for household education since educated households are disproportionately exposed to the distance-weighted income shock (Table 2), and, they may potentially be on a different outcome path in the absence of the shock.

Another threat to validity is the existence of trends in the pre-period shock. This is a concern if the trend in agricultural outcomes across households in areas with high baseline shock severity differs from that of households in areas with low baseline shock severity, regardless of whether they have resident working age males. We account for this in Panel D by including the baseline shock interacted with year fixed effects. The results remain stable.

Alternative Shifts and Shares. Next, we show that our estimates are robust to alternative shift-share instruments. First, we explore alternative specifications of the "shift". Panel E restricts to rural-urban migration only. Results remain stable, suggesting that our results are not driven by any particular stream but rather reflect internal migration in general. In Panel F, we use inverse-distance weighted nightlights as the "shift"[19]. Although nightlights units are measured in arc-seconds, coefficient magnitudes can be compared with that of the other shift-share specifications since outcomes are measured in standard deviations. Again, results remain stable, suggesting that our estimates are robust to different measures of the shift.

Second, we explore alternative measures of the "share". In Panel G, the share is the number of working age household members of *both genders*. Emigration (the explanatory variable) also covers both genders. Results are similar, which is unsurprising since women make up less than 10% of migrants (Table B2).

Inference. Table B15 reports estimates that are more robust to heteroskedasticity induced by spatial correlation across shifts and shares. As a baseline, we first compute robust standard errors. Since the shock is at the district level, we also report standard errors clustered by district to account for correlation between shocks across households in the same district. We also report errors clustered by state in case shocks are correlated across districts within the same state. Lastly, we in-

[19]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.

vestigate 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[20]. Our estimates remain stable and precise regardless of the inference method, even when allowing spatial correlation across long distances.

[20]We adapt the procedure implemented in Hsiang (2010) to an instrumental variables setting using the method developed by Colella et al. (2019).

Table B14: Robustness Checks

	Technology Index		Land (ac.)	Labour		Profits (Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)
	Expenses	Machinery	Cultivated	Wage Bill	Person-days	Crops
Panel A: State-Year FEs						
Male Migrants	-2.043***	-1.771***	-2.243***	0.240	-4.624***	-3.056***
	(0.495)	(0.473)	(0.531)	(0.271)	(0.780)	(0.587)
Observations	25928	24970	28748	20910	20910	25032
Panel B: Extensive Margin						
Male Migrants	-3.133***	-2.355***	-2.346***	0.100	-5.071***	-2.759***
	(0.651)	(0.452)	(0.456)	(0.183)	(0.712)	(0.492)
Observations	69756	68698	74706	63970	63970	65174
Panel C: Control for Education						
Male Migrants	-2.789***	-2.508***	-2.405***	0.231	-5.401***	-3.205***
	(0.723)	(0.694)	(0.715)	(0.269)	(1.089)	(0.693)
Observations	25928	24970	28748	20910	20910	25032
Panel D: Pre-period Trends						
Male Migrants	-2.629***	-2.082***	-2.424***	0.277	-5.096***	-3.079***
	(0.623)	(0.554)	(0.598)	(0.272)	(0.920)	(0.594)
Observations	25928	24970	28748	20910	20910	25032
Panel E: Rural-Urban Migration						
Male Migrants	-2.567***	-2.178***	-2.404***	0.206	-4.998***	-3.013***
	(0.608)	(0.576)	(0.584)	(0.273)	(0.925)	(0.581)
Observations	25148	24288	27596	20256	20256	24066
Panel F: Nightlights Shift-Share						
Male Migrants	-2.822***	-2.123***	-2.655***	0.143	-5.501***	-3.237***
	(0.705)	(0.609)	(0.657)	(0.247)	(1.061)	(0.609)
Observations	25928	24970	28748	20910	20910	25032
Panel G: All Gender Migration						
All Migrants	-4.031***	-3.665***	-4.182***	0.145	-7.547***	-4.417***
	(1.175)	(1.100)	(1.171)	(0.254)	(1.731)	(0.975)
Observations	25928	24970	28748	20910	20910	25032

Note: * $p < .1$, ** $p < .05$, *** $p < .01$. All outcome variables are standardized. The explanatory variable is the number of working age male migrants in the household. See tables notes of Table 4 for details on the migration instrument and other specification details). Panel A includes state-year FEs. The remaining panels only include year FEs. Panel B includes non-land-owning households in both periods. Panel C includes a control for whether the household head is educated (beyond 10th grade). Panel D controls for baseline trends in the distance-weighted income shock. Panel E restricts to rural-urban migration only. Panel F uses distance-weighted nightlights as the shift. In Panel G, the explanatory variable and the IV "share" includes both males and females. All specifications control for district drought conditions. Standard errors are clustered at the PSU level.

Table B15: Robustness: Alternative Standard Error Clustering

	Machinery	Ag. Expenses	Cultivated (ac.)	Wage Bill	Labour	Profits
Coefficient on Migration	-2.020776	-1.768078	-2.211659	.2361479	-4.641081	-3.037238
SE: Robust	.3749111	.3422357	.372162	.2542086	.6564817	.4574608
SE: District	.526225	.4904272	.4998415	.2587692	.7764696	.6278926
SE: State	.6544304	.5141994	.6556117	.2613159	.9997372	.754747
SE: Conley (200km radius)	.4857467	.367435	.4008383	.1920407	.765337	.4467658
SE: Conley (500km radius)	.3727189	.3529197	.4772027	.1809112	.4453041	.5889666

Note: The first row replicates the main coefficient estimates from Table 4. Conley standard errors are estimated in a 2SLS setting using the Colella et al. (2019) implementation.