

Madhok, Raahil; Noack, Frederik; Mobarak, Ahmed Mushfiq; Deschênes, Olivier

**Working Paper**

## Rural-urban migration and the re-organization of agriculture

Economic Growth Center Discussion Paper, No. 1095

**Provided in Cooperation with:**

Yale University, Economic Growth Center (EGC)

*Suggested Citation:* Madhok, Raahil; Noack, Frederik; Mobarak, Ahmed Mushfiq; Deschênes, Olivier (2022) : Rural-urban migration and the re-organization of agriculture, Economic Growth Center Discussion Paper, No. 1095, Yale University, Economic Growth Center, New Haven, CT

This Version is available at:

<https://hdl.handle.net/10419/265575>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Yale University

## EliScholar – A Digital Platform for Scholarly Publishing at Yale

---

Discussion Papers

Economic Growth Center

---

Fall 10-5-2022

### Rural-Urban Migration and The Re-organization of Agriculture

Raahil Madhok

Frederik Noack

Ahmed Mushfiq Mobarak

Olivier Deschenes

Follow this and additional works at: <https://elischolar.library.yale.edu/egcenter-discussion-paper-series>



Part of the [Agricultural and Resource Economics Commons](#), [Food Studies Commons](#), [Growth and Development Commons](#), [Labor Economics Commons](#), and the [Regional Economics Commons](#)

---

# Rural-Urban Migration and the Re-organization of Agriculture

Raahil Madhok\*

University of British Columbia

Frederik Noack<sup>†</sup>

University of British Columbia

Ahmed Mushfiq Mobarak<sup>‡</sup>

Yale University

Olivier Deschenes<sup>§</sup>

University of California Santa Barbara

October 5, 2022

## Abstract

This paper studies the response of agricultural production to rural labor loss during the process of urbanization. Using household microdata from India and exogenous variation in migration induced by urban income shocks interacted with distance to cities, we document sharp declines in crop production among migrant-sending households residing near cities. Households with migration opportunities do not substitute agricultural labour with capital, nor do they adopt new agricultural machinery. Instead, they divest from agriculture altogether and cultivate less land. We use a two-sector general equilibrium model with crop and land markets to trace the ensuing spatial reorganization of agriculture. Other non-migrant village residents expand farming (land market channel) and farmers in more remote villages with fewer migration opportunities adopt yield-enhancing technologies and produce more crops (crop market channel). Counterfactual simulations show that over half of the aggregate food production losses driven by urbanization is mitigated by these spillovers. This leads to a spatial reorganization in which food production moves away from urban areas and towards remote areas with low emigration.

---

\*madhokr@mail.ubc.ca

<sup>†</sup>frederik.noack@ubc.ca

<sup>‡</sup>ahmed.mobarak@yale.edu

<sup>§</sup>olivier@econ.ucsb.edu

# 1 Introduction

Movement of labor from rural to urban areas is perhaps the most pervasive feature of structural transformation. America's rise in the 19th and 20th centuries was powered by dramatic worker migrations from farms to cities to take up manufacturing and service jobs (Alvarez-Cuadrado and Poschke, 2011). China's economic boom in the late 20th century similarly featured rural workers moving *en masse* to populate urban factories. 147 million migrants were counted in China's 2005 census (11% of the population) (Gao et al., 2022). India's urban transition is currently underway (Figures 1 and 2). As migration fuels urbanization around the world, how the left-behind rural agricultural sector will adjust to labour loss remains somewhat unclear.

The goal of this paper is to deepen our understanding of how structural change shapes agricultural development. How is the food produced after workers leave farms for cities? Do those farms adopt new capital-intensive production processes to counteract labour shortages? If not, then where does crop production move? Answering these questions is crucial for understanding the economic re-organization that accompanies structural change and economic development.

We answer these questions with detailed panel data from thousands of Indian households during a period of rapid urbanization in the early 21st century. Hornbeck and Naidu (2014) and Manuelli and Seshadri (2014) also study these questions, but with more aggregated data from the historical United States. They document a process of agricultural modernization, whereby counties facing emigration shocks adopted new labor-replacing technologies like tractors. Technology adoption may be a direct response to labour loss or an indirect response to changing sectoral compositions across the broader economy. Our micro data allows us to paint that richer picture of the post-migration adjustment process.

A contribution of our study is that we characterize spatial spillovers across regions instead of assuming localized shocks. Rural regions near growing urban centers lose more labor than places further away from urban income opportunities. Previous literature suggests that labour-losing areas modernize quickly and become new centers of agricultural development. This defies traditional views of factor endowments and comparative advantage, which state that regions abundant in agricultural labor specialize in crops and become agricultural hubs. We reconcile these views by documenting both the direct effects of emigration on technology adoption as well as the indirect effects in areas less exposed to of urbanization. As such, we are able to document the internal spatial reallocation of agricultural activities following urbanization.

Our first finding is that rural households with working-aged males living near high-growth cities send migrants but, surprisingly given the prior literature, they do *not* replace the labour with capital. Nor do they hire replacement workers. Instead, they divest from agriculture entirely and cultivate less land. Using a two sector partial equilibrium model, we show that these findings are consistent with the response of a profit-maximizing farm to an urban wage shock when technology is labor complementary. Our results are relevant for developing countries where agricultural investments such as seeds and fertilizer are labor complementary rather than labor saving.

If farming households with new migration opportunities shift out of agriculture, then do other

households shift in and start cultivating more land? To trace the spatial reorganization of agriculture, we extend our model to general equilibrium and incorporate spatial heterogeneity as well as endogenous input and output prices. When some households in a village send migrants and divest from agriculture, land markets allow other villagers to expand agriculture. Crop markets allow households outside the village to react to the output contraction and grow more of those same crops.

Consistent with these model predictions, our second empirical result is that partial and general equilibrium forces draw in opposing directions and along distinct spatial lines. Food production declines compared to the national average near urban areas where emigration is high, but features compensating increases in remote areas, where emigration is low. Unlike families with migrants, non-migrant households adopt yield-enhancing technologies such as seeds and agrochemicals. This pattern characterizes the spatial reorganization of Indian agriculture in response to urban service sector growth and the accompanying outflow of labour from rural areas.

Several stylized facts about economic growth, sectoral labour reallocation, and rural-urban migration in contemporary India motivate our empirical identification strategy. India was the world's fastest growing economy between 2013 and 2018, fueled by a booming service sector. This was accompanied by a sharp decline in the agricultural share of employment. Working-aged males previously working in agriculture were attracted by new service sector opportunities, and emigrated more towards higher-growth cities.

These facts motivate a shift-share research design for identifying the impact of household emigration ([Goldsmith-Pinkham et al., 2020](#)). This design helps overcome the endogeneity of migration, since households with and without migrants differ in terms of wealth, education, or the general economic and physical environment. In shift-share verbiage, the "shift" is urban incomes weighted by proximity to the city. The "share" is each household's migration potential, defined as the number of resident prime-aged males. The interaction of shift and share yields an instrument that disentangles the stream of migration induced by urban income shocks from other determinants of migration at the origin. Related literature employs regression discontinuity designs in the context of transport infrastructure, where identification is driven by differences in access to new roads ([Asher and Novosad, 2020](#); [Asher et al., 2020](#)).

Destination income shocks could also affect household production through aggregate demand for agricultural goods, which would violate the exclusion restriction ([Borusyak et al., 2022b](#)). We therefore control for destination income shocks directly. Identification thus relies on heterogeneous exposure of households to the spatially-weighted income shocks. Although fixed differences between more- and less-exposed households are absorbed by household fixed effects, it is still possible that agricultural outcomes of more-exposed households correlate with the destination income shocks even in the absence of migration. We address this concern by showing that urban income shocks are quasi-randomly distributed conditional on controls and fixed effects ([Borusyak et al., 2022a](#)). In other words, urban income shocks are uncorrelated with the exposure variable and therefore also with potential trends in the outcomes that are correlated with the exposure variable as well. This rules out the possibility that the shock picks up differential time paths

of more- and less-exposed households.

We apply this shift-share design to the Indian Human Development Survey (IHDS), a panel survey of 42,000 households conducted in 2005 and 2012. Our analysis yields three key findings about factor reallocation in response to internal migration. First, we find sharp *declines* in technology adoption at the household level. In line with our partial equilibrium model, this is largely driven by reduced investment in labour complementary technologies such as agrochemicals, irrigation water, and work animals. In contrast, we find no effect for labour-saving technologies. Second, declines in technology adoption lead to a reduction in agricultural output and a contraction of farm size. Third, we find no evidence of increases in hired labour to replace lost family labour. This corroborates the notion that labour markets are malfunctioning in rural India ([Foster and Rosenzweig, 2022](#)). Overall, these three findings suggest that migrant-sending households divest from agriculture.

A naive interpretation of these results is that the urban sector is modernizing while the rural sector is not. Such an extrapolation is unfounded without a deeper understanding of general equilibrium effects. Inspection of a district-level map of India reveals a new stylized fact: while crop production declines in labour-losing districts, it increases elsewhere. Nightlight maps further reveal that production shifts to remote areas less exposed to emigration. Emigration is uncommon in remote areas due to high migration costs, suggesting that production increases on these farms arise indirectly. We conjecture that crop and land prices are the main indirect channel. For example, output contraction in peri-urban districts increases crop prices which incentivizes remote households to increase agricultural production.

Our general equilibrium model incorporates the demand side of the economy, village land markets, crop prices, variety, and suitability in order to generate spatial price links between households growing similar crops. This setup allows labour reallocation to affect production directly through labour loss as well as indirectly through crop and land prices. To characterize space, the model yields a migration cost threshold beyond which households do not send migrants and instead specialize in agriculture. This is the key to isolating the spatial incidence of direct and indirect effects in the empirics.

The model yields a parsimonious expression that disentangles the effect of urban labour shocks on crop production into its partial and general equilibrium components. The net effect depends on which forces dominate. Linking model parameters to variables recorded in the IHDS, we extend our empirical strategy to estimate the incidence of each force. The direct channel is measured as before: by number of household male migrants instrumented with the shift-share variable. The indirect crop channel is measured as a spatially weighted average of household migration, excluding the index household, with weights equal to household crop similarity ([Adao et al., 2019](#)). This captures the fact that households react more to changes in aggregate supply and prices of a crop if they also grow that crop, regardless of where they live. Finally, the indirect land channel is measured as the average number of migrants across other households *in the same village*, excluding the index household. This captures the fact that land prices are crop-independent yet “local”, encompassing the village.

The main result is that partial and general equilibrium effects of structural change draw in opposite directions. Whereas crop production continues to decline through the labour reallocation channel, this effect is tempered by the indirect land and crop channels. The implied magnitudes are large; increased production through the land-market channel offsets migration-induced production declines by 19%. The mechanism is through farm expansion, in line with the theory of falling land prices. Similarly, production increases through the crop-market channel offset the direct decline by 13%. We find that those households invest more in yield-enhancing technology such as seeds, agrochemicals, irrigation water, and rental equipment. We find little evidence of machinery adoption.

The results also spotlight the spatial incidence of direct and indirect effects. The (negative) direct effect dominates for households sending many migrants, who live near cities. In contrast, the (positive) indirect effects dominate for households sending zero migrants, who live farther away. Even though remote households do not “participate” in structural transformation directly, they nevertheless experience production and technology benefits triggered by indirect price effects.

Our findings therefore imply that structural transformation does not necessarily drive a shift away from agriculture altogether, but rather a spatial and inter-household redistribution toward more remote areas. Stated another way, they imply that labour reallocation *is* important for agricultural development, but that the important effects are not directly driven by labour loss and adoption of labour-saving technology. Instead, we show that agricultural development materializes indirectly in the form of increased crop production, farm expansion, and technology adoption among remote non-migrant households.

The main remaining question is whether, in aggregate, remote households fully pick up the slack left by peri-urban households that reduce food production. The answer is important for understanding whether India has a “missing food problem” (Tombe, 2015). The last part of the paper studies how much of the aggregate food decline is mitigated by general equilibrium effects. We define special cases of the empirical specification that allow aggregation to counterfactual scenarios. These include total crop value with no migration, with migration but no spillovers, with migration and land spillovers but no crop spillovers, and so on. We find that the spatial reorganization of agriculture mitigated 61% of the aggregate food decline. Quantitatively, the food “saved” by domestic reallocation amounts to Rs. 147 million, or, 48% of total crop value.

**Literature Contributions.** This paper contributes to at least four bodies of work. First, it is related to studies on rural development in India (Aggarwal, 2018; Asher and Novosad, 2020). Second, we extend existing research on labour reallocation and technology adoption in partial equilibrium. Rozelle et al. (1999) use cross-sectional survey data from China to show that out-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 adoption of high-yielding seed varieties (HYV).

Our third contribution is to decompose the agricultural response to structural transformation into partial equilibrium (labour loss) and general equilibrium (price) channels. In doing so, we

advance the literature on structural transformation that uses data aggregated at higher spatial scales. [Caprettini and Voth \(2020\)](#) show with parish-level data that the introduction of threshing-machines in England “released” agricultural labour and led to structural transformation (and riots). Similarly, [Bustos et al. \(2016\)](#) use municipality-level data to show that HYV adoption in Brazil was labour-saving and led to significant industrial growth. Both of these studies consider the reverse of our research question.

Like this paper, [Hornbeck and Naidu \(2014\)](#) and [Manuelli and Seshadri \(2014\)](#) both also study migration and agricultural development. [Hornbeck and Naidu \(2014\)](#) use county-level data to show that Black out-migration in the 1920s US South led to agricultural modernization over several decades. [Manuelli and Seshadri \(2014\)](#) use calibration to show that contractions in the US labor force was a primary driver of tractor adoption in the United States. Both papers use aggregated data that do not try to distinguish between farm-level labour effects and aggregate general equilibrium effects. We contribute to this literature by decomposing these two mechanisms.

[Emerick \(2018\)](#) studies the converse of our question and decomposes partial and general equilibrium effects. Using district-yearly rainfall shocks to instrument crop productivity, they find that positive productivity shocks increase the *non-agricultural* labour share. The explanation is that rising district incomes increase demand for non-agricultural goods. Whereas [Emerick \(2018\)](#) distinguish partial and general equilibrium as cross-sectoral differences, we do so within the same sector. And if there are structural transformation processes that materialize over longer time scales, then our seven-year study period would capture it.

Lastly, this paper joins a new literature on structural transformation and the spatial organization of agriculture. Whereas few papers separate the direct and indirect effects of structural change, even fewer document the spatial incidence of these effects as we do here. Three exceptions are [Blakeslee et al. \(2022\)](#), [Asher et al. \(2021\)](#), and [Pellegrina \(2022\)](#). [Blakeslee et al. \(2022\)](#) and [Asher et al. \(2021\)](#) study the reverse of our setting and document sectoral effects of an agricultural productivity shock in India depending on distance to the shock. Their specifications allow effects to decay with distance. [Pellegrina \(2022\)](#) use a structural model to illustrate the spatial effects of agricultural productivity shocks on the Brazilian economy. Our approach extends these papers by identifying the spatial incidence of direct and indirect effects of structural change in a single empirical framework.

The rest of this paper is organized as follows. The next section describes a series of stylized facts about structural transformation in India that motivates our model in Section 3. Section 4 describes the household panel data and provides summary statistics. Section 5 presents our instrumental variable strategy and Section 6 shows the main results. Section 7 extends the model and empirical framework to general equilibrium. Section 8 concludes.

## 2 Background

This section documents facts about India’s economic context relevant to structural change and agricultural development. Liberalization policies in the 1990s ushered in an era of rapid economic



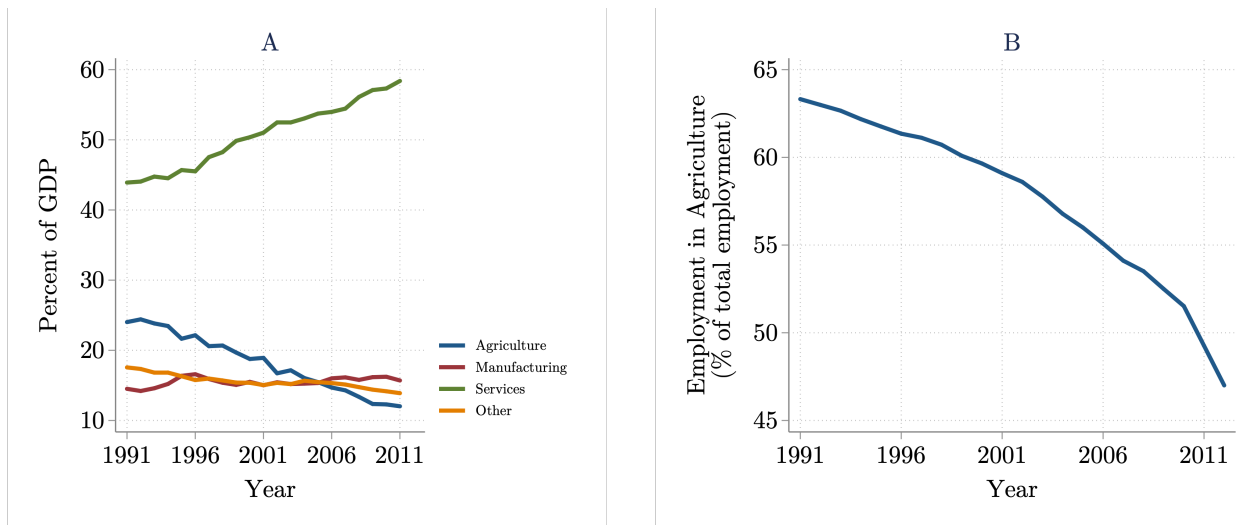


Figure 1: Economic Growth and Sectoral Re-allocation of Labour (1991-2012)

Note: Panel A reports sectoral shares of India's GDP at factor cost in 2005 prices. Data accessed from the Planning Commission of India Databook. Panel B reports percent of total employment in agriculture. Data from ILOSTAT database and obtained through World Bank Open Data Portal.

growth. Cities were the engine of growth, powered by a booming service sector. Workers left agriculture and migrated to higher-growth destinations. Despite this structural shift, agricultural output did not decline uniformly. We establish these facts using national administrative data. From an empirical standpoint, the fact that workers sort on destination growth allows us to isolate exogenous pull factors from endogenous push factors driving migration. Our instrument exploits spatial income shocks to generate plausibly random variation in migration.

## 2.1 Motivating Facts

**Fact 1: Cities were the engine of post-liberalization growth.** Landmark reforms were enacted in 1991 to liberalize India's economy and stimulate the service sector, largely concentrated in cities. GDP per capita nearly tripled in the next two decades (see Figure A1). Between 2013-2018, India overtook China as the world's fastest-growing major economy (International Monetary Fund, 2015).

Figure 1A shows sectoral shares of GDP in the post-reform period. The service share grew steadily after liberalization, making up half of GDP by 2000 and 60% by 2012. At the same time, the urban population share grew from 25% in 1991 to nearly 32% by 2012 (Figure A3). Over half of India's population is expected to be urban by 2050 (UN Population Division, 2018).

**Fact 2: Urbanization is fueled by labour reallocation.** Internal labour migration accompanied the modernization of India's economy. Agricultural productivity declined as the service sector boomed (Figure 1A). Aggregate employment in agriculture also dropped nearly 20% during the post-reform era (Figure 1B), suggesting workers exploited the arbitrage opportunity. At a local

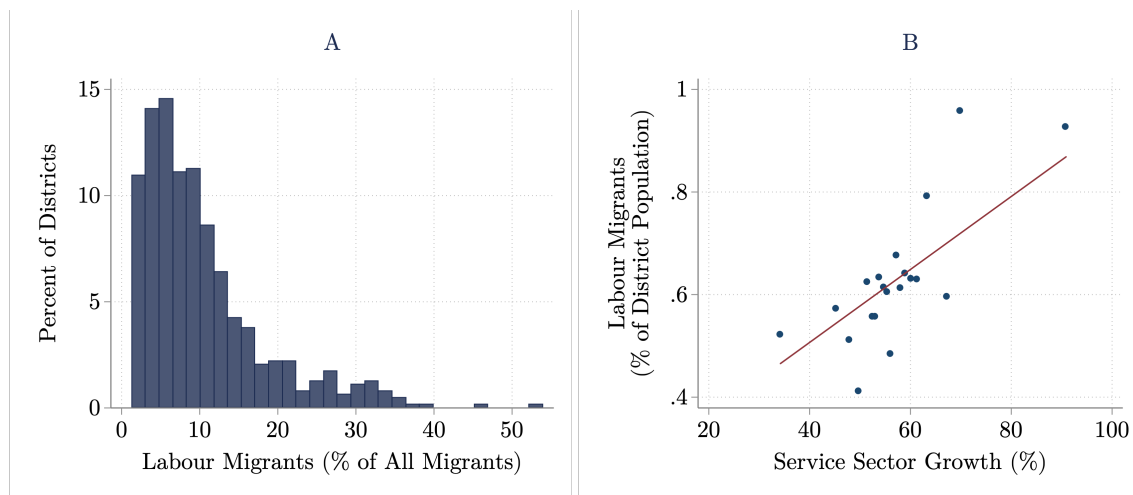


Figure 2: Labour Reallocation Across Districts

Note: Panel A shows the distribution of labour migrants in a district having arrived between 2001-2011 as a proportion of the total district migrant flow during the same period. Panel B shows a binscatter of district-level service sector GDP growth between 2004-2011 (in 2004 prices) and in-migration rates, residualized on state fixed effects. GDP data are for 310 districts available in the ICRISAT database. The y-axis is the share of labour migrants in the district population who arrived between 2006-2011. Migration data (in both panels) is accessed from the 2011 Census Migration Tables. The red line in Panel B is the best linear fit, constructed from an OLS regression of the y-residuals on the x-residuals.

level, Figure 2A shows the distribution of cross-district labour migrants among the local migrant population who arrived in the post-reform era. There is substantial variation in labour migration during the period of agricultural decline, in line with the aggregate evidence.

Despite rural-urban migration, it should be noted that labour remains misallocated in agriculture in India (Munshi and Rosenzweig, 2016). Overall, however, stylized evidence of a thriving service sector alongside rural out-migration and declining agricultural employment highlights the role of sectoral labour reallocation in fostering India’s urban-oriented economic development.

**Fact 3: High-growth districts attract more migrants.** The variation in labour migration across districts (see Fact 2) during India’s economic liberalization can be explained by several competing factors (Kone et al., 2018). The leading explanation is that migrants sort across districts according to destination productivity. A simple test is whether high-growth districts—such as those experiencing service sector booms—attract more migrants.

Figure 2B shows a binscatter of service sector GDP growth between 2004-2011 (in 2004 prices) against the population share of labour migrants who arrived during the same period using district GDP data from ICRISAT and migration data from the 2011 Census. The blue dots represent mean in-migration rates for equally-sized bins of service sector growth after adjusting for state fixed effects to account for state-level unobservables affecting all districts. The steep upward trend implies that migrants sort into higher-growth destinations. This unveils a “labour pull” force underlying structural change in India. We exploit this force to identify the impacts of labour

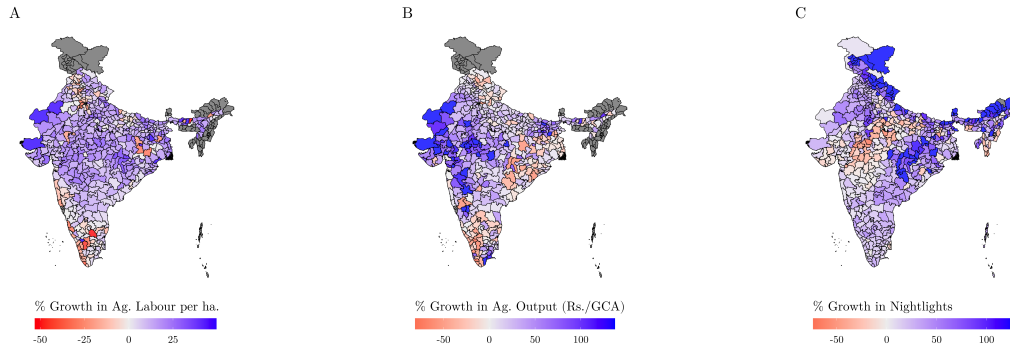


Figure 3: Labour Reallocation and Output Growth Across Districts (2001-2011)

Note: The time period for each map is 2001-2011. Panel A maps the percent change in number of agricultural labourers per hectare of cultivated area across districts. Data are from the 2001 and 2011 Census. Panel B maps the percent change in crop output across districts. Data are 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 9th percentile. Panel C maps the percent change in nightlight intensity from the VIIRS satellite product.

reallocation in the empirical analysis below.

**Fact 4: Agricultural Output Did Not Decrease Everywhere.** The flip side of migration-fuelled urban growth is labour loss from agricultural areas. The implications of urbanization for food production are ex-ante unclear. Figure 3 maps district-level growth rates over 2001-2011 for agricultural labour in Panel A, crop production in Panel B, and nightlight intensity (a proxy for economic growth (Henderson et al., 2012)) in Panel C. Panels A and B show a positive correlation between labour exit and crop production within districts. Southern, Eastern, and Northern districts experiencing labour loss (Panel A, red) also experience output contraction (Panel B, red). A simple, ‘local’ analysis would thus imply that migration leads to food scarcity.

However, an aggregate view reveals notable compensating production increases in Central and Western India (Panel B, blue). Urban growth appears to be a key driver of this spatial redistribution. Whereas labour exit and output contraction (Panels B, C; red) are concentrated in high-growth areas (Panel C, blue), production increases are concentrated in low-growth areas (Panel C, red). Put differently, food production declines in peri-urban, migrant-sending districts, but farmers in stagnant districts “pick up the slack”. We turn to a formal exploration of these phenomena in Section 7.

## 2.2 Summary of Facts

These stylized facts underscore key features of structural change in India: a thriving service sector (Fact 1), labour outflows from agriculture (Fact 2), and more migration towards higher-growth destinations (Fact 3). These features are associated with a spatial reorganization of agriculture (Fact 4). As crop production fell in labour-losing districts, agriculture shifted elsewhere. This opens important questions about which direct and indirect forces determine the spatial shift. A

deeper understanding of these processes also reveals who benefits most from structural change.

We proceed in two parts. The next part (Sections 3-6) theoretically analyzes household responses to labour reallocation in partial equilibrium. We then test model predictions empirically by packaging Facts 1-3 into an instrument for migration. The second part (Section 7) extends the model and empirics to characterize the spatial organization of agricultural in general equilibrium.

### 3 Model

This section develops a two-sector partial equilibrium model of household production. Workers face mobility costs across locations but are freely mobile across sectors. The purpose of the model is to elicit theoretically-grounded predictions about technology responses to labour reallocation. Factor prices are fixed to enable labour-capital substitution as the only margin of adjustment. The model is parsimonious, integrates the stylized facts from Section 2, and yields testable predictions. The general equilibrium extension with flexible prices is left for Section 7.

#### 3.1 Setup

**Environment.** There is one urban centre and a large number rural regions. Each rural region is inhabited by a large number of farmers. The farmers allocate time between agriculture and manufacturing, which earns urban wage  $w$ . Rural labour markets are absent, given malfunctioning labour markets in India (Foster and Rosenzweig, 2022). Appendix A.1.1 elaborates on this assumption. Land markets are also absent, given the thinness of land markets in India and other developing economies (Fernando, 2022). Functioning land markets are allowed in the model extension in Section 7. Product and capital markets are frictionless such that their prices are constant across regions.

Households supply one unit of labour inelastically. Working in manufacturing involves a migration cost,  $\tau$ , which increases with distance. Assuming iceberg costs (a fraction of the wage is lost), farmers earn net wage  $\frac{w}{\tau}$  in the city. We assume  $\tau \geq 1$ , with  $\tau = 1$  describing zero-cost migration since the farmer earns the full wage.

**Profit Maximization.** Agricultural production  $f(l, a, \theta)$  is concave and increasing in labor,  $l$ , land,  $a$ , and technology,  $\theta$ . We depart from the standard exposition and distinguish between two types of technologies. Technology is labor saving if it reduces the marginal productivity of labor and labor complementary if it increases the marginal productivity of labor (Acemoglu, 2010).

**Definition 1.** A technology is labor saving if  $f_{l\theta} < 0$  and it is labor complementary if  $f_{l\theta} > 0$ .

With this set up, the profit-maximizing farmer solves:

$$\max_{\theta, l} \quad pf(l, a, \theta) - v\theta + \frac{w}{\tau}(1 - l) \tag{1}$$

taking agricultural prices,  $p$ , technology prices,  $v$ , and wages,  $w$ , as given. Note that migrant wages  $w$  are assumed to be entirely remitted. This assumption has no bearing on profit-maximization since  $\frac{1}{\tau}$  can be re-interpreted as the share of wages available to the household decision maker.

This simple setup theoretically replicates key features of our setting. Since  $w$  is exogenous, we can solve Equation (1) to trace the impact of urban wage shocks on labour reallocation and technology adoption.

### 3.2 Testable Predictions

**Labour Reallocation.** In the following, subscripts denote partial derivatives. Totally differentiating the first order conditions (FOCs) of Equation (1) with respect to  $w$  and solving for  $l_w$  yields an expression for the response of agricultural labour to urban wage shocks:

$$l_w = \frac{1}{p\tau} \left( \frac{f_{\theta\theta}}{f_{ll}f_{\theta\theta} - f_{l\theta}^2} \right) \quad (2)$$

**Lemma 1.** *Increasing urban wages causes labor reallocation from agriculture to urban production. The impact of increased urban wages on agricultural labor declines with distance to the urban center.*

*Proof.* The proof follows directly from the sufficiency conditions for the farmers optimization problem i.e.  $f_{ll}f_{\theta\theta} - f_{l\theta}^2 > 0$  and  $f_{\theta\theta} < 0$  (full proof in Appendix A.1.2).  $\square$

This result directly informs our empirical strategy. Lemma 1 implies that positive urban wage shocks constitute a credible instrument for migration. Migration induced by these shocks (pull migration) is orthogonal to that induced by drivers of migration at the origin (push migration).  $\tau$  controls the strength of the urban pull, exerting a stronger force as  $\tau$  declines. We build this into the structure of the instrument with inverse distance weights.

**Technology Adoption.** A similar strategy characterizes the impact of urban wage shocks on technology adoption. Solving the totally differentiated FOCs of Equation (1) for  $\theta_w$  yields:

$$\theta_w = -l_w \frac{f_{\theta l}}{f_{\theta\theta}} \quad (3a)$$

$$= \frac{1}{p\tau} \left( \frac{-f_{\theta l}}{f_{ll}f_{\theta\theta} - f_{l\theta}^2} \right) \quad (3b)$$

**Proposition 1.** *Positive urban productivity shocks reduce (increase) agricultural technology use if the technology is labor complementary (saving). The impact of urban income shocks on agricultural technology use declines with the distance to the urban center.*

*Proof.* The first part follows from Definition 1 and the sufficiency conditions for the farmer optimization problem, i.e.  $f_{ll}f_{\theta\theta} - f_{l\theta}^2 > 0$ . The second part follows from the fact that  $\tau \geq 1$   $\square$

This result also directly informs our empirical strategy.  $\theta_w$  in Equation (3a) denotes the reduced form impact of urban wage shocks on rural agricultural technology adoption. Since prices are fixed, the impact operates entirely through a first stage labour effect as workers migrate to cities ( $-l_w$ ), scaled by ensuing changes to the marginal product of capital.

**Model Summary.** The model shows that, under a fixed price regime, households respond to positive urban wage shocks by reallocating labor from agriculture to urban production. The profit-maximizing response to this labour loss is to adopt more labour saving technology and dis-invest in labour complementary technology. This is especially true for urban and peri-urban households, whereas remote households experience less reallocation because of prohibitive migration costs.

## 4 Data and Measurement

Testing the model predictions requires three key data: i) labour migration, ii) agricultural development, and iii) disaggregated productivity shocks to construct the instrument. The IHDS panel survey is the only source of micro-data in India with detailed data on all three. This section outlines the data, variable construction, and provides summary statistics.

### 4.1 IHDS Survey

Data on individual migration are typically sparse and incomplete in India. For example, the census reports district-level gross immigration, but without information on origins to districts flows. The National Sample Survey (NSS) reports household emigration but does not define a migrant nor a reference period.<sup>1</sup> Further, few data sources combine information on migration and agricultural production. To this end, we use the two-wave IHDS household panel, a nationally representative survey, which provides rich information on all three types of data our analysis requires. Wave I (2004-05) surveyed 41,554 households, of which 83% were re-located in Wave II (2011-12). These households span 65% of districts across all states and territories<sup>2</sup>.

There are at least three advantages of using IHDS. First, it is one of the few Indian surveys documenting out-migration in detail. Second, the same households are interviewed in each round, enabling us to include household fixed effects in our analysis to control for time-invariant unobserved heterogeneity across households, such as caste or baseline poverty. Lastly, IHDS provides high quality disaggregated income measures. Survey staff calculate income from over 50 sources and compile them into eight categories (e.g. crop income, business income, etc.) according to a standardized procedure.

However, IHDS is not without limitations. First, exact migrant destinations are not documented, preventing us from exploiting productivity shocks from linked origin-destination pairs. We instead define a flexible instrument that allows *all* districts to exert a pull force, albeit less so

---

<sup>1</sup>The NSS 64th Round asks: “Did any former member of the household migrate out at any time in the past?”.

<sup>2</sup>With the exception of Andaman and Nicobar Islands and Lakshadweep, which contain < 1% of the population.

the further away it is (see section 5.2). Second, there is a 17% attrition rate. While attrition is common in most household panels, it poses minimal concern here. 80% of attritors did not own land in Wave I and would have been excluded from our sample. Third, seasonal migration (<6 months) is not reported in Wave II, restricting our analysis to medium term migration. Short-term migration is less likely to impact capital investment decisions of agricultural households.

## 4.2 Measurement

**Migration.** Migrants are household members living elsewhere for over 6 months of the past year. We arrive at this definition starting with the non-resident roster, which characterizes all household non-residents. IHDS defines a household as those sharing the same roof for at least 6 months of the past year. Thus, migrants are non-resident family members who *do not* meet this criteria.

Our definition is better-suited to characterize longer-term migration spells<sup>3</sup>. A migrant could be the household head's oldest son who has been living and working in another city for the past five years. It could also be his youngest son who worked there as a taxi driver for the past eight months and then returned. If the younger son instead worked for five months and then returned, he is a household resident and not a migrant. The non-resident roster reports migrant age, sex, and destination.

We apply two restrictions to the migrant sample. First, we restrict to internal migrants, who constitute over 90% of India's migrants<sup>4</sup>. Second, we keep working age males aged 15-60. Males make up 85% of migrants in this age group. To justify the working age window, Table A1 tabulates migrant types by age cohort. The student migrant share sharply drops after age 14, at which time the share of employed sons jumps nearly five times. This suggests that migrant males transition from school to work around age 15. The upper bound coincides with India's retirement age.

**Agricultural Outcomes.** Farmers report information on capital, labour, and crop revenue. The capital expense sheet includes: seeds, fertilizer, pesticides, irrigation water, and hired animals or equipment. The capital stock sheet includes: tubewells, electric/diesel pumps, bullock carts, tractors, and threshers. The labour expense sheet includes person-days of hired labour, wages paid, and person-days of unpaid family labour in the past year. We express all expenses in 2005 prices. The rural and urban price deflator is based on the CPI for agricultural and manufacturing labour, respectively.

We also collapse capital investment and stock into two indices to allay concerns of multiple hypothesis testing. This also simplifies exposition. We follow Anderson (2008) by first demeaning component outcomes and then dividing by the standard deviation of non-migrant households. The index is a weighted sum of demeaned values with weights equal to the row sum of the inverse covariance matrix. Intuitively, each variable influences the index in proportion to the information it adds. Several similar studies have implemented this method, e.g., (Asher and Novosad, 2020).

---

<sup>3</sup>Short-term (<6 months) circular migration is only documented in Wave II.

<sup>4</sup>359 international migrants (8.1% of migrants) were dropped in wave 1 and 881 (6.7%) were dropped in wave 2.

Table 1: Migrant Profile

	IHDS-1 (2004-05)			IHDS-II (2011-12)		
	Obs.	Mean	SD	Obs.	Mean	SD
<i>A: Migrant Type</i>						
Husband/Son	4402	0.51	0.50	13364	0.49	0.50
Wife/Daughter	4402	0.06	0.24	13364	0.05	0.23
Student	4402	0.26	0.44	13367	0.27	0.44
Other	4402	0.16	0.37	13364	0.19	0.39
<i>B: Destination</i>						
Within State	4393	0.64	0.48	13345	0.65	0.48
Other State in India	4393	0.36	0.48	13345	0.35	0.48
<i>C: Migration Stream</i>						
Rural-Rural	4377	0.25	0.43	13276	0.32	0.47
Rural-Urban	4377	0.47	0.50	13276	0.44	0.50
Urban-Rural	4377	0.12	0.33	13276	0.11	0.32
Urban-Urban	4377	0.11	0.31	13276	0.05	0.22

Note: The sample consists only of migrants. "Obs." is the total number of migrants. In Panel A, categories are mutually exclusive and describe a relation to an existing household member. In Panel C, "Rural-Rural" indicates the origin and destination were rural, "Rural-Urban" indicates the origin is rural and destination is urban, and so on.

We use the farm income variable constructed by IHDS, define as crop income minus expenses, where income is the product of crop quantity and prices. Over half of farmers report the price they *would* receive if sold at market, since over half of farm-crop pairs are not sold. We acknowledge that a more reliable farm income measure would use market-level crop prices. However, crop-wise data are unreleased for Wave II. We therefore use estimated farm income (deflated to 2005 prices), which is available for both periods<sup>5</sup>.

**Other Variables.** As we explain in details below, our instrument for migration is the inverse-distance population weighted district income. District income is mean household income reported in IHDS at the district level. Distances are computed between each district's centroid based on 2001 boundaries. Population weights are obtained from the 2001 Census. Section 5.2 provides details of instrument construction.

### 4.3 Summary Statistics

Our sample frame consists of 40,018 households interviewed in both periods. We restrict our sample to households that own land in both periods, which reduces the sample to 19,203 households.

**Migration Trends and Profiles.** Household migration more than doubled between survey waves. Nine percent of households ( $N = 3747$ ) had a migrant in Wave I, jumping to 23% ( $N = 9112$ ) in Wave II. Table 1 Panel A characterizes the typical migrant in each period. Categories are mutually exclusive. Husbands or sons of household residents are the most common migrant type, account-

<sup>5</sup>Five percent of households report negative farm income, which we recode as missing



Table 2: Capital and Labour Summary Statistics

	Obs.	Mean	SD
<i>Land</i>			
Area cultivated (ac.)	30091	3.50	4.49
Yield (Rs./ac.)	24062	8312.47	29823.97
<i>Expenses/Acre</i>			
Inputs	26097	2009.31	7380.80
Water	25884	259.89	1337.56
Rented Equipment	26467	638.17	2396.33
<i>Ownership (Num./Acre)</i>			
Pumps	25674	0.14	0.94
Tractors	25676	0.01	0.14
Bullock Carts	25762	0.06	0.27
<i>Labour (past yr.)</i>			
Farm Labour/Acre	24474	250.42	679.86
Wage Bill (Rs.)	26909	2802.66	8672.75

Note: Data is a household level panel for land-owning households in both periods. Observations vary due to missing data. Holding (ac.) is acres owned plus acres rented in minus acres rented out. Inputs include seeds, fertilizer, and pesticides. Water is purchased irrigation water. Rentals include machinery and work animals. Pumps are electric and diesel water pumps. Labour is measured in person-days. Farm labour is both hired labour and family labour.

ing for half of all migrants in both periods. Just 6% of migrants are wives or daughters, providing additional justification for restricting the analysis to males. The bulk of remaining migrants leave for education. Our working age window of 15-60 excludes the majority of student migrants (see Table A1) and keeps focus on labour migration.

Panels B and C help characterize the spatial dimension of migration in India. The majority of migration is within the state. Interestingly, rural-rural migration is an important migration stream, accounting for 25% and 32% of migrants in Waves I and II, respectively.

Figure A2 splits streams by inter- and intra-state travel, revealing a novel fact: among inter-state migrants, rural-urban migrants dominate rural-rural by seven to one, whereas rural-rural migration is more common among intra-state migrants. This is why we focus on internal migration instead of only rural-urban; because 64-65% of migration is within-state, where rural-rural migration is more prevalent than rural-urban.

**Agricultural Development.** Table 2 describes agricultural outcomes among land-owning households. Indian farming is generally small-scale, with the typical household cultivating 3.5 acres. Cultivation of yields farm income of approximately 8000 Rs. per acre each year on average.

Capital expenses and ownership are expressed in per-acre terms to account for scale differences. Inputs (agrochemicals and seeds) are the largest expense category at 2009 Rs. per acre, which accounts for 20% of farm income on average. Water pumps and bullock carts are the dominant technologies in our sample, while tractors are rarely owned.

## 5 Empirical Framework

This section describes our empirical framework for estimating the direct effect of labour migration on agricultural development. We first propose a shift-share instrument for internal migration based on Facts 1-3 (Section 2.1) and Lemma 1. We then test instrument validity through a variety of validation exercises. The 2SLS framework in this section assumes fixed prices. We extend the model and estimation strategy to allow endogenous crop and factor prices in Section 7.

### 5.1 Baseline Specification

To proceed, consider a simple linear relationship between labour outflows and agricultural development that can be written as:

$$Y_{iot} = \beta_1 Migrants_{iot} + \alpha_i + \gamma_{st} + \mu_{iot} \quad (4)$$

where  $Y_{iot}$  are agricultural outcomes for household  $i$  in origin district  $o$  at time  $t$ .  $Migrants_{iot}$  is the number of working age males who emigrated from the household.  $\alpha_i$  and  $\gamma_{st}$  are household and state-year fixed effects, respectively. Household fixed effects absorb time constant differences between households such as wealth levels and location, which could affect both migration and agricultural outcomes. The state year fixed effects control for local shocks such as agricultural policies, which are largely determined at the state level. To obtain baseline results, we estimate this equation with OLS and cluster errors by population sampling unit (PSU) (Abadie et al., 2017).

The coefficient of interest is  $\beta_1$  and its sign depends on part on whether technology is labour saving or complementary. It also depends on the curvature of the production function with respect to technology (see equation 3a). Based on the model we expect a negative slope for labor complementary technologies and a positive slope otherwise.

The regression in Equation 4 is likely to be confounded by omitted variable bias and reverse causality. Local economic or environmental shocks such as droughts may simultaneously determine migration as well as agricultural outcomes. Technology adoption itself may also release surplus labour, introducing simultaneity. The inclusion of household and state-year fixed effects only partially solve these problems if the unobserved confounders at the household level vary sufficiently over time and the environmental shocks have sufficiently high local spatial variation.

### 5.2 A Shift-Share Instrumental Variables Approach

We overcome these OLS identification concerns by constructing a shift-share instrument for migration informed by the stylized facts presented earlier (section 2.1) and the summary statistics. We denote the instrument as  $z_{iot}$ , for migration outflows from household  $i$  in origin district  $o$  at time  $t$ . The instrument combines income shocks,  $s_{odt}$ , at each destination  $d \in \Theta$  (where  $\Theta$  is the set of all destination districts) with the migration potential of the origin household at baseline,  $\lambda_{io}$ :

$$z_{iot} = \sum_{d \in \Theta / \{o\}} s_{odt} \cdot \lambda_{io} \quad (5)$$

**The Shift.** Income shocks form the “shift” of our shift-share design. The shock is constructed at the district level by first averaging incomes across all households in a district. For each origin district, we then aggregate over destination shocks weighting by the inverse distance between origin-destination centroids <sup>6</sup>:

$$s_{odt} = \sum_{d \in \Theta / \{o\}} \frac{inc_{dt} \times pop_d}{distance_{od}} = s_{ot} \quad (6)$$

$inc_{dt}$  is mean per-capita income in destination  $d \neq o$  and  $distance_{od}$  is the distance between the origin and destination. Distance weighting builds in a gravity structure where potential migrants are attracted more to nearby destinations for a given income. It also assumes that  $\Theta$  spans all districts, not only urban ones, which allows rural-rural migration in the underlying structure (see Table 1).  $s_{ot}$  is the leave-one-out distance-weighted income shock experienced at the origin.

We interact the shift with baseline district population,  $pop_d$  from the 2001 Census. Otherwise, rural and urban destinations are treated as equally attractive, which is incorrect since rural-urban migration is twice as common as rural-rural (Table 1). The interaction of distance and population weights ensures that potential migrants are drawn more towards urban destinations, although less so the further away they are. A time-constant population weight also improves instrument exogeneity by ensuring it captures income shifts and not population growth.

**The Share.** The number of working age males living at home at baseline is the “share” of our shift-share design. Combining the shift with this share is necessary because otherwise all district households are equally exposed to the destination shocks, including those with no potential migrants. In this sense,  $\lambda_{io}$  reflects the migration potential of each household, as their “potential” for exploiting destination income shocks by sending migrants grows with the number of working age males at home. We experiment with alternative exposure measures in the robustness checks.

### 5.3 Two-Stage Least Squares

Equipped with the shift-share instrument, we specify the effect of labour reallocation on agricultural outcomes in a standard two-stage least squares regression:

$$Migrants_{iot} = \mu_1 z_{iot} + \mu_2 s_{ot} + \mu_3 inc_{ot} + \alpha_i + \gamma_{st} + \varepsilon_{iot} \quad (7)$$

$$Y_{iot} = \beta_1 \widehat{Migrants}_{iot} + \beta_2 s_{ot} + \beta_3 inc_{ot} + \alpha_i + \gamma_{st} + \eta_{iot} \quad (8)$$

---

<sup>6</sup>Distances are between district centroids from the official 2001 District Census shapefile. All distances are measured in kilometres using the WGS84 projection.

Equation (7) is the first stage relating labour outflows from the origin,  $Migrants_{i ot}$  to the instrument,  $z_{i ot}$ , while controlling for the independent effect of the income shock ( $s_{ot}$ ). We also control for origin district per-capita incomes ( $inc_{ot}$ ), which accounts for spatial correlation of the destination income shocks and origin income.  $\alpha_i$  and  $\gamma_{st}$  are household and state-year fixed effects, respectively. To the extent that  $z_{i ot}$  is exogenous (see next section), predicted migration,  $\widehat{Migrants}_{i ot}$ , generated from this equation represents the “labour pull” component of worker reallocation that is orthogonal to migration incentives at the origin.  $\mu_1$  therefore represents the empirical analogue of  $l_w$  from the model (Equation 2). Empirical evidence on  $\mu_1 > 0$  would thus validate Lemma (1).

Equation (8) models the impact of labour reallocation generated by  $z_{i ot}$  on agricultural development at the origin, represented by  $Y_{i ot}$ . Proposition (1) states that the sign of  $\beta_1$ , the main coefficient of interest, depends on whether technology is labour complementary or labour saving. We estimate the first and second stage together using a two-stage least squares (2SLS) estimator.

Regardless of technology type, Equation (3b) from the theory implies that 2SLS will only yield a consistent estimate of  $\beta_1$  if migration is the only channel through which  $z_{i ot}$  affects  $Y_{i ot}$ . We explore instrument validity in the next section.

## 5.4 Instrument Validity

The instrument  $z_{i ot}$  combines a common shock with individual exposure to the shock, in line with the idea of shift-share instruments. A 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. Our design differs from the standard one in that we have two measures of exposure: distance to all potential destinations and migration potential. We discuss the validity of our instrument in the context of the shift-share literature below.

A potential concern with our instrument is that it implicitly treats labour migration as a unilateral choice, whereas it is bilateral in practice (Borusyak et al., 2022b). If origin incomes increase when destination incomes increase, then migrants have little incentive to move. This attenuates  $\mu_1 \rightarrow 0$  in the first stage even if the true migration elasticity with respect to local economic conditions is high (Borusyak et al., 2022b). Controlling for average incomes in the origin district,  $inc_{ot}$ , addresses this by ensuring migration decisions are based on destination shocks *relative to* the origin shock.

### 5.4.1 Endogenous Shares

The location of the household relative to urban centers is possibly endogenous. Households living near cities have different incentives and opportunities than their more remote counterparts. The time constant differences between locations are absorbed by the household fixed effects. However, agricultural outcomes may also develop differently than agriculture in remote areas, and this development may be correlated with the distance weighted income shocks (see e.g., Goldsmith-Pinkham et al. (2020)).

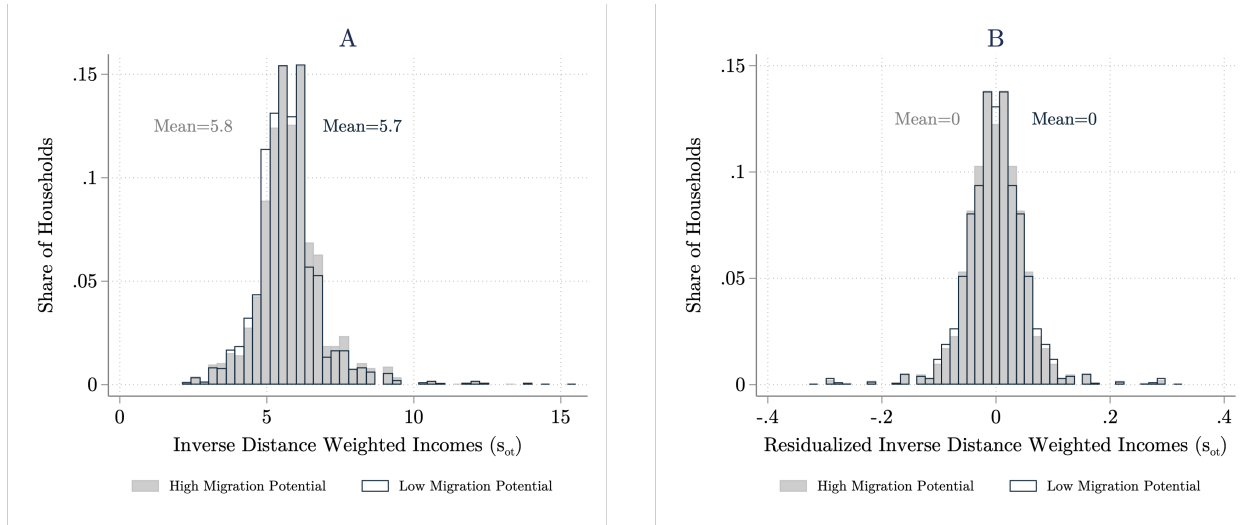


Figure 4: Distribution of Shocks

Note: Figures show distribution of inverse-distance weighted income shocks ( $s_{ot}$ ) 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 state-year fixed effects.

To help mitigate this potential bias we include  $s_{ot}$ , the distance-weighted income shocks, directly as a control. Identification thus relies on differences in household exposure to the urban income shocks, conditional on the shock itself.

The potential endogeneity of migration potential,  $\lambda_{io}$ , is more difficult to address. Households with more working-age males may differ from those with no males in many ways. The time-invariant nature of these differences are absorbed by household fixed effects. The identification concern, then, is that agricultural outcomes of households with high  $\lambda_{io}$  are correlated with the distance-weighted income shocks,  $s_{ot}$ , even in the absence of migration. We address this concern with visual evidence and a series of falsification tests.

#### 5.4.2 Exogenous Shifts

Given that some of the shares may be endogenous, the validity of our research 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 condition holds, we can rule out the possibility that  $s_{ot}$  picks up differential time paths of households with high and low migration potential in the baseline period.

**Distribution of Shocks.** We characterize the distribution of  $s_{ot}$  across households with high (above-median) and low (below-median) migration potential. This informs whether either type of household differentially experiences the shock, in which case the shock is non-random (Borusyak et al., 2022a). Panel A of Figure (4) shows that the distribution of  $s_{ot}$  is similar across both groups of households. The similarity is more evident after residualizing household and state-year fixed

Table 3: Shock Balance Tests

	(1)	(2)	(3)	(4)	(5)	(6)
	Males	Educated	HH Size	Farm Size	Ag. HH	Landowner
$\Delta$ Wt. Income	0.071 (0.090)	0.174* (0.103)	-0.450* (0.267)	-0.449 (0.582)	0.035 (0.045)	-0.074 (0.060)
$\Delta$ Origin Income	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	✓	✓	✓	✓	✓	✓
N	38589	38589	38589	18092	38589	38588
$R^2$	0.021	0.019	0.061	0.121	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 the number of household members. Column 4 is land cultivated (acres). Column 5 is an indicator for primary income coming from agriculture. Column 6 is a dummy for land ownership. Standard errors clustered at the PSU level.

effects (Panel B). The majority of households in both groups are exposed to similar levels of the shock, which helps rule out the possibility that the shock picks up correlated characteristics of more and less exposed households.

**Falsification Tests.** Next, we formalize the visual evidence for the as-good-as-random assignment of the shock. While Figure (4) provides supporting evidence, it is still possible that  $z_{iot}$  is correlated with other, unobserved household characteristics that affect outcomes through other channels. For example, if districts more exposed to the income shock have a large share of land-owning households at baseline, then these households may be subject to different agricultural development dynamics even in the absence of migration.

We follow [Borusyak et al. \(2022a\)](#) and [Xu \(2022\)](#) and rule out such violations of the exclusion restriction through a set of falsification tests. We choose a set of confounders at baseline, which are potential proxies for  $\eta_{iot}$  in Equation (8), and regress them on the change in the shock between 2005 and 2012,  $\Delta s_o$ , in a pooled cross-section:

$$y_{io} = \beta_1 \Delta s_o + \beta_2 \Delta inc_o + \gamma_s + \varepsilon_{io} \quad (9)$$

$\Delta inc_o$  is the change in the origin district's income, which accounts for demand effects (see Section 5.4.1), and  $\gamma_s$  is a state fixed effect.  $\beta_1$  is the balance coefficient. If  $\Delta s_o$  is as-good-as-randomly assigned to households within states, we expect it not to predict the chosen potential confounders. In that case, we should fail to reject the null hypothesis that  $\beta_1$  equals zero.

Table (3) shows balance tests for: number of household working age males (the migration potential), number of educated household members, household size, farm size, a dummy for earning income primarily from agriculture, and a dummy for owning land. The first column presents the key result: it shows that the shock is uncorrelated with the exposure variable. The other columns show the correlation of the shock with potential confounders. Overall, we find little correlation between these variables and the shock. Indeed we fail to reject the null hypothesis that  $\beta_1$  equals

Table 4: First Stage: Distance-Weighted Productivity Shocks and Internal Migration

	(1)	(2)	(3)
Wt. Income $\times$ Working Age Males	0.074*** (0.009)		0.065*** (0.009)
Wt. Income $\times$ Num. Educated		0.068*** (0.013)	0.042*** (0.013)
Wt. Income	Yes	Yes	Yes
Origin Income	Yes	Yes	Yes
Outcome Mean	0.201	0.201	0.201
HH FEs	✓	✓	✓
State $\times$ Year FEs	✓	✓	✓
KP (2006) F-Stat	67.74	27.59	35.99
Observations	25032	25032	25032

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . The outcome is number of working age male migrants in the household. “Wt. Income” is inverse-distance, population weighted income at the district level ( $s_{ot}$  in Equation 6). In column 1, this is interacted with number of resident working age males at baseline. In column 2, it is interacted with number of educated household members. In column 3, both instruments are included together. All regressions include controls for the direct shock and mean origin district income. Standard errors clustered by PSU.

zero at the 5% significance level in all six specifications. There is weaker evidence that households more exposed to the shock are larger and more educated. We report stress tests controlling for these two variables in the robustness checks below.

## 6 Results: Labour Reallocation, Technology Adoption, and Output

This section quantifies the direct impact of labour outflows on agricultural investment and output. In contrast to the developed world experience, Indian households do not replace labour with capital. Instead, we find that they decrease agricultural investment and food production. We explore indirect impacts of labour reallocation in Section 7. This section focuses on IV estimates (Equations 7 and 8) because of potential biases of the OLS estimates. OLS results are in Table A2.

### 6.1 First Stage: Destination Income Shocks Trigger Labour Reallocation

Although we provided evidence to support the credibility of the instrument in Section 5.4), we also need to demonstrate instrument relevance; that the shift-share variable must predict migration in the first place. Table 4 reports the first-stage estimates as well as the KP F-statistics for instrument relevance. The instrument is standardized in all specifications to ease interpretation and the regressions also include controls for the direct shock ( $s_{ot}$ ) and origin district income ( $inc_{ot}$ ) to address demand effects and attenuation, respectively (see Section 5.4.1).

Table 5: Second Stage—Impact of Migration on Agricultural Development

	Capital (index)		Land (ac.)	Profits (Rs.)	Labour	
	(1)	(2)	(3)	(4)	(5)	(6)
	Investments	Stock	Cultivated	Crops	Wage Bill	Person-days
Male Migrants	-2.007*** (0.500)	-1.703*** (0.461)	-2.234*** (0.538)	-3.101*** (0.611)	0.269 (0.277)	-4.631*** (0.783)
Wt. Income	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income	Yes	Yes	Yes	Yes	Yes	Yes
Outcome SD	1.109	1.037	4.392	30291.547	547.124	344.519
HH FEs	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓
Observations	25928	24970	28748	25032	20910	20910
F-Stat	62.8	55.9	57.0	64.2	58.1	58.1

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. Columns 1 and 2 are indices for technology expenses and stock, respectively (see Section 4.2 for details of index construction). Crop profits (column 4) are net of expenses. Wage bill (column 5) is total wages paid to all workers in the past year. Person-days (column 6) includes both household and hired labour. Standard errors clustered at the PSU level.

Column 1 shows that the instrument strongly predicts labour outflows, consistent with Lemma 1. The point estimate is economically significant. Intuitively, it implies that a one standard deviation increase in destination income shocks increases out-migration from the origin by 36% relative to the mean ( $=0.074/0.201$ ). Another interpretation arises from the fact that the average household has 1.76 resident working-age males. The point estimate thus implies that a marginal income shock pulls 4% ( $=0.074/1.74$ ) of potential migrants away to join the destination labour force.

The remaining two columns test the sensitivity of the first-stage relationship. The point estimate is virtually unchanged when the share is number of educated household members as opposed to working age males (Column 2). The latter also measures migration potential since skilled workers are generally better able to fill destination jobs (Young, 2013). Estimates are also similar when both shift-shares are included together (Column 3). The F-statistic in all specifications is well above rule-of-thumb levels.

To proceed, we use the specification in column (1) as our preferred first stage. Table 3 showed that educated households are more exposed to the shock, and may be on a different technology path (although the statistical significance was at the 10% level only). In contrast, the number of resident working age males passed the balance test. Moreover, the F-statistic in the first column is substantially larger than in the other two, mitigating the risk of weak instrument bias.



## 6.2 Second Stage: Internal Migration Reduces Agricultural Development

We now turn to the 2SLS estimates of household responses to sectoral labour reallocation. Table 5 presents estimates of Equation (8). Again, all outcomes are expressed in standard deviations. After instrumenting migration, we find that labour migration causes households to uniformly *divest* from agriculture, starkly contrasting the OLS estimates.

Columns 1 and 2 show that the loss of a marginal worker causes statistically significant reductions in capital investment and capital stock, respectively. Appendix Table A4 decomposes estimates by the specific technology underlying each index. The investment result is driven primarily by reduced spending on seeds, agrochemicals, irrigation water, and rented equipment. The stock result is driven by reduced ownership of tubewells, water pumps, bullock carts, and threshing machines. At least eight of the ten technologies in Table A4 are labour complementary, indicating that the estimates are consistent with Proposition 1 from the model. Note that rentals include both hired equipment, such as tractors, as well as animals. Although disaggregated rentals are unavailable, case studies show that Indian farmers often rent power tillers, rotovators, and threshers, all of which require complementary labour for operation.

Columns 3-6 of Table 5 explore additional margins of response. Declining technology use reduces the marginal productivity of land, which may prompt farm contraction. Column 3 corroborates this logic: the point estimate implies that loss of a marginal worker causes households to reduce cultivated area by 2 standard deviations. Given the downward adjustment of both factors (technology and land), we expect output contraction. Column 4 shows that agricultural profits in the past year decline by 3 standard deviations, corresponding to about Rs. 90,000.

If labour markets function well, households would replace migrants by hiring workers at a higher wage. Column 5, however, shows no impact on origin wages. An alternative explanation is that loss of a single worker is insufficient to affect equilibrium wages. Column 6 shows a decline in labour person-days. If the migrant was replaced, there would be no effect. The estimate, then, is likely driven by the migrant himself, with no market replacement. Overall, these results provide evidence of India's poorly functioning rural labour market (Foster and Rosenzweig, 2022).

## 6.3 Summary of Results

The results of this section reveal a new and important fact: Indian households reduce technology use in response to labour loss. They also contract farm size, output, and employ less labour. Estimates are powered by a credible shift-share instrument for labour outflows and thus can reasonably be interpreted as causal. Overall, we document an overall decline in agricultural development as labour shifts from the countryside to the city. Our model showed that this response is rational if technology is labour complementary. If it were labour-saving, there would be technology uptake. Indeed, technology is labour complementary for nearly all technologies in our sample, and in India generally.

Yet, our results clearly do not generalize, given that urbanization and agricultural mechanization were observed together in the 20th century across developed countries (Alvarez-Cuadrado

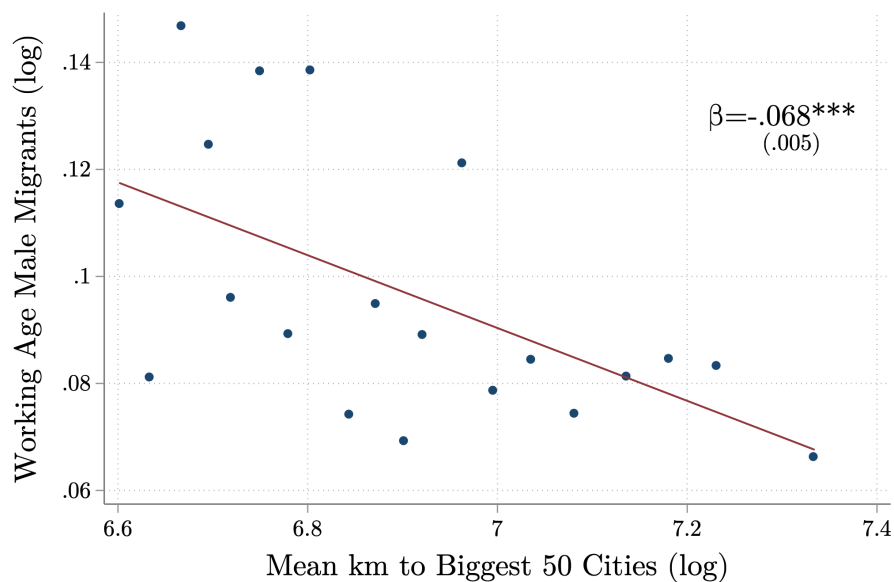


Figure 5: Number of Migrants and Household Remoteness

Note: The figure shows a binscatter of household migration against distance to 50 biggest cities by population. The y-axis is log number of of working age male migrants in a household from the IHDS panel. The x-axis is log of the straight-line distance from each IHDS district to the 50 biggest districts by population size. Year fixed effects are partialled out.

and Poschke, 2011). They also imply that food scarcity may become more of an issue as India urbanizes further. However, we have thus far only considered the direct impacts of labour migration. Next, we turn to an investigation of indirect impacts through economy-wide changes in factor prices.

## 7 General Equilibrium: The Spatial Re-organization of Agriculture

At first glance, our findings in the previous section suggest that India is failing to modernize the agricultural sector. They also question how the food needs of 1.5 billion people will be met as urbanization continue to unfold. Our results thus far only provide part of the answer, as they identified *direct* effects through a labour channel only (i.e. partial equilibrium). This section extends the analysis to general equilibrium and frames a deeper discussion about the *indirect* implications of internal migration for agricultural development. We document a counterbalancing pattern of agricultural modernization in remote regions, leading to a spatial reorganization of agriculture.

We showed in Figure 3 (Fact 4, Section 2.1) that while production declines in migrant-sending areas, in agreement with the estimates of the previous section, it increases in other parts of India. Figure 3B showed that production shifted to places experiencing minimal economic activity. Why does production move to remote regions (Figure 3)? A corollary of Lemma 1 is that remote households specialize in agriculture because they are less affected by the pull of urban growth

due to high movement costs. Figure 5 shows a binscatter of district remoteness (mean distance to the 50 biggest cities) and number of migrants residualized on year fixed effects. The inverse association implies that remote households indeed send significantly fewer migrants. The slope is the regression coefficient: households twice as remote send 7% fewer migrants.

Figure 5 is evidence of high mobility costs to distant, higher-growth destinations. Since the labour-loss channel is virtually nonoperational among remote households, production increases here must arise indirectly. We conjecture that prices are the main indirect channel. Intuitively, output contraction in peri-urban districts increases crop prices which, in turn, incentivizes remote households to increase production. Land prices are another potential indirect channel. The next section formally characterizes spatial spillovers via factor prices.

## 7.1 Model Extension

We now extend the model from Section 3 to allow endogenous prices to compete with the labour reallocation mechanism. We introduce the demand side of the economy, as well as land markets, to allow labour reallocation to affect technology choices directly (Table 5), as well as indirectly through crop and land prices. The model characterizes the spatial distribution of these two forces and structures our subsequent empirical analysis.

### 7.1.1 Set-Up

**Environment.** There are many rural regions (villages) indexed by  $j \in J$  and one urban centre. Each village  $j$  is endowed with land  $\bar{A}_j$  and inhabited by a set of heterogeneous households. Whereas we modelled production in isolation in Section 3, we now allow household production in each location to be spatially linked. We do this by introducing crop prices, variety, and suitability into the model. Households choose from a vector of crops  $k' \in [1, \dots, K]$  based on local crop suitability,  $\omega_{jk}$ . Spatial links arise through crop prices, and are stronger between households that grow similar crops. Crop suitability is homogeneous within villages and uncorrelated with distance to the center and across different crops. Several villages can be equally suitable for the same crop.

Households continue to supply one unit of labour inelastically. We also assume that agricultural labor productivity is homogeneous but that urban labour productivity is heterogeneous across households. This generates heterogeneous responses to urban productivity shocks.

Space in our economy is characterized by distance to the urban center. This determines the opportunity cost of agricultural labor. Residents of remote regions face lower opportunity costs of agriculture than urban and peri-urban households.

**Preferences.** Utility from consumption is CES across crops and Cobb-Douglas across sectors (agriculture and manufacturing). Homotheticity implies that we can consider a representative con-

sumer or the population of the economy. Assuming market clearing, consumption utility is:

$$U = \eta \log(Y^A) + (1 - \eta) \log(Y^M) \quad \text{s.t.} \quad \sum_{k=1}^K p_k Y_k + Y^M \leq E$$

$$\text{with } Y^A = \left( \sum_{k=1}^K Y_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where  $Y^M$  and  $Y_k$  are aggregate supply of manufactured goods and crop  $k$ , respectively.  $\eta$  is the agricultural expenditure share,  $E$  is total spending, and  $\sigma$  is the elasticity of crop substitution.

### 7.1.2 Crop Prices: The First Indirect Channel Linking Migration to Food Production

We derive a simple expression for  $p_k$  to show that crop prices are an indirect channel through which labour reallocation affects agricultural production. Following [Borusyak et al. \(2022b\)](#), we assume frictionless product markets, such that  $p_k$  is constant across regions:

$$p_k = \left( \frac{\eta E}{Y_k P^{1-\sigma}} \right)^{\frac{1}{\sigma}} \quad (10)$$

where  $P$  is the Dixit-Stiglitz price index,  $P := \left( \sum_{k'} p_{k'}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$  (see [Appendix A.2.1](#) for derivation). Equation (10) shows that crop prices decrease in aggregate supply and increase in total expenditure. [Table 5](#) already showed that labour loss causes output contraction. Moreover, urbanization is accompanied by increased expenditure. We thus expect that labour outflows will generate a positive indirect effect on crop production through rising crop prices.

### 7.1.3 Household Production

We now characterize household production with endogenous crop and land prices. There are a finite number of crops indexed by  $k' \in [1, \dots, K]$  that can be produced in village  $j$ . Production is a function of labor, land and technology. Household  $i$  chooses how to allocate factors to each crop  $k$  following a constant returns to scale Cobb-Douglas production function:

$$f(l_{ijk}, a_{ijk}, \theta_{ijk}) = \omega_{jk} a_{ijk}^{\alpha} l_{ijk}^{\beta} \theta_{ijk}^{1-\alpha-\beta} \quad (11)$$

where  $l_{ijk}$  is labor of household  $i$ , in village  $j$  allocated to crop  $k$ . Similarly,  $a_{ijk}$  is land, and  $\theta_{ijk}$  is technology. The parameter  $\omega_{jk}$  is village  $j$ 's agricultural suitability for growing crop  $k$ .  $\alpha$  and  $\beta$  are the output elasticities of land and labor, respectively.

Rural labour markets remain absent. However, in contrast to [Section 3](#), land markets are perfectly functioning to allow farm consolidation as an adjustment margin. The profit-maximizing household  $i$  in village  $j$  solves:

$$\max_{a_{ijk}, \theta_{ijk}, l_{ijk}} \sum_{k=1}^K p_k \omega_{jk} a_{ijk}^{\alpha} l_{ijk}^{\beta} \theta_{ijk}^{1-\alpha-\beta} - r_j a_{ijk} - v \theta_{ijk} - \frac{w \varphi_{ij}}{\tau_j} l_{ijk} \quad (12)$$

where  $p_k$  is the homogenous price of crop  $k$ ,  $r_j$  is the land price in village  $j$ , and  $v$  is the exogenously given rental rate of capital.  $\varphi_{ij}$  is urban labor productivity of household  $i$  from village  $j$ . As before,  $w$  is the urban wage and  $\tau_j \geq 1$  is the migration cost which increases with distance. The following lemmas summarize the solution to the farmers problem.

**Lemma 2.** *Each farmer produces only one crop*

**Lemma 3.** *All farmers within one region produce the same crop.*

**Lemma 4.** *Several regions may produce the same crop.*

*Proof.* Proofs are in Appendix A.2.2. □

Households in each village  $j$  specialize in a single crop (Lemma 2) because the production function is homogenous of degree one. They specialize in the most suitable crop since there are no benefits from spreading inputs across more crops. Moreover, since all households on  $\bar{A}_j$  face the same crop suitability and have homogenous agricultural skills, they all produce the same crop (Lemma 3). By this logic, villages with the same most suitable crop will specialize in it (Lemma 4). We therefore drop the  $k$  subscripts in the remainder of the model.

#### 7.1.4 Land Prices: The Second Indirect Channel Linking Migration to Food Production

Land markets are competitive such that land price,  $r_j$ , equals the marginal productivity of land. Using the FOCs from Equation (12) with the fact that  $\sum_{i=1}^{n_j} a_{ij} = \bar{A}_j$ , the equilibrium land price is:

$$r_j = p_j \alpha \left( \frac{\sum_{i=1}^{n_j} \left( l_{ij}^\beta \theta_{ij}^{1-\alpha-\beta} \right)^{\frac{1}{1-\alpha}}}{\bar{A}_j} \right)^{1-\alpha} \quad (13)$$

where  $p_j$  is the price of the crop grown in village  $j$ , and  $n_j$  is the number of households in village  $j$ .<sup>7</sup> The derivation is in Appendix A.2.3.

Intuitively,  $r_j$  is a function of effective labor per unit of available land. Suppose technology is labour complementary. Then a decline in effective labour (numerator in Equation 13) reduces land productivity which, in turn, deflates land prices in village  $j$ . This introduces a second indirect channel (the first was crop prices) through which labour outflows affect household production. Households in villages with high migration, but who do not send migrants themselves, increase farm size in response to declining land prices.

#### 7.1.5 The Spatial Organization of Agriculture

We now characterize space in the economy. Villages differ in crop suitability and crop specialization (Lemmas 2-3). The degree to which household  $i$  in village  $j$  specializes in agriculture versus

---

<sup>7</sup>For identical individuals, the expression simplifies to  $r_j = p \alpha n_j \frac{l_{ij}^\beta \theta_{ij}^{1-\alpha-\beta}}{\bar{A}_j^{1-\alpha}}$ .

manufacturing depends on the opportunity cost of agricultural labour. This cost is a combination of their value of  $\varphi_{ij}$  (urban productivity) and  $\tau_j$  (distance to the urban centre). There is a cost threshold beyond which households will not send migrants.

**Proposition 2.** *Households with  $\varphi_{ij} < \bar{\varphi}_j$  will not participate in the urban labor market and specialize in agriculture.  $\bar{\varphi}_j$  is the level of  $\varphi_{ij}$  that makes the marginal household indifferent between agricultural and urban specialization:*

$$\bar{\varphi}_j = \frac{\tau_j}{w} \phi_0 \quad (14)$$

where  $\phi_0$  is a constant.

*Proof.* The proof is in Appendix A.2.4. □

Equation (14) characterizes the extensive margin of the spatial economy.  $\bar{\varphi}_j$  increases with distance to the centre such that more of the population specializes in agriculture in remote areas.

### 7.1.6 General Equilibrium: Urban Productivity Shocks and Agricultural Production

To close the model, we combine all features of the economy to characterize structural change, agricultural development, and its spatial distribution. We explicitly decompose direct labour reallocation mechanisms as well as indirect price mechanisms in way that can easily recovered from the data. To start, note that the impact of an urban productivity shock on agricultural production in village  $j$  is the sum of production responses by each household  $i$  in the village:

$$\frac{dY_j}{dw} = \sum_{i=1}^{n_j} \frac{dy_{ij}}{dw}$$

The crop subscript is dropped because of Lemmas 2 and 3. In the rest of this section, subscripts denote partial derivatives. Appendix A.2.5 shows that the optimal response of each household to the urban shock can be written as a weighted sum of their labor, land and capital adjustments:

$$y_w = \frac{a_w}{a} \alpha y + \frac{l_w}{l} \beta y + \frac{\theta_w}{\theta} (1 - \alpha - \beta) y \quad (15)$$

To make explicit that these adjustments are not only determined by the opportunity costs of agricultural labour, but also by changes in crop and land prices (Equations 10 and 13), we totally differentiate the FOCs from Equation (12) with respect to  $w$ . We then express  $\frac{a_w}{a}$  and  $\frac{\theta_w}{\theta}$  in terms of  $l_w$  and insert them into Equation (15). This yields a parsimonious expression for the general equilibrium effect of urban labour shocks on agricultural production:

$$y_w = \underbrace{\phi_1 l_w}_{\text{partial equilibrium}} + \underbrace{\phi_2 p_w + \phi_3 r_w}_{\text{general equilibrium}} + \underbrace{\phi_4}_{\text{household FE}} \quad (16)$$

where  $\phi_1 := \frac{y}{l}$ ,  $\phi_2 := -\frac{y}{p} \frac{(1+\alpha)}{\alpha}$ ,  $\phi_3 := \frac{y a}{\alpha y}$  and  $\phi_4 := \frac{\phi l (\alpha + \beta)}{\tau y \beta \alpha} (\alpha + y \alpha - y)$ . The derivations are in Appendix A.2.5. The first term is the partial equilibrium effect from Equation (2), which states that

the direct impact of increased urban wages is through sectoral labour reallocation. Table 5 showed that, in response, households decrease labour complementary technology and production.

The second and third term describe general equilibrium effects. Labour reallocation increases aggregate food demand ( $E$ ) and decreases aggregate food supply ( $Y_k$ ), both which increase crop prices and production (Equation 10). Labour loss also decreases land prices (Equation 13), which increases farm size and production. These indirect channels both counterbalance the partial equilibrium effect, leaving the net effect unclear. The fourth term is a household intercept.

Whether the partial or general equilibrium effect dominates depends on the spatial structure of the economy. Recall that households with  $\varphi_{ij} < \bar{\varphi}_j$  do not engage in urban work, and that  $\bar{\varphi}_j$  increases with distance to the centre. Therefore, remote households circumvent the direct labour effect, but still respond to indirect price changes. The overall impact of urban wage-induced shocks to labour reallocation depends on the spatial incidence of direct and indirect effects.

## 7.2 From Theory to Empirics

The extended model establishes that the effect of labour reallocation on agricultural development is an empirical question, due to the existence of opposing partial and general equilibrium forces. We already showed in Section 6 that the partial equilibrium effect is negative: households reduce technology use and contract output in response to labour loss. We turn to estimating general equilibrium effects by mapping features of the model economy to observables.

### 7.2.1 Linking to Observables

**Interpreting  $\varphi_{ij}$  As An “Exposure”.** The objective is to empirically decompose the direct and indirect channels shown in Equation (16). We know from Proposition 2 that the indirect effect is concentrated on remote households, who are more likely to satisfy  $\varphi_{ij} < \bar{\varphi}_j$  and evade the direct effect.  $\varphi_{ij}$  represents household  $i$ 's urban labour productivity i.e. the ability to take advantage of the urban shock. In shift-share language, it measures exposure to the shock. We use number of baseline resident working age males to proxy  $\varphi_{ij}$  to be consistent with previous sections.

$\varphi_{ij}$  allows us to disentangle the direct and indirect channels, but it does not help decompose the two indirect channels (crop and land prices). For example, if  $\varphi_{ij} = 0$  (no resident males), then household  $i$  sends zero migrants in response to the urban shock (first stage). Therefore, any change in their technology or production must come through price effects.

**Decomposing the Crop Price Channel.** To empirically isolate the crop price channel, we use the fact that prices are spatially linked across regions growing the same crop. From Equation (10), we know that  $p_k$  changes in response to shifts in aggregate supply ( $Y_k$ ) of crop  $k$ . We also know from Lemmas 2-4 that several regions can specialize in  $k$ . Therefore, household production responds to changes in  $Y_k$  and  $p_k$  if they also grow  $k$ .

We map from theory to data using the spatial aggregation technique of Adao et al. (2019). We compute a leave-one-out weighted average of household labour reallocation with weights equal

to the similarity of crop portfolios at baseline. Crop similarity is defined by euclidean distance:

**Definition 2.** Let  $\vec{v}$  and  $\vec{u}$  be two  $K \times 1$  indexed vectors with  $K$  equal to the number of possible crops (in our data  $K=78$ ).  $\vec{v}_i$  is the value share of crop  $i$ , equal to zero if  $i$  is not grown. Crop similarity is defined as inverse euclidean distance between two portfolios:  $d(\vec{u}, \vec{v})^{-1} = \left( \sqrt{(\vec{u}_1 - \vec{v}_1)^2 + (\vec{u}_2 - \vec{v}_2)^2 + \dots + (\vec{u}_K - \vec{v}_K)^2} \right)^{-1}$ .

These weights generate a measure of aggregate labour reallocation between similar-crop growing regions, which is our empirical analog for  $p_w$  in Equation (16). Intuitively, if urban shocks triggered widespread labour reallocation from a rice growing region, reducing  $Y_{rice}$ , then household  $i$  elsewhere in the country will respond to the price increase if they also grow rice.

**Decomposing the Land Price Channel.** To empirically isolate the land price channel, we use the fact that land prices are a function of effective labour per unit of available land *in the village* (Equation 13). Land prices thus operate at the village level in our model and decrease with village-level labour outflows. Whereas the crop price effect was location-independent but crop-dependent, the opposite is true for the land price effect. Therefore, we proxy  $r_w$  in Equation 16 with the leave-one-out average number of migrants across households in village  $i$ .

## 7.2.2 Estimation Strategy

We now specify our empirical analogue to Equation (16). We estimate the total effect of labour reallocation on agricultural development with the following 2SLS regression:

$$Y_{iot} = \underbrace{\beta_1 \widehat{Migrants}_{iot}}_{\text{labour channel}} + \beta_2 \underbrace{\sum_{j \in N/i} d(\vec{i}, \vec{j})^{-1} Migrants_{jot}}_{\text{crop price channel}} + \beta_3 \underbrace{\sum_{j \in N_j/i} Migrants_{jot}}_{\text{land price channel}} \quad (17)$$

$$+ \beta_4 s_{ot} + \beta_5 inc_{ot} + \alpha_i + \gamma_t + \varepsilon_{iot}$$

where  $Y_{iot}$  is the agricultural outcome of household  $i$  in origin district  $o$  at time  $t$ . The first term is the same as in Equation (8) and captures the direct effect of migration on agricultural production through the labour reallocation channel. It is instrumented with the same shift-share instrument based on distance-weighted urban income shocks (Section 5.2).

The second term is the weighted average number of migrants sent from all other households  $j \neq i$ , with weights equal to crop similarity between  $i$  and  $j$  (Definition 2). In our main specification, we define “other households” as those within the state, due to complex market restrictions preventing crop prices from equilibrating across states (Chatterjee, 2021).  $\beta_2$  thus captures household responses through the indirect crop channel. The third term is the average number of migrants across households  $j \neq i$  within the same village.  $\beta_3$  thus captures household responses through the indirect land channel.  $s_{ot}$ ,  $inc_{ot}$ , and  $\alpha_i$  are as defined earlier.

Since the indirect crop channel is measured by a leave-one-out weighted average at the state level, it already constitutes a deviation from the state mean. Therefore, include only year fixed-effects,  $\gamma_t$  since state-year fixed effects are redundant.  $\gamma_t$  is also crucial for identifying  $\beta_2$  because



Table 6: Partial and General Equilibrium Effects (State Price Channel)

	(1) Profits (Rs.)	(2) Farm Size (ac.)	(3) Investment	(4) Assets
Male Migrants (direct labour channel)	-1.689*** (0.389)	-1.230*** (0.324)	-1.361*** (0.320)	-1.102*** (0.283)
Village emigration (indirect land channel)	0.314*** (0.087)	0.255*** (0.068)	0.268*** (0.068)	0.258*** (0.061)
Other-HH emigration (indirect crop channel)	0.226*** (0.064)	0.190*** (0.055)	0.282*** (0.063)	0.060 (0.055)
Wt. Income	Yes	Yes	Yes	Yes
Origin Income	Yes	Yes	Yes	Yes
HH FEs	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓
N	20150	26524	25662	24632
F-Stat on Migrants	48.0	52.6	61.7	54.8

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . *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. *Other-HH emigration* is the leave-one-out number of migrants in the state weighted by crop similarity. All variables are standardized. Crop profits are net of expenses (Column 1). Column 2 is cultivated area. Columns 3 and 4 are indices for technology expenses and stock, respectively (see Section 4.2 for details of index construction). Standard errors clustered by PSU.

it disentangles shifts in aggregate supply from aggregate demand. Equation (10) states that equilibrium  $p_k$  is determined by aggregate supply of  $k$  and total expenditure. Since expenditure is not crop-specific, and rises uniformly with urbanization, it is absorbed by  $\gamma_t$ . This leaves  $\beta_2$  to capture household responses through a supply channel only. Variation in supply is retained even after including  $\gamma_t$  because supply is crop-specific.

**Main Identification Assumption.** Note that household migration (first term) is instrumented while the aggregate counterparts (second and third term) are not. We do this to address our main identification concern: the endogeneity of the household migration decision. The main identification assumption of Equation (17) then is that the migration decision of other households  $j$  is orthogonal to that of the focal household  $i$ , conditional on controls and fixed effects. The biggest concern is that households  $i$  and  $j$  in the same district experience the same destination income shocks. However, inclusion of  $s_{ot}$  as a control alleviates this concern.

### 7.3 General Equilibrium Results

This section quantifies partial and general equilibrium effects of labour reallocation on agricultural development in a unified empirical framework. Although the theory emphasized production as

the key outcome, we also present estimates of Equation (17) with technology and land as the outcome in order to infer mechanisms. Table 6 presents the results. All variables are standardized.

**Results Overview.** As predicted by the theory, the main result is that partial and general equilibrium effects draw in opposing directions. Across all columns, the direct effect (row 1) is negative, mirroring the agricultural decline observed in Table 5. However, the indirect channels temper the decline. Starting with column 1, declining output<sup>8</sup> through the labour channel is partially offset through the indirect land channel (row 2). Our model showed that the mechanism is through declining land prices (Equation 13) which prompts farm expansion. The positive coefficient in row 2 when farm size is the outcome (column 2) corroborates the theoretical result. Lastly, column 1 row 3 shows that direct output declines are also offset through the indirect crop channel. Our theory showed that this occurs through rising crop prices induced by the direct effect.

The point estimates are economically significant. Whereas households reduce production by 1.7 standard deviations in response to their own labour loss (column 1), output increases through the indirect land channel offset this by 19%. The indirect crop effect is 13% of the direct labour reallocation effect. The counterbalancing pressures are similar for farm size changes (column 2).

How are households able to expand production through the two indirect channels? Column 3 shows that they increase expenditure on yield-enhancing technology. Table A6 decomposes the technology index by the component variables. Households respond to rising crop prices (row 3) by spending more on seeds, fertilizer, pesticide, irrigation water, and rental equipment. This fuels the output expansion which, in turn, puts offsetting pressure on migration-induced output declines. In contrast, the indirect crop channel has no impact on the stock of machinery (Table A6 Column 4, row 3). Among the underlying variables, there is only weakly significant adoption of water pumps (Table A6 column 7).

**The Spatial Incidence of Direct and Indirect Effects.** The results in Table 6 have important implications for the spatial reorganization of agriculture in response to structural change. We establish this point by distinguishing which households experience the direct and indirect effects. The direct effect dominates for households sending many migrants. These households experience large output declines (column 1 row 1) that are not offset by the indirect effects. In contrast, households sending zero migrants will exclusively experience the indirect effects of other households' migration. The labour channel is "switched off" for non-migrant households but they still increase production and technology use in response to the indirect effects.

Therefore, agricultural production is expected to decline in areas with high migration and increase in areas with low migration. Figure 5B established that migration follows a spatial gradient; it declines with remoteness due to higher movement costs. This was built into our model in Proposition 2 which revealed a distance-based threshold beyond which households do not send migrants. Our results thus imply that agricultural production declines near urban areas, where

---

<sup>8</sup>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.

migration is highest, and compensate with increases in remote areas, where migration is lowest. This characterizes the spatial reorganization of agriculture in response to labour reallocation.

#### 7.4 Counterfactual Simulations: The Aggregate Extent of Spatial Spillovers

The previous results quantify the redistribution of production across households. We now aggregate our estimates to study national changes in food supply. We find that 63% of the decline in aggregate food supply near urban centres is compensated by production booms in remote areas.

**Methods.** We define a set of counterfactuals that allow us to disentangle the direct and indirect effect channels. For each household  $i$  in year  $t$ , we use the coefficients from Equation (17) 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  $t_1$ :

$$Y_{iot_2}^{NM} = \hat{\beta}_1 \widehat{Migrants}_{iot_1} + \hat{\beta}_2 \sum_{j \in N/i} d(\vec{i}, \vec{j})^{-1} Migrants_{jot_1} + \hat{\beta}_3 \sum_{j \in N_j/i} Migrants_{jot_1} + \alpha_i + \gamma_{t_2}$$

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

$$Y_{iot_2}^{LO} = \hat{\beta}_1 \widehat{Migrants}_{iot_2} + \hat{\beta}_2 \sum_{j \in N/i} d(\vec{i}, \vec{j})^{-1} Migrants_{jot_1} + \hat{\beta}_3 \sum_{j \in N_j/i} Migrants_{jot_1} + \alpha_i + \gamma_{t_2}$$

In the same way, we define a *Labour and Land* (LL) scenario where the indirect crop effect is fixed at  $t_1$ , as well as a *Labour and Crop* (LC) scenario where the indirect land effect is fixed at  $t_1$ .

We aggregate these local predictions into national counterfactuals:  $NatY_{iot_2}^{NM}$ ,  $NatY_{iot_2}^{LO}$ ,  $NatY_{iot_2}^{LL}$  and  $NatY_{iot_2}^{LC}$ , the national production value without migration, with migration but no spatial spillovers, with migration plus land spillovers, and with migration plus crop spillovers, respectively. Comparisons with in-sample fitted values  $NatY_{iot_2}$  (the scenario with all channels operational) yield three statistics of interest. The first and second are national values of food production with and without spatial spillovers relative to the total possible value absent migration:

$$PctChange^{SS} = 100 \cdot \left( \frac{NatY_{iot_2} - NatY_{iot_2}^{NM}}{NatY_{iot_2}^{NM}} \right) \quad PctChange^{NSS} = 100 \cdot \left( \frac{NatY_{iot_2}^{LO} - NatY_{iot_2}^{NM}}{NatY_{iot_2}^{NM}} \right) \quad (18)$$

The third is the amount of food decline through the labour channel that is offset by spatial spillovers, as a percentage of the counterfactual change absent spillovers:

$$PctOffset = 100 \cdot \left( \frac{PctChange^{NSS} - PctChange^{SS}}{PctChange^{NSS}} \right) \quad (19)$$

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

**Simulation Results.** Figure 6A shows simulation estimates of the aggregate change in food supply

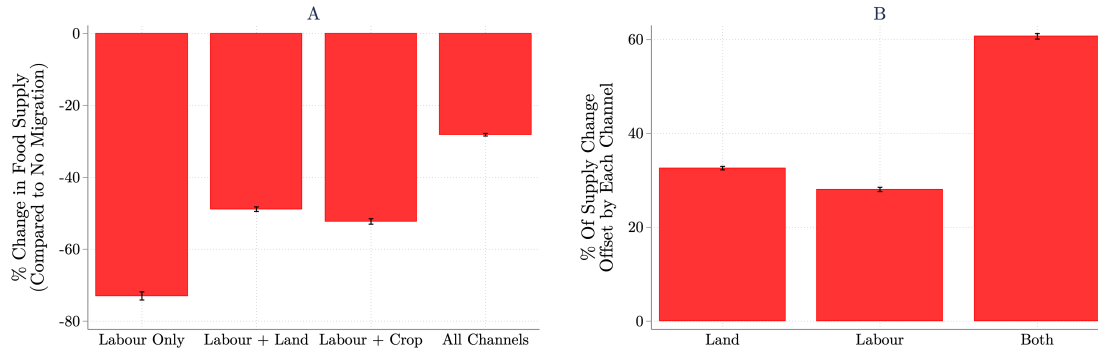


Figure 6: Aggregate Influence of Spatial Spillovers

Note: Panel A displays the aggregate change in food supply from migration under the four scenarios (Equation 18). Labour only means the crop and land channels are held constant, labour + crop means the land channel is held constant, and so on. Panel B shows the percent of the *Labour Only* food decline mitigated by the indirect land and crop forces (Equation 19). Confidence intervals computed from 1000 bootstrap draws.

with and without indirect effects (Equation 18). Under the *Labour Only* scenario, with general equilibrium effects shut off, aggregate migration would have caused a 73% reduction in food supply (compared to the *No Migration* counterfactual scenario). This amounts to Rs. 240 million worth of food. When the labour, land, and crop channels all operate, the supply contraction becomes two-and-a-half times smaller. Error bars correspond to 95% confidence intervals from bootstrapping the prediction and calculation procedure with 1000 draws.

Panel B shows the extent to which migration-induced food shortages are mitigated by indirect price effects (Equation 19). 33% of the damage is mitigated by reduced land prices, and 28% by increased crop prices. Both indirect channels together mitigate 63% of the direct effect of migration. The recovered food amounts to Rs. 147 million, or, 26% of the total in-sample value of crop production.

The simulation results have important distributional implications. We showed in Table 6 that indirect production increases are concentrated among non-migrant households in remote areas (Figure 5B). We also know that India has one of the world’s largest rural-urban wage gaps (Munshi and Rosenzweig, 2016), with rural households substantially poorer than their urban counterparts. Our results therefore imply that the Rs. 147 million of savings accrue to the rural poor in the form of increased crop production, technology adoption, and agricultural development in general.

## 8 Conclusion

The reallocation of labour from farms to cities is an emblematic feature of the economic development process. Many studies document the implications of structural change for manufacturing growth in destinations or agricultural growth at the origin. We study the latter, extending existing work by quantifying spatial spillovers in regions that did not urbanize. In doing so, we provide

new evidence on structural change and the spatial reorganization of agriculture.

We track labour reallocation and agricultural development using a detailed household panel survey from India between 2005 and 2012, a period of rapid economic modernization. We instrument out-migration with a shift-share design based on distance-weighted destination income shocks that draw workers away from agriculture (the shift). This is interacted with a measure of household migration potential that proxies for exposure to the shock (the share). We find that migration causes a contraction of agricultural technology, food production, and farm size.

Whether these results imply aggregate food scarcity requires an investigation of general equilibrium. In aggregate, declining household food production drives up crop prices. Similarly, increased land availability reduces land prices. Guided by a two-sector 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 two forces draw in opposing directions. Importantly, we find that households with no migrants, who live in remote areas, are the ones increasing production in response to price effects. Documenting the incidence of these spillovers allow us to characterize the spatial reorganization of agriculture toward remote areas.

Our best estimate is that spatial spillovers mitigate 60% of the decline in food production induced by labour loss between 2005-2012. The spatial redistribution of agriculture through market-based forces is therefore economically significant but not a panacea. Although India has one of the largest agricultural workforces in the world, 40% of the decline in food supply is left uncompensated by domestic smallholders.

Two important issues in this paper deserve further study. The first is where the “missing food” comes from. Logically, it either comes from imports or industrial farming, which is not captured in our sample. The second relates to distributional consequences of structural change. We showed that while agricultural development declines in peri-urban areas, it surges in remote areas where migration is low and poverty is widespread. We thus expect that structural transformation promotes income redistribution toward even those who do not directly participate. These are both open areas for future research.

## References

- Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge**, “When should you adjust standard errors for clustering?,” Technical Report, National Bureau of Economic Research 2017.
- 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.
- Aggarwal, Shilpa**, “Do rural roads create pathways out of poverty? Evidence from India,” *Journal of Development Economics*, 2018, 133, 375–395.
- 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.
- 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*, 2021.
- **and Paul Novosad**, “Rural roads and local economic development,” *American Economic Review*, 2020, 110 (3), 797–823.
- **, Teevrat Garg, and Paul Novosad**, “The ecological impact of transportation infrastructure,” *The Economic Journal*, 2020, 130 (629), 1173–1199.
- Blakeslee, David, Aaditya Dar, Ram Fishman, Samreen Malik, Heitor Pelegrina, and Karan Singh**, “Irrigation and the Spatial Pattern of Local Economic Development in India,” *Mimeo*, 2022.
- 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.
- 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.

- Chatterjee, Shoumitro**, “Market Power and Spatial Competition in Rural India,” *Mimeo*, 2021.
- 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.
- 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.
- International Monetary Fund**, “World Economic Outlook, April 2015,” Technical Report, International Monetary Fund 2015.
- Kone, Zovanga L, Maggie Y Liu, Aaditya Mattoo, Caglar Ozden, and Siddharth Sharma**, “Internal borders and migration in India,” *Journal of Economic Geography*, 2018, 18 (4), 729–759.
- Manuelli, Rodolfo E and Ananth Seshadri**, “Frictionless technology diffusion: The case of tractors,” *American Economic Review*, 2014, 104 (4), 1368–91.
- Mendola, Mariapia**, “Migration and technological change in rural households: Complements or substitutes?,” *Journal of Development Economics*, 2008, 85 (1-2), 150–175.
- Munshi, Kaivan and Mark Rosenzweig**, “Networks and misallocation: Insurance, migration, and the rural-urban wage gap,” *American Economic Review*, 2016, 106 (1), 46–98.
- Pellegrina, Heitor S**, “Trade, productivity, and the spatial organization of agriculture: Evidence from Brazil,” *Journal of Development Economics*, 2022, 156, 102816.
- Rozelle, Scott, J Edward Taylor, and Alan DeBrauw**, “Migration, remittances, and agricultural productivity in China,” *American Economic Review*, 1999, 89 (2), 287–291.
- 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.

**UN Population Division**, "World Urbanization Prospects: 2018 Revision," Technical Report, United Nations Department of Social Affairs 2018.

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

**Young, Alwyn**, "Inequality, the urban-rural gap, and migration," *The Quarterly Journal of Economics*, 2013, 128 (4), 1727–1785.



## A Proofs

### A.1 Partial Equilibrium Proofs

#### A.1.1 Labor markets

Here we discuss the impact of rural labor markets on the results. First, we introduce perfect rural labor markets into our model such that the problem of the household becomes

$$\begin{aligned} \max_{\theta, l} & pf(l^a, a, \theta) - v\theta - hw^h + \frac{w}{\tau}(1-l) \\ & l^a = l + h \end{aligned}$$

i.e. agricultural labor is the sum of hired labor,  $h$ , and household agricultural labor  $l$ . Labor can be hired for the rural wage rate  $w^h$ . The FOCs are then given by:

$$\begin{aligned} pf_{l^a} - \frac{w}{\tau} &= 0 \\ pf_{l^a} - w^h &= 0 \\ pf_{\theta} - v &= 0 \end{aligned}$$

Combining the first and second condition implies  $w^h = \frac{w}{\tau}$ . Therefore, adding a perfect functioning rural labor market does not change our results.

### A.1.2 Proof of Lemma 1

Using subscripts to denote partial derivatives, the first order conditions (FOCs) are:

$$pf_l - \frac{w}{\tau} = 0$$

$$pf_\theta - v = 0$$

The first condition states that the farmer sends labour to the city as long as the manufacturing wage exceeds marginal product of labour if the migrant were instead to work on the family farm. The second condition is the standard FOC for capital.

We first totally differentiate the FOCs with respect to urban wages  $w$  to see how farmers re-optimize in response to destination wage shocks.

$$pf_{ll}l_w + pf_{l\theta}\theta_w - \frac{1}{\tau} = 0 \tag{20}$$

$$pf_{\theta\theta}\theta_w + pf_{\theta l}l_w = 0. \tag{21}$$

Substituting (21) into (20) and solving for  $l_w$  yields

$$l_w = \frac{1}{p\tau} \left( \frac{f_{\theta\theta}}{f_{ll}f_{\theta\theta} - f_{l\theta}^2} \right) \tag{22}$$

From the sufficiency conditions for the farmer's optimization problem we know that  $f_{ll}f_{\theta\theta} - f_{l\theta}^2 > 0$  and  $f_{\theta\theta} < 0$  which completes the proof.

## A.2 General Equilibrium Proofs

### A.2.1 Prices

The problem of the consumer can therefore be written as

$$L = \left( \sum_{k=1}^K Y_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - \lambda \left( \sum_{k=1}^m Y_k p_k - \eta E \right) \quad (23)$$

with the first order condition

$$\left( \sum_{k=1}^K Y_k^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} Y_k^{\frac{-1}{\sigma}} = \lambda p_k$$

dividing the condition for  $k$  by the condition for  $k'$  yields

$$\begin{aligned} \left( \frac{Y_k}{Y_{k'}} \right)^{-\frac{1}{\sigma}} &= \frac{p_k}{p_{k'}} \\ Y_k &= p_k^{-\sigma} p_{k'}^{\sigma} Y_{k'} \\ Y_k p_k &= p_k^{1-\sigma} p_{k'}^{\sigma} Y_{k'} \end{aligned}$$

summing over all  $k$ 's yield

$$\begin{aligned} \sum_k Y_k p_k &= p_{k'}^{\sigma} Y_{k'} \sum_k p_k^{1-\sigma} \\ \eta E &= p_{k'}^{\sigma} Y_{k'} \sum_k p_k^{1-\sigma} \\ p_{k'}^{\sigma} &= \frac{\eta E}{Y_{k'} \sum_k p_k^{1-\sigma}} \\ p_{k'} &= \left( \frac{\eta E}{Y_{k'} \sum_k p_k^{1-\sigma}} \right)^{\frac{1}{\sigma}} \end{aligned}$$

The price of crop  $k$  can then be expressed as

$$p_k = \left( \frac{\eta E}{Y_k P^{1-\sigma}} \right)^{\frac{1}{\sigma}}$$

where  $P$  is the Dixit-Stiglitz or ideal price index defined as  $P := \left( \sum_{k'} p_{k'}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ . The crop prices declines in the total production of that crop.

### A.2.2 Specialization

First note that the production function is homogeneous of degree one and that the first order conditions are given by

$$\begin{aligned}\alpha p_k \omega_{jk} a_{ijk}^{\alpha-1} l_{ijk}^\beta \theta_{ijk}^{1-\alpha-\beta} - r &= 0 \\ \beta p_k \omega_{jk} a_{ijk}^\alpha \beta l_{ijk}^{\beta-1} \theta_{ijk}^{1-\alpha-\beta} - \frac{w \varphi_{ij}}{\tau_j} &= 0 \\ (1 - \alpha - \beta) p_k \omega_{jk} a_{ijk}^\alpha l_{ijk}^\beta \theta_{ijk}^{-\alpha-\beta} - v &= 0.\end{aligned}$$

The first order conditions therefore imply  $a_{ij1} = a_{ij2} = \dots = a_{ijK}$  for  $p_1 \omega_{1j} = p_2 \omega_{2j} = \dots = p_K \omega_{jK}$ . Denoting the solutions of the first order conditions by  $a^*$ ,  $l^*$  and  $\theta^*$  and assuming  $p_1 \omega_{1j} = p_2 \omega_{2j} = \dots = p_K \omega_{jK}$  yields

$$\sum_{k=1}^K p_k \omega_{jk} a_{ijk}^{*\alpha} l_{ijk}^{*\beta} \theta_{ijk}^{*1-\alpha-\beta} = K p_{k'} \omega_{jk'} a_{ijk'}^{*\alpha} l_{ijk'}^{*\beta} \theta_{ijk'}^{*1-\alpha-\beta} = p_{k'} \omega_{jk'} (K a_{ijk'}^*)^\alpha (K l_{ijk'}^*)^\beta (K \theta_{ijk'}^*)^{1-\alpha-\beta}$$

for any  $k' \in [1, \dots, K]$ . Finally  $p_{k'} \omega_{jk'} a_{ijk'}^\alpha l_{ijk'}^\beta \theta_{ijk'}^{1-\alpha-\beta} > p_{k'} \omega_{jk'} a_{ijk'}^\alpha l_{ijk'}^\beta \theta_{ijk'}^{1-\alpha-\beta}$  for  $p_{k'} \omega_{jk'} > p_{k'} \omega_{jk'}$  completes the proof of part a).

Part b) follows from the fact that  $p_k \omega_{jk}$  is the same across all farmers within one region and farmers do not differ with respect to their agricultural productivity.

Part c) follows from the fact that two regions  $j'$  and  $j''$  with  $p_{k'} \omega_{j'k'} > p_k \omega_{j'k}$  and  $p_{k'} \omega_{j''k'} > p_k \omega_{j''k}$  for any  $k \neq k'$  will both specialize in  $k'$ .

### A.2.3 Land markets

Use the first order condition and rearrange to arrive at

$$p\alpha a^{\alpha-1}l^{\beta}\theta^{1-\alpha-\beta} = r \iff a = \left( \frac{p\alpha l^{\beta}\theta^{1-\alpha-\beta}}{r} \right)^{\frac{1}{1-\alpha}}.$$

Aggregate demand in region  $j$  is constrained by the total land endowments such that

$$\sum_i a_{ij} = \bar{A}_j.$$

Use the individual demand for land in the constraint and rearrange to arrive at

$$r = p\alpha \left( \frac{\sum_i (l^{\beta}\theta^{1-\alpha-\beta})^{\frac{1}{1-\alpha}}}{\bar{A}} \right)^{1-\alpha}.$$

#### A.2.4 Spatial organization

There is a level of  $\varphi$  which makes a farmer indifferent between engaging in urban production or specialization in agriculture. This level is defined by equalizing the marginal productivity in urban and agricultural production when  $l^* = 1$ . This level is defined by

$$\beta p_k \omega_{jk} a_{ijk}^{*\alpha} l_{ijk}^{*\beta-1} \theta_{ijk}^{*1-\alpha-\beta} = \frac{w \bar{\varphi}_j}{\tau_j}.$$

where  $a_{ijk}^*$ ,  $l_{ijk}^*$  and  $\theta_{ijk}^*$  solve the first order conditions. In the following we drop region and farmer indices and use subscripts to denote partial derivatives. By definition  $l^* = 1$ . Next, divide the first order conditions by  $py$  and solve for  $a^*$ ,  $l^*$  and  $\theta^*$  respectively:

$$\begin{aligned} a^* &= \frac{\alpha py}{r} \\ l^* &= \frac{\beta py \tau_j}{w \varphi} \\ \theta^* &= \frac{(1 - \alpha - \beta) py}{v} \end{aligned}$$

Next, set  $l^* = 1$  and rearrange the first order condition for labor:

$$py = \frac{w \varphi}{\beta \tau_j}$$

Insert this expression in the first order conditions for technology and land:

$$\begin{aligned} a^* &= \frac{\alpha w \varphi}{r \beta \tau_j} \\ \theta^* &= \frac{(1 - \alpha - \beta) w \varphi}{v \beta \tau_j} \end{aligned}$$

Lastly, insert those expression into the condition for  $\bar{\varphi}$  and solve the expression for  $\bar{\varphi}$ :

$$\bar{\varphi} = \frac{\tau_j}{w} \left[ p \omega \left( \frac{\alpha}{r} \right)^\alpha \left( \frac{\beta(1 - \alpha - \beta)}{v} \right)^{1-\alpha-\beta} \right]^{\frac{1}{\beta}}$$

$$\text{Define } \phi := \left[ p \omega \left( \frac{\alpha}{r} \right)^\alpha \beta^{1-\alpha-\beta} \left( \frac{1-\alpha-\beta}{v} \right)^{1-\alpha-\beta} \right]^{\frac{1}{\beta}}$$

### A.2.5 Urban productivity shocks and the organization of agriculture

The response of regional agricultural production is determined by

$$\begin{aligned}
\frac{dY_{jk}}{d\tau} &= \sum_i \frac{dy_{ijk}}{d\tau} \\
&= \sum_{i=1}^{n_j} \alpha a_{ijk}^{\alpha-1} l_{ijk}^\beta \theta^{1-\alpha-\beta} \frac{da_{ijk}}{d\tau} + \sum_{i=1}^{n_j} \beta a_{ijk}^\alpha l_{ijk}^{\beta-1} \theta^{1-\alpha-\beta} \frac{dl_{ijk}}{d\tau} + \sum_{i=1}^{n_j} (1-\alpha-\beta) a_{ijk}^\alpha l_{ijk}^\beta \theta^{-\alpha-\beta} \frac{d\theta_{ijk}}{d\tau} \\
&= \alpha \sum_i \frac{da_{ijk}}{d\tau} \frac{y_{ijk}}{a_{ijk}} + \beta \sum_i \frac{dl_{ijk}}{d\tau} \frac{y_{ijk}}{l_{ijk}} + (1-\alpha-\beta) \sum_i \frac{d\theta_{ijk}}{d\tau} \frac{y_{ijk}}{\theta_{ijk}}
\end{aligned}$$

Next, totally differentiate the first order condition with respect to  $\tau$  and rearrange terms to express changes in land and technology as a function of labor changes. Here we use subscripts to denote partial derivatives and omit the individual, location and crop indices:

$$\begin{aligned}
\alpha \frac{a_\tau}{a} &= \alpha \frac{l_\tau}{l} + \frac{\phi l(\alpha + \beta)}{\tau y \beta} - \frac{p_\tau}{p} \\
(1 - \alpha - \beta) \frac{\theta_\tau}{\theta} &= \frac{r_\tau a}{\alpha y} - \frac{p_\tau}{p \alpha} + (1 - \alpha - \beta) \frac{l_\tau}{l} + \frac{\phi l(\alpha + \beta)(\alpha - 1)}{\alpha \tau y \beta}
\end{aligned}$$

We then insert these expressions in the equation for the individual changes in agricultural production to arrive at

$$\begin{aligned}
y_\tau &= \left[ \alpha \frac{l_\tau}{l} + \frac{\phi l(\alpha + \beta)}{\tau y \beta} - \frac{p_\tau}{p} \right] y + \frac{l_\tau}{l} \beta y + \left[ \frac{r_\tau a}{\alpha y} - \frac{p_\tau}{p \alpha} + (1 - \alpha - \beta) \frac{l_\tau}{l} + \frac{\phi l(\alpha + \beta)(\alpha - 1)}{\alpha \tau y \beta} \right] y \\
&= l_\tau \phi_1 + p_\tau \phi_2 + r_\tau \phi_3 + \phi_4
\end{aligned}$$

with  $\phi_1 := \frac{y}{l}$ ,  $\phi_2 := -\frac{y}{p} \frac{(1+\alpha)}{\alpha}$ ,  $\phi_3 := \frac{y a}{\alpha y}$  and  $\phi_4 := \frac{\phi l(\alpha + \beta)}{\tau y \beta \alpha} (\alpha + y \alpha - y)$ .

## B Appendix Tables

Table A1: Migrant Type by Age Cohort

	Student	Employed Son	Husband
0-4	0.69	0.09	0.06
5-9	0.96	0.01	0.00
10-14	0.91	0.05	0.00
15-19	0.61	0.29	0.02
20-24	0.33	0.44	0.10
25-29	0.07	0.44	0.25
30-34	0.01	0.37	0.33
35-39	0.01	0.32	0.36
40-44	0.00	0.30	0.40
45-49	0.00	0.22	0.46
50+	0.00	0.08	0.29
Total	0.27	0.30	0.20

Note: Data from IHDS wave 1 (2004-05). Each row denotes an age cohort. Values denote the share of migrants in each age group belonging to each migrant type.

Table A2: OLS—Impact of Labour Migration on Agricultural Development

	Capital (index)		Land (ac.)	Profits (Rs.)	Labour	
	(1)	(2)	(3)	(4)	(5)	(6)
	Investments	Stock	Cultivated	Crops	Wage Bill	Person-days
Male Migrants	0.039*	0.020	0.032**	0.020	0.009**	-0.002
	(0.022)	(0.015)	(0.013)	(0.015)	(0.005)	(0.018)
Outcome SD	1.102	1.037	4.376	30186.600	542.390	342.733
HH FEs	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓
Observations	26336	25362	29238	25588	21282	21282
$R^2$	0.687	0.734	0.749	0.705	0.503	0.632

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . The explanatory variable is the number of working age male migrants in the household. All outcomes are in standard deviations. Columns 1 and 2 are indices for technology expenses and stock, respectively (see Section 4.2 for details of index construction). Crop profits (column 4) are net of expenses. Wage bill (column 5) is total wages paid to all workers in the past year. Person-days (column 6) includes both household and hired labour. Standard errors clustered at the PSU level.



Table A3: OLS—Impact of Labour Migration on Technology Adoption

	Investment Expenses (Rs.)					Stock (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants	0.042** (0.018)	0.032*** (0.011)	0.014 (0.016)	0.036 (0.038)	0.027 (0.022)	0.018 (0.017)	0.029* (0.016)	0.028* (0.016)	0.034** (0.016)	-0.018 (0.019)
Outcome SD	0.035	0.027	0.030	0.019	0.031	0.045	0.051	0.064	0.041	0.024
HH FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	26240	25850	24454	23886	25002	25076	24224	24402	24242	24106
R <sup>2</sup>	0.703	0.647	0.728	0.551	0.634	0.677	0.690	0.688	0.742	0.667

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . The explanatory variable is the number of working age male migrants in the household. All 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. Standard errors clustered at the PSU level.

Table A4: Second Stage—Impact of Migration on Technology Adoption

	Investment Expenses (Rs.)					Stock (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants	-1.214*** (0.387)	-1.653*** (0.393)	-1.249** (0.633)	-0.701* (0.381)	-1.980*** (0.566)	-0.988** (0.398)	-1.246** (0.485)	-2.139*** (0.480)	-0.953 (0.643)	-1.163*** (0.392)
Wt. Income	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Income	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	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	25852	25454	24054	23496	24602	24692	23862	24042	23882	23746
F-Stat	63.0	60.9	57.8	56.6	59.5	55.8	55.3	53.7	53.0	52.6

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. 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. Standard errors clustered at the PSU level.

Table A5: Partial and General Equilibrium Effects (Nationwide Price Channel)

	(1) Profits (Rs.)	(2) Farm Size (ac.)	(3) Investment	(4) Assets
Male Migrants (direct labour channel)	-1.666*** (0.381)	-1.156*** (0.292)	-1.055*** (0.267)	-0.864*** (0.232)
Village emigration (indirect land channel)	0.308*** (0.085)	0.240*** (0.060)	0.212*** (0.055)	0.194*** (0.049)
Other-HH emigration (indirect crop channel)	0.135 (0.237)	-0.061 (0.130)	-0.255* (0.137)	-0.277** (0.116)
Wt. Income	Yes	Yes	Yes	Yes
Origin Income	Yes	Yes	Yes	Yes
HH FEs	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓
N	20150	26524	25662	24632
F-Stat on Migrants	50.7	57.3	67.8	60.9

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . *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. *Other-HH emigration* is the leave-one-out number of migrants across India weighted by crop similarity. All variables are standardized. Crop profits are net of expenses (Column 1). Column 2 is cultivated area. Columns 3 and 4 are indices for technology expenses and stock, respectively (see Section 4.2 for details of index construction). Standard errors clustered by PSU.

Table A6: Partial and General Equilibrium Effects on Technology (State Price Channel)

	Investment Expenses (Rs.)					Assets (Num. Owned)				
	(1) Seeds	(2) Fertilizer	(3) Pesticide	(4) Water	(5) Rentals	(6) Tubewell	(7) Pumps	(8) Bullcart	(9) Tractor	(10) Thresher
Male Migrants (direct labour channel)	-0.713*** (0.228)	-1.019*** (0.239)	-1.043*** (0.401)	-0.500** (0.225)	-1.218*** (0.332)	-0.688*** (0.242)	-0.841*** (0.316)	-1.128*** (0.267)	-0.525 (0.360)	-0.646*** (0.227)
Village emigration (indirect land channel)	0.162*** (0.045)	0.194*** (0.047)	0.198** (0.086)	0.092** (0.045)	0.240*** (0.069)	0.197*** (0.053)	0.173*** (0.064)	0.244*** (0.060)	0.123 (0.078)	0.124** (0.051)
Other-HH emigration (indirect crop channel)	0.163*** (0.046)	0.198*** (0.053)	0.280*** (0.081)	0.155*** (0.047)	0.190*** (0.064)	-0.078 (0.051)	0.123* (0.067)	0.076 (0.049)	0.081 (0.059)	0.016 (0.043)
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 Migrants	61.6	59.4	55.1	53.0	57.5	54.2	53.1	52.4	50.9	50.7

Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . *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. *Other-HH emigration* is the leave-one-out number of migrants in the state weighted by crop similarity. All variables are standardized. 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. Standard errors clustered at the PSU level.

## C Appendix Figures

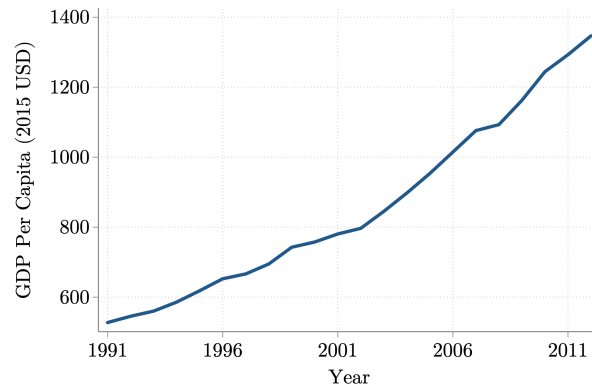


Figure A1: India Per-capita GDP Growth

Note: Indian GDP in 2015 prices divided by mid-year population. Data accessed from World Bank Open Data Portal.

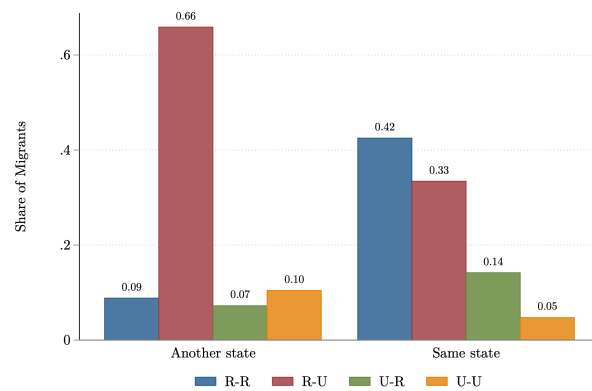


Figure A2: Migration Streams by Distance

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

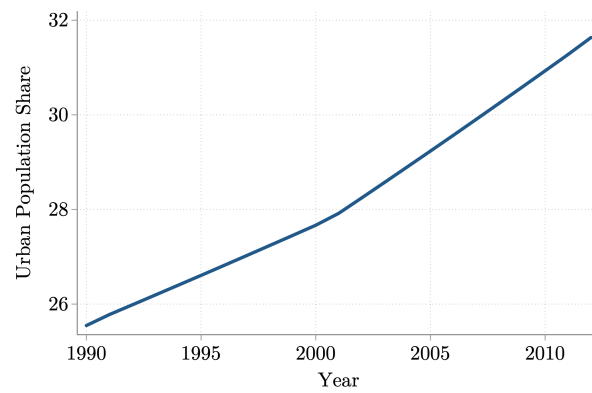


Figure A3: Urban Population Share (1990-2010)

Data source: United Nations Population Division. World Urbanization Prospects: 2018 Revision. Accessed through World Bank Open Data