NBER WORKING PAPER SERIES

MELITZ MEETS LEWIS:
THE IMPACTS OF ROADS ON STRUCTURAL TRANSFORMATION AND BUSINESSES

Joseph P. Kaboski
Jianyu Lu
Wei Qian
Lixia Ren

Working Paper 32448
http://www.nber.org/papers/w32448

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2024

We have benefited from comments from Tasso Adamapoulos, Kevin Donovan, Chaoran Chen, and Daniel Xu. We thank CEPR's Structural Transformation and Economic Growth (STEG) Research programme for supporting this research. On behalf of all authors, the corresponding author states that there is no conflict of interest. The views expressed are those of the authors and do not necessarily represent the views of the Central Bank of Chile or its board members. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32448

Melitz Meets Lewis: The Impacts of Roads on Structural Transformation and Businesses
Joseph P. Kaboski, Jianyu Lu, Wei Qian, and Lixia Ren
NBER Working Paper No. 32448
May 2024
JEL No. F15,O13,O18,O41

ABSTRACT

This paper examines the impact of roads on structural transformation and business composition theoretically and empirically. We develop a two-sector model of regional trade with endogenous firm entry that highlights two opposing forces. *Ceteris paribus* lower trade costs in non-agriculture lead to fewer firms, but cheaper agricultural imports releases labor from local agricultural production leading to more firms. Using major highway programs in India and China, we find results broadly consistent with the theory, with declines in the number of businesses where structural transformation is weak, and increases where it is strong.

Joseph P. Kaboski
Department of Economics
University of Notre Dame
3039 Nanovic Hall
Notre Dame, IN 46556
and CEPR
and also NBER
jkaboski@nd.edu

Jianyu Lu
Agustinas 1180
Banco Central de Chile
Santiago, RM
Chile
willjianyu@gmail.com

Wei Qian
Haverford College
370 Lancaster Ave Chase Hall
Haverford, PA 19041
wqian0901@gmail.com

Lixia Ren
Kellogg Institute for International Studies
University of Notre Dame
lren2@nd.edu

A data appendix is available at http://www.nber.org/data-appendix/w32448

1 Introduction

Two common problems for development are the large role of a low productivity agricultural sector (e.g., Gollin et al., 2014), especially subsistence farming among people in the hinterlands, and poor road infrastructure linking these hinterlands to markets (e.g., Asher and Novosad, 2020). Indeed, the two issues are deeply interrelated: specialization in subsistence production, regardless of one's productivity, can be a necessity for those not integrated into markets. Roads investments are quite costly, so the stakes are high, and public infrastructure investments are generally made with the hope that they will spur corresponding private investment. However, the links between road investment and structural transformation are not well understood. In particular, how do roads impact businesses in newly connected areas and how does this relate to (inter-regional) trade and structural transformation? While trade can open regions to new markets, it can also bring in competition, and closing firms can have distributional impacts. Indeed, to the extent that roads lower trade costs, standard trade theory (e.g., Krugman, 1980; Melitz, 2003) predict a decrease in the number of businesses. Yet, given the food problem, the structural transformation literature insists that roads are essential promoting business activity outside of agriculture (e.g., Gollin and Rogerson, 2014). While both literature view these opposing effects as indicative of well functioning economies benefiting from lower trade costs, structural transformation has clear consequences for the distribution of income across factors. How might we reconcile these two ideas, both theoretically and empirically?

This paper does exactly that. Integrating these two literature, we develop a two-sector model of regional trade with endogenous firm entry and subsistence requirements from agriculture. The model highlights two forces: i) a channel in which lower trade costs leads to increased demand to export, expansion of firms and in equilibrium fewer firms in non-agriculture (the "Melitz channel") and ii) a channel in which access to cheaper agricultural imports releases labor from local agricultural production and leads to more non-agricultural firms in equilibrium (the "Lewis channel"). The larger the initial share of agriculture, the stronger is the second channel, and the channels have distributional consequences as wages increase through the first channel, but labor's share declines via the second. We then evaluate this empirically, using the construction of major highway systems in India and China, to evaluate these channels. The number of firms contracts in India but not in China. We find, however, that the theory can help reconcile these very different findings for India and China: Structural transformation is stronger in China, and the growth of firms is stronger in regions where agriculture was more prominent.

The model we develops mingles standard approaches in trade and structural transformation. We view regions as small, open, two-sector economies. Specifically, we embed a Melitz (2003) model, with free entry and heterogeneous, differentiated producers of non-agricultural goods, and fixed costs of selling outside of the region, into a typical two-sector model with a subsistence requirement for agricultural goods but with trade in agriculture. With the price of agriculture in other regions fixed for the small open economy, the share of labor in the non-agriculture is high when agricultural productivity is high and/or trade costs are low. The model formally demonstrates the two channels above and shows that in the limiting cases when the non-agricultural (agriculture) sector uses most resources, the first (second) channel dominates and the number of firms declines (increases) as trade costs fall.

We evaluate these predictions using two of the largest highway expansions in the world as empirical changes in trade costs: the Golden Quadrilateral (GQ), India's expressway expansion initiative connecting Delhi, Calcutta, Chennei, and Mumbai, and China's National Trunk Highway System (NTHS) expansion. These are two recent expansions in the two largest developing economies in the world. For India, we combine firm and labor data from the Annual Survey of Industries and National Sample Surveys, while for China we utilize data from Provincial Statistical Yearbooks. All of these are available at local levels which we combine with geospatial data on road location. Empirically, we use a difference-in-difference approach, treating both countries as symmetrically as possible. Dividing available regions into terciles, and using those closest to the road, we interact the spatial variation with time variation, and using 80% completion to identify post-road years.

We find that the number of establishments drops markedly (10 to 14 log points) in regions closer to new roads in India, while they increase (by roughly 14 log points) in China. There is suggestive evidence that the average number of employees per firm increases in both countries, reflecting some role for the Melitz channel. Consistent with the theory, where

Melitz channels dominate in India but Lewis mechanisms in China, we see a strong drop in agricultural employment in China (roughly 32 log points), but no statistical change in India. Finally, dividing regions into terciles by their initial share of the labor force in agriculture, we find that within country variation is also consistent with the theory. Lewis channels are stronger in high agriculture regions, leading to relatively more businesses in response (e.g., 24 log points more in China), whereas Melitz channels are stronger in low agriculture regions, leading to relatively fewer businesses in response (e.g., 18 log points less in India).

Related Literature

We contribute to and integrate multiple literature.

Foremost, we contribute to the macro development literature on the structural transformation out of agriculture and its impact on the rest of the economy. The classic citation, Lewis (1954), posited an effectively unlimited supply of labor to feed industrial development. His ideas still impact current research (e.g., Storesletten et al., 2019), and a key theme in the larger literature is the importance of moving out of subsistence agriculture — a stagnant and low-productivity sector with too much labor — for growth (Matsuyama, 1991, 1992a,b; Caselli and Coleman II, 2001; Gollin et al., 2007; Restuccia et al., 2008; Vollrath, 2009; Lagakos and Waugh, 2013; Gollin et al., 2014; Donovan, 2021). The role that road infrastructure can play in alleviating potential misallocation has also been quantified as substantial (Gollin and Rogerson, 2014; Van Leemput, 2021). Typically these models start with agriculture as a necessity good. Wedding such a model to a classic trade model, we show the implications for establishments of structural transformation out of agriculture.[1] Firms can be indicators of job creation, dynamism, and private investment, all of which are often anticipated as private complements to public investment in infrastructure. In our model, declines in trade costs are always welfare-enhancing whether the number of firms rises or falls. However, the distribution of income across wages and profits is impacted by structural transformation.

[1] Somewhat analogously, Gollin et al. (2002) wed a Solow model non-agricultural sector to an agricultural sector with subsistence preferences and demonstrate how the movement of labor out agriculture can lead to slow convergence.

We therefore also relate to strands of the trade literature. One strand is quantitative work emphasizing the role of internal trade costs and highway infrastructure (e.g., Allen and Arkolakis, 2014, 2022). A second emphasizes an important role for international trade in structural transformation (e.g., Matsuyama, 2009; Uy et al., 2013), respectively. Most closely, however, our paper contributes to a trade literature on the impact of trade costs on the distribution of firms (e.g., Bernard et al., 2003; Krugman, 1980; Melitz, 2003). The empirical work in the past 25 years has largely focused on the heterogeneous behavior of firms in their production and exporting firms, including the exit of low productivity firms in response to trade liberalizations (e.g., Pavcnik, 2002). We augment firm-level trade models with an agricultural sector, and show that outflow from agriculture is important to understand the varied responses of the size and number of businesses in regions of India and China to increased trade possibilities from highway expansion.

Lastly, a growing applied micro literature studies the impacts of roads on rural development.[2] In particular, India's GQ has been evaluated for its impacts on various outcomes, including product market competition (Asturias et al., 2019) and labor market power (Brooks et al., 2021). Most relevant to us in India, Ghani et al. (2016) find that the GQ spurs business activity. This may appear to contradict our findings, but doesn't: given their emphasis on location choice, they focus on close proximity to the roads, only examining states and districts with high existing manufacturing activity. Given our interest in regional development and structural transformation, we examine larger bands around roads, include more rural areas, and exclude larger regions around the targeted large urban areas. Also, Grover Goswami et al. (2024) find young firms expand but old and low productivity firms exit in response to GQ connection, consistent with our results. For China, Faber (2014) and Banerjee et al. (2020) have both evaluated the impact of NTHS on connected regions with somewhat different conclusions. Most relevant to us is Lu (2023) who finds positive impacts of the highway on firm growth, productivity, and input usage. Our connecting the impacts on firms to structural transformation out of agriculture is unique.

[2]Other transportation investments have also been examined recently, including colonial railroads in India (Donaldson, 2018), bridge construction in Nicaragua Brooks and Donovan (2020), and community boat investments in the Amazon (Bartkus et al., 2022).

2 Model

We model a within-country region as a small, open economy with two sectors, A (agriculture) and N (non-agriculture). Demand in the agricultural sector follows Armington (1969), aggregating local and traded varieties produced competitively, while the non-agricultural sector follows Melitz (2003) with a variety of differentiated non-agricultural goods. A measure $\bar{L}$ of households are endowed with one unit of labor and can freely choose to supply labor to either sector. The small, open economy trades with the "rest of the world", which includes the other regions in the country as well, all aggregated into a single supplier.

2.1 Preferences

The preferences of a representative household are described by the following function:

$$U = \beta \log (Q_A - \bar{a}) + (1 - \beta) \log Q_N$$

where Q_A and Q_N are the consumption bundle of agriculture and non-agriculture goods, $\bar{a}$ is the subsistence level of agricultural goods, and β is the preference weight on agricultural goods. We further assume that $\beta \to 0$, so that, although the the agricultural good is a necessity good, the household places no value on any consumption beyond the necessity requirement, $\bar{a}$.

Both Q_A and Q_N are aggregates of varieties. Namely, agricultural consumption follows an Armington (1969) formulation with a symmetric constant elasticity of substitution (CES) aggregator over agricultural goods produced in the local region q_A and in the rest of the world, $q_{A_{ROW}}$:

$$Q_A = \left(q_A^{\frac{\gamma-1}{\gamma}} + \left(\frac{q_{A_{ROW}}}{\tau} \right)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} \tag{1}$$

where τ is the iceberg trade cost that occurs for interregional trade. That is, $\tau > 1$ units must be purchased in order to receive and consume one unit of the goods from the rest of the world.

Non-agricultural consumption is described by a CES function over a continuum of locally-produced differentiated goods indexed by ω:

$$Q_N = \left[\int_{\omega \in \Omega} q_N(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where σ is the elasticity of substitution between any two goods with $\sigma > 1$.

For simplicity, we assume that the region demands agricultural goods from the rest of the world but does not demand non-agricultural good imports. This simplifies the formulas and notation without impacting the qualitative results (as long as import demand were not large).

2.2 Technology

Aggregate agricultural goods production is constant returns to scale in labor:[3]

$$q_A = TL_A$$

where T is the agricultural productivity and L_A is the amount of labor used in agriculture.

Inside the region exists an endogenous continuum of firms, each producing a distinct variety ω with only one input, labor. Production requires paying a fixed cost of f units of labor and then firms produce with marginal productivity, φ. The total amount of labor l required to produce an output q is therefore:

$$l(\varphi) = f + \frac{q(\varphi)}{\varphi}$$

An infinite supply of potential entrant firms exists, but entry requires paying a fixed entry cost f_e to discover one's productivity parameter, φ, which is drawn from a Pareto distribution $G(\varphi)$:

$$G(\varphi) = 1 - \varphi^{-\kappa}$$

with with shape parameter $\kappa > \sigma - 1$ as in Chaney (2008).

[3]Given the importance of land for agricultural production, the assumption of diminishing returns to scale production in labor is more realistic, but it complicates the algebra without adding additional insights of note.

Entrants can sell locally, but there is an additional cost, f_{ROW}, to sell in the rest of the world. After paying this f_{ROW}, each firm faces a symmetric downward-sloping, constant-elasticity demand for their variety in the rest of the world[4]:

$$q_N^{ROW}(\omega) = Bp_N^{ROW}(\omega)^{-\sigma} \tag{3}$$

Again, for simplicity we omit demand for the local agricultural variety from the rest of the world. As long as this demand were small, e.g., either through preferences or low agricultural productivity T, this simplifies formulas and notation without impacting the qualitative results.

2.3 Equilibrium and the Impacts of Trade Costs

We present the optimization of the household, farm, and firms, respectively, and then apply market clearing before summarizing comparative statics with respect to iceberg trade costs.

Household's Problem

The household takes prices p_A, $p_{A_{ROW}}$, $p_N(\omega)$, $p_{N_{ROW}}$ and w as given. We assign the local agricultural good as the numeraire, $p_A = 1$. The households' problem is:

$$\max_{\substack{q_A, q_{A_{ROW}}, q_N(\omega), \\ L_A, l_n(\omega)}} \beta \log(Q_A - \bar{a}) + (1-\beta)\log Q_N$$

s.t.

$$p_A q_A + p_{A_{ROW}} q_{A_{ROW}} + \int_{\omega\in\Omega} p_N(\omega) q_N(\omega) \leq w_A L_A + \int_{\omega\in\Omega} w_N(\omega) l_n(\omega),$$

$$L_A + \int_{\omega\in\Omega} l_n(\omega) = \bar{L},$$

and aggregators (1) and (2).

The household's problem leads to several results. First, the $\beta \to 0$ assumption implies

[4] For prices and quantities, we use superscript to denote the region of consumption and subscript to denote the region of production.

that $Q_A = \bar{a}$, and the first-order conditions (FOCs) with respect to q_A and $q_{A_{ROW}}$ imply that the optimal consumption of local agricultural goods is:

$$q_A = \frac{\bar{a}}{\left[1 + \left(\frac{1}{\tau}\frac{1}{p_{A_{ROW}}}\right)^{\gamma-1}\right]^{\frac{\gamma}{\gamma-1}}}$$

Second, optimal choices of non-agricultural varieties imply the well-known demand functions for each variety:

$$q_N(\omega) = \left[\frac{p_N(\omega)}{P}\right]^{-\sigma} \frac{X}{P} \tag{4}$$

where X denote the total expenditure on non-agricultural goods, and P is the aggregate price index:

$$P = \left[\int_{\omega \in \Omega} p_N(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$$

where $p_{N_{ROW}}$ is the exogenous price of imported non-agricultural goods.

Third, FOCs with respect to labor supply require that any interior labor supply decisions earn the same wage: $w_N(\omega) = w_A = w$.

Agricultural Problem

Since locally produced agriculture is the numeraire, the representative competitive farm takes its price and the wage as given and solves:

$$\max_{L_A} p_A q_T(L_A) - wL_A$$

The optimality condition determines the wage, $w = T$.

Firm's Problem

Upon realizing its productivity, each firm ω decides whether to produce, and whether to pay the fixed cost to sell to the export market, f_{ROW}. Small relative to the continuum of producers, it takes the wage, w, and overall price level, P, as given, but it sets its own price in each market, p_N^{ROW}.

A firm that has paid the entry cost f_e and drawn productivity φ solves the following problem:

$$\max\{0, \max_{\{p_N(\varphi), p_N^{ROW}(\varphi)\}} \{\pi_d(\varphi) - wf + \max\{0, \pi_{ROW}(\varphi) - wf_{ROW}\}\}\} \tag{5}$$

The three maximizations in the firm's problem (5) capture the firm's various decisions. The first maximization captures the decision to not produce or pay the fixed cost, f, and earn domestic profits $\pi_d(\varphi)$, the second maximization sets prices optimally, and the third maximization is the decision of whether to only sell locally or pay the exporting fixed cost, f_{ROW}, and earn additional profits, $\pi_{ROW}(\varphi)$, from selling to the rest of the world. These profits are:

$$\pi_d(\varphi) = \max_{p_N} p_N q_N - \frac{w}{\varphi} q_N \tag{6}$$

and

$$\pi_{ROW}(\varphi) = \max_{p_N^{ROW}} p_N^{ROW} q_N^{ROW} - \frac{\tau w}{\varphi} q_N^{ROW}, \tag{7}$$

respectively. Here the latter expression incorporates the iceberg transportation cost, i.e., the firm must send τ units in order for one to arrive, so its marginal cost is $\tau w/\varphi$.

Substituting local and external demand using equations (4) and (3) into the profit functions, the optimal pricing rules are then:

$$p_N(\varphi) = \frac{\sigma}{\sigma - 1}\frac{w}{\varphi} \qquad p_N^{ROW}(\varphi) = \frac{\sigma}{\sigma - 1}\frac{\tau w}{\varphi}$$

Firms charge a constant markup $\sigma/(\sigma - 1)$ over marginal cost in both markets, but the trade cost τ increases the effective marginal cost of supplying to the external market.

The well-known result of this problem is that both local and external profits are increasing in φ and the fixed costs to produce and export define two thresholds, φ^* and φ^*_{ROW}, above which the firm produces and exports, respectively:

$$\pi_d(\varphi^*) = wf \qquad \pi_{ROW}(\varphi^*_{ROW}) = wf_{ROW} \tag{8}$$

Market Clearing

We now apply market clearing conditions. Given the demand for the local agricultural good and the agricultural production function, equilibrium employment in the agriculture sector is:

$$L_A = \frac{q_A}{T}\bar{L} = \frac{\frac{\bar{a}}{T}\bar{L}}{\left[1 + \left(\frac{1}{\tau}\frac{1}{p_{A_{ROW}}}\right)^{\gamma-1}\right]^{\frac{\gamma}{\gamma-1}}} \tag{9}$$

L_A depends on the price ratio $\frac{p_A}{p_{A_{ROW}}}$, iceberg trade cost τ, the elasticity of substitution between agricultural goods produced locally and non-locally γ, and the ratio $\frac{\bar{a}}{T}$, which measures how many labors it takes to feed one household.

The necessary assumption that agricultural productivity, T is high enough to produce the subsistence level of agricultural consumption, $\bar{a}$ now becomes clear:

Assumption 1

$$T \geq \underline{T} \equiv \frac{\bar{a}}{\left[1 + \left(\frac{1}{\tau}\frac{1}{p_{A_{ROW}}}\right)^{\gamma-1}\right]^{\frac{\gamma}{\gamma-1}}}.$$

Market clearing in the labor market therefore dictates that $L_N = \bar{L} - L_A$. Market clearing for L_N is dictated by the wage, fixed at T, and the number of firms, M. In equilibrium, the number of firms is determined by the free entry condition: a potential entrant's expected profit after entry is equal to the entry cost f_e:

$$\bar{\pi} = \left(\int_{\varphi^*}^{\infty} (\pi_d(\varphi) - wf)\, d\varphi + Prob_{ROW} \int_{\varphi^*_{ROW}}^{\infty} (\pi_{ROW}(\varphi) - wf_{ROW})\, d\varphi\right) = \frac{wf_e}{1 - G(\varphi^*)} \tag{10}$$

where $Prob_{ROW} = \frac{1-G(\varphi^*_{ROW})}{1-G(\varphi^*)}$ is the probability a successful firm exports.

Combined with the zero profit conditions in Eq.(8) and the Pareto distribution $G(\varphi)$, the mass of firms in equilibrium is given by:

$$M = \frac{wL_N}{\sigma(\bar{\pi} + wf + wProb_{ROW} f_{ROW})} = \frac{(\kappa - (\sigma - 1))L_N}{\sigma\kappa\left(f + f_{ROW}\tau^{-\kappa}\left(\frac{f_{ROW}}{f}\frac{X}{BP^{\sigma+1}}\right)^{-\frac{\kappa}{\sigma-1}}\right)} \tag{11}$$

Changes in Trade Costs

As trade costs τ fall two things change. First, the effective cost of imported agricultural goods falls, and the household substitutes toward $q_{A_{ROW}}$. Consequently q_A falls, lowering the amount of labor needed to produce the regional agricultural good in order to meet subsistence. Second, the marginal cost of selling regional non-agricultural varieties to other regions falls, so each firm's labor demand increases. These forces act in opposite directions on the number of firms in equilibrium. We now summarize the channels of how trade costs impact the equilibrium number of firms.

Proposition 1 *Given Assumption 1, the model implies:*

1. *The amount of non-agricultural labor rises as trade costs fall or, in parallel fashion, as the price of the agricultural good in the rest-of-the-world falls, i.e.,* $\partial L_N/\partial \ln \tau = \partial L_N/\partial \ln p_{A_{ROW}} < 0$.

2. *Melitz Channel: As trade costs τ fall, for a fixed non-agricultural labor (i.e., increasing $p_{A_{ROW}}$ to keep $\tau p_{A_{ROW}}$ and L_N constant as τ falls), the number of firms, M, falls (and the average firm size increases).*

3. *Lewis Channel: For a fixed τ, as non-agricultural labor rises (i.e., via a decrease in $\tau p_{A_{ROW}}$), the number of firms, M, rises (and the average firm size declines).*

4. *For any finite value of trade costs, there exists a threshold value of agricultural labor (and corresponding T) above which the number of firms increases as trade costs fall, and below which the number of firms decreases as trade cost fall.*

The first result captures the trade forces (in addition to the obvious force of technology, T) behind structural transformation in the model. The second result, the Melitz channel, is well understood and the status quo for thinking about the impact of trade cost reductions on the number of firms. Indeed, the result does not require the heterogeneity of Melitz (2003) but is already present in Krugman (1980). We call the third result the Lewis result, after Lewis (1954), who posited that "capitalist" development proceeded from an elastic supply of labor coming from a "subsistence" sector (largely agricultural) at a fixed wage. The influx of labor from agriculture leads to an increase in the number of non-agricultural firms. The final result demonstrates that this Lewis channel dominates when agriculture plays an important role in the economy.

The key intuition for this last result comes from the expression of $\partial M/\partial\tau < 0$ which is proportional to the sum of a negative and positive term:

$$\frac{\partial M}{\partial \tau} \propto -\frac{\gamma\tau^{-\gamma}\frac{\bar{a}}{T}(\frac{1}{p_{A_{ROW}}})^{\gamma-1}}{\left(1+(\frac{1}{\tau}\frac{1}{p_{A_{ROW}}})^{\gamma-1}\right)^{\frac{2\gamma-1}{\gamma-1}}}\left(f+\tau^{-\kappa}\left(\frac{f_{ROW}}{f}\frac{X}{BP^{\sigma-1}}\right)^{-\frac{\kappa}{\sigma-1}}f_{ROW}\right)$$

$$+\kappa\tau^{-\kappa-1}\left(\frac{f_{ROW}}{f}\frac{X}{BP^{\sigma-1}}\right)^{-\frac{\kappa}{\sigma-1}}f_{ROW}\left(1-\frac{\bar{a}}{T\left(1+\left(\frac{1}{\tau}\frac{1}{p_{A_{ROW}}}\right)^{\gamma-1}\right)^{\frac{\gamma}{\gamma-1}}}\right)$$

The top-term is negative, reflecting the Lewis channel, while the bottom term is positive, reflecting the Melitz channel, leaving an ambiguous sign. However, by inspection, clearly $\partial^2 M/\partial\tau\partial T > 0$. Moreover, as $T \to \infty$, no labor is in the agricultural sector, so the Lewis channel is non-existent (the top term is zero), while when T is low enough that all labor is in the agricultural sector, the Lewis force dominates (the bracketed part of the bottom term is zero). See Online Appendix for formal proofs.

The remaining paper evaluates the predictions of Proposition 1 using major highway infrastructure investments as an exogenous decrease in trade costs.

3 Policy Background and Data

This section describes the policy background of the GQ in India and NTHS in China, major highway construction projects in one of the largest and most populous countries of the world, and our data sources.

3.1 The Golden Quadrilateral Initiative

The GQ Initiative is part of India's National Highways Development Project (NHDP) launched in 2001. The name was chosen because it connects four largest metropolitan areas across India: Delhi in the north, Calcutta in the east, Chennai in the south, and Mumbai in the west in a circuit. The main objective was to reduce the traveling times between major cities in India. The Quadrilateral spans over 5,800km and is ranked the fifth-longest highway in the world.

The construction of the Quadrilateral started in 2001 and approximately 95% of the total project was completed by 2006. The project costed the government 250 billion rupees ($5.3 billion in 2001), which constituted about 1 percent of India's GDP. Figure B.1 in the appendix demonstrates Golden Quadrilateral quick completion by displaying the extent of the network in 2001 and 2006.

The Golden Quadrilateral improved road conditions as compared to the previously existing road networks, replacing mostly single-lane and two-lane highways with four and six-lane roads. The Quadrilateral also comes with additional safety features such as grade separators, over-bridges, and bypasses, and it indeed reduced travel times between connected areas significantly. For example, travel time between Gurgaon (near the Delhi-Haryana border) and Delhi was reduced from 60 minutes to approximately 20 minutes.

3.2 China's National Trunk Highway System

China's NTHS was originally planned in 1992 by the Chinese State Council, with the somewhat more ambitious goal of connecting *all* major cities in one single network. The initial network was known as "5-7" network, which refers to five vertical (north-south) and seven horizontal (east-west) routes. It had the objective of connecting all provincial capitals and cities with an urban registered population above 500,000 with the nation's capital Beijing. The original network had a total length of 35,000 km and was completed ahead of schedule. Therefore, in late 2004 the central government issued a revision to expanded the original plan. The revised plan, known as the "7-9-18" system, combined 7 radial expressways departing from the national capital of Beijing, 9 north-south expressways, and 18 east-west expressways. Figure B.2 in the appendix displays the network of NTHS in 2000 and 2011.

The NTHS is a highly improved system of expressways compared to previously existing road networks. The NTHS routes are limited access toll roads with at least four lanes (sometimes six lanes or even eight lanes). These expressways are superior in road condition and driving speed relative to any pre-existing national and provincial highways and also reached new areas. The expressways in NTHS have a maximum speed limit of 120km/h and a minimum of 70km/h, much faster than other highways and roads in China usually have a maximum speed limit of 70km/h. Therefore, the expressways of NTHS significantly reduce the travel time. For example, when the MeiHe Highway opened in 2006, the driving time between Guangzhou and Meizhou, a within province drive, decreased from 6 hours to 4 hours.

3.3 Data

We use three datasets for India. The first dataset is the panel version of the Annual Survey of Industries collected by India's Ministry of Statistics and Programme Implementation. These data are at the establishment level with accurate locations for production. The data include all plants with over 50 employees and a sample of smaller plants that depends on the industry and the number of plants within that industry and state. We focus on the manufacturing sector data from 2000 to 2011. The second dataset is the Employment and Unemployment Survey from the National Sample Survey (NSS) Organisation. We use six rounds of these data between 2000 and 2012.[5] The NSS data have been the primary source of data for estimating the aggregate employment trends in India, and we use individual household employment information to calculate the district-level agricultural employment. We linearly interpolate missing values for those years when data are unavailable. The last dataset is the road and geospatial data coming from Asturias et al. (2019). The dataset contains two kinds of distances to the GQ from each Indian district. One is the straight-line distance between the actual highway and the most populous city in the district. The other is the distance to a hypothetical highway layout that would have been chosen if GQ

[5]We use the 55th round, the 61th round, the 62th round, the 64th round, the 66th round, and the 68th round of the NSS data.

was established based on the minimum distance between the nodal areas.[6] We use both distances in our analysis.

We use two datasets for China. The first dataset is the prefecture-level city socioeconomic records are taken from Provincial Statistical Yearbooks from the years 2000 to 2011, obtained from the University of Michigan's China Data Center. The Provincial Statistical Yearbook series report prefecture-level city level variables such as output, local population, number of firms (revenues$\geq$RMB5 million), average number of employees per firm, and number of people working in agriculture. The second dataset is the road and geospatial data coming from Lu (2023), and the corresponding least-cost network. China has around 300 prefecture-level cities, and the highway data measures the shortest straight-line distance between the highway and geometric center within each prefecture.

We focus on India and China as two of the most important economies of the world. Importantly, our firm data for both countries are based on datasets that do not truncates coverage of small firms. Nevertheless, this limits the level of analysis that we can with firm heterogeneity. Specifically, we lack representative data on productivity in China to examine heterogeneous reallocation.

4 Empirical Analysis

In this section, we present our regression specification and empirical results. We use a difference-in-differences estimation, that is, we compare the change in the districts that are close to the highway (where trade costs presumably fell the most) to the change in the districts that remain further away from the highway (where trade costs were less impacted) before and after the highway connection. The regression specification is as follows:

$$ln(y_{it}) = \beta * \chi_{\text{tercile}=1,2} * \chi_{t>t_c} + \chi_i + \chi_t + \gamma Z_{it} + \varepsilon_{it}, \tag{12}$$

where $\chi_{\text{tercile}=1,2}$ is an indicator for whether the district is located in the closest or the middle terciles of locations from the highway. $\chi_{t>t_c}$ indicates whether the date is after the highway

[6]The GQ that was established based on the minimum distance between the nodal areas would simply be the straight-line connection of the four nodal areas: Delhi, Mumbai, Calcutta, and Chennai.

completion date t_c. We choose the completion year to be 2004 for India, when the GQ achieved at least 80% completion rate.[7] For China, we choose the completion year to be the one when at least 80% of NTHS segments planned within a province has been constructed. We include χ_i and χ_t location- and year-fixed effects. To isolate the effect of the highway, we control for several time-varying district characteristics (Z_{it}) that may correlate with the location of the highway, such as the (log) level of district output and working-age population.

Estimating equation (12) directly by OLS could suffer from endogeneity concerns, however. Provincial governments, responsible for raising credit and paying for its construction, may have placed roads in areas expected to grow faster in the future. To address this potential endogenous placement, we instrument the proximity to highway with a distance to a hypothetical highway layout that central planners would have chosen if construction cost were the only driving factor. For India, we assume that the least-cost highway is the straight-line connection of the GQ's four nodal areas. For China, we follow the strategy developed by Lu (2023) and construct a hypothetical least-cost path spanning tree network based on the global minimum construction cost to connect all the targeted cities in a single network.[8] Our instrument is the shortest straight-line distance between the least-cost counterfactual highway and geometric center within each prefecture.

Dropping the prefecture-level cities located within 50km of the targeted urban centers, we divide the remaining districts into terciles based on their distance to the least-cost highway. Figure 1 displayed the three terciles in different colors.

Table H.3 presents the effect of highway on the number of manufacturing establishments in India and China. Columns 1-3 show the IV results for India, and columns 4-6 show the IV results for China. The coefficient of our regressor of interest, $\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$ is significant and

[7] In Brooks et al. (2021), we use 2006 as the completion year for GQ. We change the completion year to 2004 because we want to make the analysis between China and India as symmetric as possible. Using 2006 as the completion year for GQ does not change our results qualitatively.

[8] The OLS regressions for India, which use distance to the actual highway, produce qualitatively the same results (see Online Appendix E). The OLS results for China, included in Online Appendix D, produce similar qualitative patterns to our IV with the exception of weakened significance (10% level) in agricultural employment declines (the analog to Table 3) and an increase in firm size (the analog to Table 2), indicating endogenous targeting of the road toward locations with growing firms. Event studies for both countries, including a staggered difference-in-differences approach for China, are presented in Online Appendices F and G, respectively. Finally, Online Appendix H shows that omission of small firms in the Chinese data is not a source of the differing results between India and China.

negative for India: the number of establishments reduces more after connection in districts that are close to the newly constructed highway than in districts that remain further away from the highway. In Columns 2 and 3, we include time-varying district characteristics, such as total output and working-age population (which could control for commuting or migration, respectively), in the regression.[9] The results are robust to these additional controls. In contrast, the NTHS in China appears to increase the number of establishments in city-prefectures. The lower panel shows the first-stage results for our IV, indicating that terciles based on the counterfactual NTHS are strong predictors of terciles based on the actual NTHS placements. The upper panel of Column 4 shows that, relative to the unconnected districts in the third tercile, connected districts experienced over 14 log points increase in the number of establishments after connection. The estimated effect of connection remains essentially unchanged with the inclusion of additional controls.

The results in Table 2 illustrate the impact of highways on the number of employees in an establishment. In India, establishments increase employment after districts become connected to the highway, by about 6% relative to the establishments in the unconnected districts after connection. In contrast, the impact on average employment in China is smaller and not significant. In sum, the standard Melitz channel of fewer but larger firms remaining as a result of a fall in trade costs seems consistent with what we observe in India, but not consistent with what we observe in China.

Melitz forces additionally predict that this happens by differential culling of small firms and an expansion of large firms. For China, we lack microdata to assess whether this holds, but for India we indeed find that the number of small firms fall disproportionately, while the number of large firms actually rises significantly (see Online Appendix C).

We also test the effect of highways on agricultural employment and the results are presented in Table 3. In India, the estimated coefficients on the regressor of interest, $\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$, are negative but not statistically significant. We cannot reject that the agricultural employment in India does not change after connection in areas that are close to the highway relative to the areas that remain further away from the highway. However,

[9] Brooks et al. (2021) argue that the GQ impacted labor market competition, while Banerjee et al. (2020) attribute tepid growth in response the NTHS to spatially immobile factors in China. We find output to be a significant control but not population.

in China, agricultural employment decreases significantly, by roughly 32 log points in locations connected by the highway. The Lewis channel appears to be strong in China but not in India. We now evaluate whether both forces may be at play in both countries. Recall our model predicts that given a large enough agricultural sector, the number of firms will increase in response to a lower transportation cost. We therefore investigate whether the effect of highways in India and China varies with the size of agricultural sector. We split the sample regions into terciles based on the share of people working in agriculture in 2000.[10] Table H.4 shows the effect of highway on the number of establishments using those subsamples. In both India and China, there is evidence that the effect of highway on the number of establishments becomes more positive (or less negative) as the size of agricultural sector increases. In India, the significant decrease in establishments is concentrated among districts with initially small agricultural sector (Columns 1 and 2). In China, the significant increase is concentrated in cities where the initial agricultural employment share is highest (Column 6).

5 Conclusion

We have demonstrated that trade costs and structural transformation can interact with each other in nuanced ways that impact the firm composition. In particular, while trade in the non-agricultural sector leads to fewer but larger firms, trade in the agricultural sector releases labor into non-agriculture allowing the number of firms to expand. Using the expansion the national highways in China and India, we evaluate these predictions, finding that firms decline in areas where the agricultural sector was smaller and where structural transformation was weaker. The theory can therefore reconcile otherwise seemingly conflicting results across regions.

Not all predictions line up neatly, however. Indeed, a standing question is why China experienced much stronger structural transformation out of agriculture near roads, while India did not, despite the fact that India is more agricultural overall. It may well be that

[10] We calculate the ratio of agricultural employment relative to the total population aged between 15 and 64. The results are robust to replacing working-age population with total employment in the calculation.

India's comparative advantage was indeed in agriculture. Although the model is stylized, introducing frictions (e.g., financial frictions) or externalities (as in Matsuyama, 1991, 1992c) might yield insights, and, vice versa, the empirical results might lead to insights into the frictions. All of these considerations may be important for thinking about the efficiency and distributional welfare implications of roads on firm activity. We leave these for future research.

References

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy ," *The Quarterly Journal of Economics*, 05 2014, *129* (3), 1085–1140.

_ **and** _ , "The Welfare Effects of Transportation Infrastructure Improvements," *The Review of Economic Studies*, 02 2022, *89* (6), 2911–2957.

Armington, Paul S., "A Theory of Demand for Products Distinguished by Place of Production (Une théorie de la demande de produits différenciés d'après leur origine) (Una teoría de la demanda de productos distinguiéndolos según el lugar de producción)," *Staff Papers (International Monetary Fund)*, 1969, *16* (1), 159–178.

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, 2020, *110* (3), 797–823.

Asturias, Jose, Manuel Garcia-Santana, and Roberto Ramos, "Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral in India," *Journal of the European Economic Association*, 2019, *17* (6), 1881–1940.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the road: Access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, 2020, *145*, 102442.

Bartkus, Viva Ona, Wyatt Brooks, Joseph P. Kaboski, and Carolyn Pelnik, "Big fish in thin markets: Competing with the middlemen to increase market access in the Amazon," *Journal of Development Economics*, 2022, *155*, 102757.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, September 2003, *93* (4), 1268–1290.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *The Review of Economic Studies*, 02 2024, p. rdae007.

Brooks, Wyatt and Kevin Donovan, "Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua," *Econometrica*, 2020, *88* (5), 1965–1997.

Brooks, Wyatt J., Joseph P. Kaboski, Illenin O. Kondo, Yao Amber Li, and Wei Qian, "Infrastructure Investment and Labor Monopsony Power," *IMF Economic Review*, September 2021, *69* (3), 470–504.

Caselli, Francesco and Wilbur John Coleman II, "The US structural transformation and regional convergence: A reinterpretation," *Journal of political Economy*, 2001, *109* (3), 584–616.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, September 2008, *98* (4), 1707–21.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, April 2018, *108* (4–5), 899–934.

Donovan, Kevin, "The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity," *Review of Economic Studies*, 2021, *88* (5), 2275–2307.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *The Review of Economic Studies*, 03 2014, *81* (3), 1046–1070.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr, "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing," *The Economic Journal*, 2016, *126* (591), 317–357.

Gollin, Douglas and Richard Rogerson, "Productivity, transport costs and subsistence agriculture," *Journal of Development Economics*, 2014, *107* (C), 38–48.

_ , David Lagakos, and Michael Waugh, "The Agricultural Productivity Gap," *Quarterly Journal of Economics*, 2014, *129* (2).

_ , Stephen L. Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review Papers and Proceedings*, 2002, *92* (2).

_ , _ , and _ , “The Food Problem and the Evolution of International Income Levels,” *Journal of Monetary Economics*, 2007, *54*, 1230–1255.

Goswami, Arti Grover, William Maloney, and Stephen D. O’Connell, *Which Firms Drive the Gains from Connectivity and Competition? The Impact of India’s Golden Quadrilateral across the Firm Life Cycle*, The World Bank, 2024.

Krugman, Paul, “Scale Economies, Product Differentiation, and the Pattern of Trade,” *American Economic Review*, 1980, *70* (5), 950–59.

Lagakos, David and Michael Waugh, “Selection, Agriculture, and Cross-Country Productivity Differences,” *American Economic Review*, 2013, *103* (2), 948–80.

Lewis, W Arthur, “Economic development with unlimited supplies of labour,” *The Manchester School*, 1954, *22* (2), 139–191.

Lu, Will Jianyu, “Transport, Infrastructure and Growth: Evidence from Chinese Firms,” *Working Paper*, 2023.

Matsuyama, Kiminori, “Increasing Returns, Industrialization, and Indeterminacy of Equilibrium,” *Quarterly Journal of Economics*, 1991, *106* (2), 617–50.

_ , “Agricultural productivity, comparative advantage, and economic growth,” *Journal of Economic Theory*, 1992, *58* (2), 317–334.

_ , “Agricultural productivity, comparative advantage, and economic growth,” *Journal of Economic Theory*, 1992, *58* (2), 317–334.

_ , “Agricultural productivity, comparative advantage, and economic growth,” *Journal of Economic Theory*, December 1992, *58* (2), 317–334.

_ , “Structural Change In An Interdependent World: A Global View Of Manufacturing Decline,” *Journal of the European Economic Association*, 2009, *7* (2-3), 478–486.

Melitz, Marc J, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 2003, *71* (6), 1695–1725.

Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 01 2002, *69* (1), 245–276.

Restuccia, Diego, Dennis Yang, and Xiaodong Zhu, "Agriculture and Aggregate Productivity," *Journal of Monetary Economics*, 2008, *55* (2), 234–250.

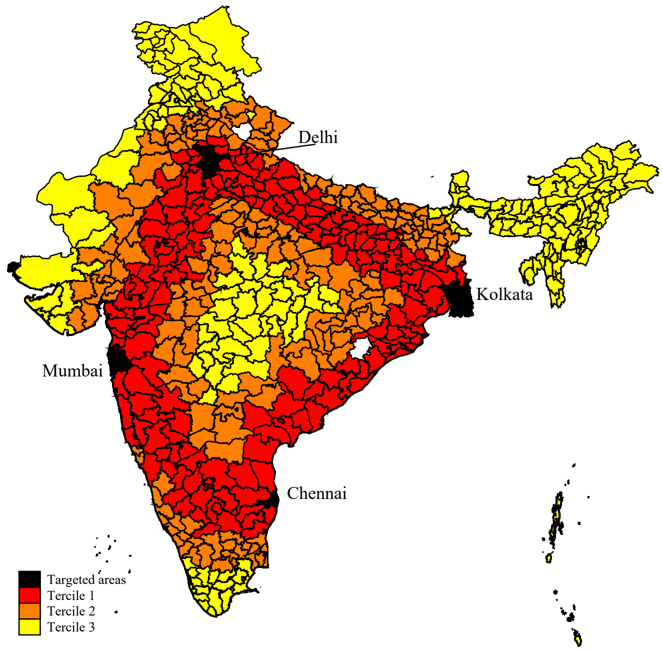
Storesletten, Kjetil, Bo Zhao, and Fabrizio Zilibotti, "Business Cycle during Structural Change: Arthur Lewis' Theory from a Neoclassical Perspective.," Working Paper 26181, National Bureau of Economic Research August 2019.

Uy, Timothy, Kei-Mu Yi, and Jing Zhang, "Structural change in an open economy," *Journal of Monetary Economics*, 2013, *60* (6), 667–682.

Van Leemput, Eva, "A passage to India: Quantifying internal and external barriers to trade," *Journal of International Economics*, 2021, *131*, 103473.

Vollrath, Dietrich, "How Important are Dual Economy Effects for Aggregate Productivity?," *Journal of Development Economics*, March 2009, *88* (2), 325–334.

(a) Indian Districts by GQ Proximity Terciles.



(b) Chinese cities by NTHS Proximity Terciles.

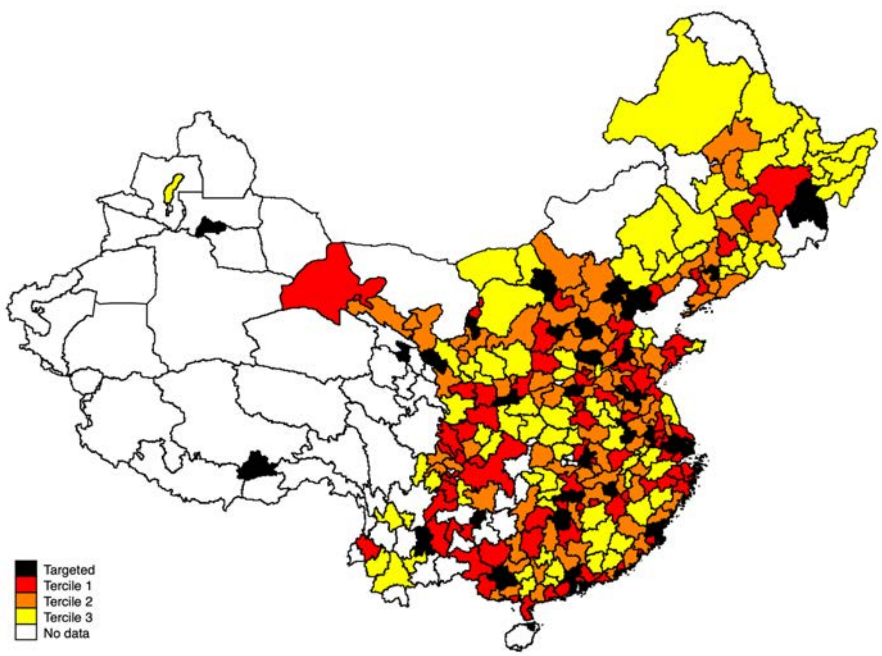


Figure 1: Location Terciles. The three terciles displayed in the panels are based on the shortest straight-line distance between the counterfactual highway and geometric center within each prefecture.

Table 1: Impact of Highway on the Number of Establishments

	Dependent variable = log(number of establishments)					
	India			China		
	(1)	(2)	(3)	(4)	(5)	(6)
$\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$	-0.136**	-0.106**	-0.100**	0.197***	0.195***	0.194***
	(0.061)	(0.050)	(0.049)	(0.0571)	(0.0551)	(0.0549)
log(output)		0.420***	0.431***		0.519***	0.514***
		(0.019)	(0.019)		(0.0877)	(0.0857)
log(population)			-0.491			(0.436)
			(0.061)			(0.326)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,676	5,676	5,611	2,925	2,925	2,925
	First stage			First stage		
$\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$	0.837***	0.837***	0.838***	0.666***	0.666***	0.667***
	(0.028)	(0.028)	(0.028)	(0.0416)	(0.0416)	(0.0415)
log(output)		0.004	0.004		-0.0446	-0.0503
		(0.003)	(0.004)		(0.0865)	(0.0885)
log(population)			0.033			0.265
			(0.021)			(0.414)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,676	5,676	5,611	2,925	2,925	2,925
Adj. R^2	0.824	0.824	0.824	0.801	0.801	0.801
First Stage F-Stat	892.5	897.8	899.7	255.73	256.23	257.45

Note: Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table 2: Impact of Highway on the Average Number of Employees

	Dependent variable = log(average number of employees)					
	India			China		
	(1)	(2)	(3)	(4)	(5)	(6)
$\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$	0.045	0.066	0.069	0.0439	0.0434	0.0446
	(0.048)	(0.045)	(0.046)	(0.0277)	(0.0276)	(0.0273)
log(output)		0.202***	0.199***		0.0444	0.0357
		(0.018)	(0.020)		(0.0419)	(0.0444)
log(population)			0.060			0.411
			(0.051)			(0.284)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,669	5,669	5,605	2,919	2,919	2,919
	First stage			First stage		
$\chi_{\text{tercile}=1,2} * \chi_{t>t_c}$	0.837***	0.837***	0.837***	0.665***	0.665***	0.665***
	(0.028)	(0.028)	(0.028)	(0.0417)	(0.0417)	(0.0416)
log(output)		0.004	0.004		-0. -0.0416	-0.0472
		(0.003)	(0.004)		(0.0877)	(0.0896)
log(population)			0.033			0.260
			(0.021)			(0.415)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,669	5,669	5,605	2,919	2,919	2,919
Adj. R^2	0.823	0.824	0.824	0.798	0.798	0.798
First Stage F-Stat	891.2	896.6	898.7	253.55	254.16	255.36

Note: Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table 3: Impact of Highway on Agricultural Employment

	Dependent variable = log(agricultural employment)					
	India			China		
	(1)	(2)	(3)	(4)	(5)	(6)
$\chi_{\text{tercile}=1,2}{}^{*}\chi_{t>t_c}$	-0.041	-0.042	-0.009	-0.316***	-0.315***	-0.317***
	(0.044)	(0.044)	(0.035)	(0.0958)	(0.0954)	(0.0955)
log(output)		-0.004	-0.005		-0.113	-0.0973
		(0.008)	(0.006)		(0.221)	(0.216)
log(population)			0.905***			-0.769
			(0.046)			(0.830)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,601	5,601	5,601	2,909	2,909	2,909
	First stage			First stage		
$\chi_{\text{tercile}=1,2}{}^{*}\chi_{t>t_c}$	0.835***	0.836***	0.837***	0.664***	0.664***	0.664***
	(0.028)	(0.028)	(0.028)	(0.0419)	(0.0419)	(0.0418)
log(output)		0.004	0.004		-0.0412	-0.0470
		(0.004)	(0.004)		(0.0877)	(0.0896)
log(population)			0.034			0.276
			(0.021)			(0.416)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,601	5,601	5,601	2,909	2,909	2,909
Adj. R^2	0.822	0.823	0.824	0.796	0.796	0.797
First Stage F-Stat	871.7	878.4	889	251.01	251.63	252.81

Note: Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table 4: Number of Establishments and Agricultural Employment Share

	Dependent variable = log(number of establishments)					
	India			China		
	(1)	(2)	(3)	(4)	(5)	(6)
	Low agr	Med agr	Large agr	Low agr	Med agr	Large agr
$\chi_{\text{tercile}=1,2} {}^{*} \chi_{t>t_c}$	-0.154**	-0.182**	0.0314	-0.116	0.158*	0.347***
	(0.069)	(0.089)	(0.109)	(0.0941)	(0.0856)	(0.0977)
log(output)	0.305***	0.302***	0.294***	0.510***	0.335***	0.489***
	(0.036)	(0.028)	(0.034)	(0.177)	(0.124)	(0.154)
log(population)	0.081	-0.063	0.051	0.292	-0.798	-2.706***
	(0.099)	(0.116)	(0.099)	(0.353)	(0.500)	(0.940)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,918	1,855	1,838	899	899	885
	First stage			First stage		
$\chi_{\text{tercile}=1,2} {}^{*} \chi_{t>t_c}$	0.904***	0.747***	0.841***	0.673***	0.559***	0.646***
	(0.035)	(0.064)	(0.046)	(0.0738)	(0.0819)	(0.0808)
log(output)	-0.001	0.002	0.009	-0.133	0.141	-0.162
	(0.005)	(0.005)	(0.006)	(0.217)	(0.185)	(0.112)
log(population)	0.035	0.010	0.057	-0.0280	-0.774	1.945**
	(0.033)	(0.037)	(0.047)	(0.528)	(0.753)	(0.839)
location FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,918	1,855	1,838	899	899	885
Adj. R^2	0.453	0.322	0.788	0.453	0.322	0.481
First Stage F-Stat	685.8	136.8	335.1	82.99	46.53	63.83

Note: Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.