Quality Upgrading and Productivity Gains from Domestic Market Access Changes in India

Sang Hoon Kong*

October 30, 2024

Abstract

Does greater access to domestic higher-income markets lead to quality upgrading and gains in productivity among local firms? Taking advantage of a major highway project in India, I find that firms responded differently to large and small increases in access to higher-income markets. When shocks were large, local firms produced higher-quality output by using more skilled labor, capital, and expensive materials, leading to productivity gains. For small shocks, local firms also produced higher-quality output but only by using more quantities of input and without any signs of productivity gains. Seen through the lens of the model, these results suggest that firms face substantial adjustment costs that prevent them from adopting more sophisticated production processes when shocks are small. Since income differences found within India are relatively narrow, my analysis suggests that developing countries do not need to rely solely on demand from high-income countries to incentivize quality upgrading and to capture associated productivity benefits.

JEL Codes: D22, L60, O14, O18

1 Introduction

Developing countries often seek greater access to high-income countries like the U.S. or members of the E.U. as a pathway to local economic development. One rationale is that high-income consumers in these countries are willing to pay more for product quality. Hence, increasing market access to these countries can provide greater incentives for local firms to upgrade quality and capture associated productivity benefits. For example, Atkin et al. (2017b) document that an intervention matching Egyptian rug makers to buyers from high-income countries—who requested higher-quality rugs—caused increases in productivity among local workers. However, for the vast majority of firms in developing countries that do not sell to high-income countries, the demand from domestic high-income consumers or those in countries at similar stages of development are likely more relevant. Can narrower

*Department of Economics, Columbia University; 420 W 118th, New York, NY 10027; s.kong@columbia.edu. I am indebted to David Weinstein, Donald Davis, and Réka Juhász for their advice and guidance. I am grateful for comments from Jonathan Dingel, Eric Verhoogen, Nicolás de Roux, and participants at the Trade and Development Colloquiums at Columbia University. I received help from the Lehman Social Sciences Library at Columbia University in accessing the Indian manufacturing data (ASI). All errors are my own.

income differences such as those found within the borders of a country still incentivize quality upgrading and lead to productivity gains? Answers to these questions are important in informing developing countries on whether policies that increase access to markets outside high-income countries can promote local economic development through quality upgrading.

In this paper, I take advantage of a major highway project in India to answer these questions by examining if increased access to domestic higher-income markets caused quality upgrading and productivity gains among local firms. I report two main findings. First, increased access to higher-income markets caused quality upgrading among local firms in Indian districts. In 2001, the 75/25th percentile ratio of per-capita income across Indian districts was 1.6, only a tenth of the same ratio measured using international per-capita income differences. Therefore, my first finding shows that relatively narrow income differences found within India were sufficient in inducing firm-level quality upgrading. Second, I find that only large increases in access to higher-income markets converted these quality upgrading responses into productivity gains. Small increases in access to higher-income markets induced quality upgrading without signs of productivity gains. When seen through the lens of the model, these empirical findings suggest that adjustment costs associated with adopting more sophisticated production processes prevented most Indian firms from realizing productivity gains despite upgrading quality. In order to also achieve productivity gains, my analysis suggests that the increase in access to narrowly higher-income markets has to be achieved through large reductions in trade costs or be accompanied by lowering of adjustment costs associated with adopting sophisticated production processes.

India initiated its National Highway Development Project (NHDP) in 2000, which upgraded existing national highways to four-lane roads in order to improve road speeds. For some districts, these upgraded roads meant that they were now better connected to other high-income districts in India. By digitizing the Indian road network and the upgraded roads, I am able to estimate the resulting changes in trade costs between every pair of districts from 2000 to 2009. I use a model to map these trade cost changes and data on district income into measures of changes in access to higher-income markets for each district, which I relate to the local manufacturing plants' output quality and productivity changes. I address identification concerns regarding the non-random placement of road upgrades by restricting the estimation sample to districts where NHDP roads were built. This avoids comparing locations close to the NHDP roads to locations far away. The identification assumption is then that within this subset of districts, the changes in access to higher-income markets resulting from the NHDP road upgrades are uncorrelated with other factors that can affect local firms' output quality and productivity.

Product quality is not observed in India's Annual Survey of Industries (ASI) that I use for this study and changes in output quality have to be inferred. My strategy for detecting variety-level (unique plant and product pairs) changes in quality is to use both the changes in marginal cost and product appeal. The marginal cost of a variety is measured as the per-unit cost of materials and labor while product appeal is derived from demand residuals recovered using data on prices and quantities (Khandelwal, 2010; Hottman et al., 2016). Each measure is imperfect. Increases in marginal cost may reflect decreases in productivity (Garcia-Marin and Voigtländer, 2019) while increases in product appeal may reflect demand shocks unrelated to quality. However, a concurrent increase in a variety's marginal cost and product appeal makes a compelling case that the variety's quality was improved since alternative explanations would have to simultaneously account for decreases in productivity and increases in product appeal. Firm productivity is measured by revenue per labor and revenue total factor productivity (TFP).

Using this strategy, I first find that increased access to higher-income markets caused local plants to spend more on materials and labor to produce each unit of their variety. For example, I estimate that a standard deviation increase in access to higher-income markets caused 20% higher increases in the marginal costs of local varieties relative to those produced in unaffected locations. Moreover, I find that these plants devoted more quantities of inputs into producing each unit of output, which is consistent with the findings in Atkin et al. (2017b) that more hours of labor were used to produce higher-quality rugs.[1] These outputs were not only more costly but also more appealing. I interpret these increases in marginal cost and product appeal as increases in quality.

Turning to productivity, I estimate that large and small increases in access to higher-income markets had differential effects on measured productivity, skill use (ratio of non-production to production labor), value of capital, and material input prices. Large increases in access to higher-income markets caused relatively higher increases in all of these variables, suggesting that these firms undertook reorganization of their production process to produce higher-quality output. I estimate that plants in districts exposed to the largest shocks, comparable to a 6.5% reduction in trade costs, experienced 54% higher growth in revenue per labor and 19% higher revenue TFP growth relative to plants in unaffected locations. In contrast, small increases in access to higher-income markets had no effects on any of these variables. In response to small shocks, local firms upgraded their output quality merely by using more quantities of inputs, without showing visible changes to their production processes.

These results indicate that relatively narrower income differences found within India were able to incentivize firm-level quality upgrading. At the same time, these quality upgrading responses did not translate into productivity gains for most Indian firms. To understand why this was the case, the model in this paper conceptualizes two ways of quality upgrading. In the first, a firm upgrades output quality through using more quantities of input per output but without changing the nature of the production process, which I refer to as increasing "input intensity". For example, one way in which the Egyptians produced higher-quality rugs in Atkin et al. (2017b) was by increasing the thread count. Producing a rug with a higher thread count requires more materials and hours of labor, but does not change the nature of tasks that workers engage in. As a result, there are no changes to the firm's productivity or skill use. In the second, a firm upgrades output quality through adopting a more sophisticated production process that requires more skilled workers and higher-quality materials, which I refer to as "sophistication". For example, another way that the Egyptians produced higher-quality rugs was by creating "modern designs", which would require the firm to hire workers who can design or train existing workers on designing.

If firms face a fixed adjustment cost to change the level of their sophistication but are free to adjust input intensity, the model predicts that small shocks will only cause firms to adjust their input intensity to upgrade quality. In comparison, large shocks make it profitable for firms to pay the adjustment cost and change both their input intensity and sophistication to upgrade quality. The model also predicts that these firms will display increased skill use and higher revenue productivity as a result of increasing sophistication. When seen through

[1] Although it is not the main focus of this study, these results reinforce a point raised by Atkin et al. (2019) and Verhoogen (2023) that quantity based measures of productivity (e.g., TFPQ) fail to reflect firm efficiencies when there are quality differences. A few researchers (De Loecker et al., 2016; de Roux et al., 2021) have begun incorporating output and input qualities into existing methods of production function and productivity estimation, and my findings indicate that these solutions are relevant for a wide set of products.

the lens of the model, the empirical findings suggest that adjustment costs associated with adopting more sophisticated production processes prevented most Indian firms from realizing productivity gains despite upgrading quality.

To the extent that productivity gains are the end goal, quality upgrading through increasing input intensity is insufficient. In order to encourage firms to upgrade quality through sophistication and realize productivity gains, the analysis in this paper suggests two possibilities. First, large increases in access to higher-income markets, achieved either through large reductions in trade costs with narrowly higher-income markets or through wider income differences. Second, small increases in access to higher-income markets coupled with lowering of the adjustment costs associated with sophistication. On the one hand, these suggestions justify the importance of increasing demand from high-income countries. On the other hand, they highlight how demand from narrowly higher-income markets can also be leveraged to promote both quality upgrading and productivity gains. Specifically, the analysis suggests that increasing access to narrowly higher-income markets should be achieved through large reductions in trade costs or be accompanied by lowering of the adjustment costs that firms face. Thus, developing countries do not need to rely solely on high-income countries to promote local economic development through quality upgrading.

The findings in this paper add to the growing evidence on the effect of access to higher-income markets on quality upgrading. Previous research have documented that increased access to higher-income markets lead to productivity gains (Atkin et al., 2017b), skill upgrading (Verhoogen, 2008; Brambilla et al., 2012), and use of more expensive material inputs (Bastos et al., 2018; Hansman et al., 2020) through the quality upgrading mechanism. These studies have relied on exogenous variation in foreign demand (e.g., exchange rate shocks) for identification and so they were looking at much wider international income differences. This paper complements these analyses by exploiting exogenous variations in domestic demand for output quality. My findings confirm that much narrower income differences are still able to incentivize quality upgrading and that sufficiently large reductions in trade costs can also generate comparable results on productivity, skill upgrading, and material prices. These studies make important advances to the earlier literature that found cross-sectional evidence on countries and firms that sell more expensive and higher-quality outputs to higher-income destinations (Hallak, 2006; Bastos and Silva, 2010; Hallak and Schott, 2011; Manova and Zhang, 2012; Kugler and Verhoogen, 2012; Dingel, 2017).

This paper joins a large body of research that relates market integration to economic development. In particular, a recent line of work (Faber, 2014; Allen and Arkolakis, 2014; Donaldson and Hornbeck, 2016; Donaldson, 2018) has focused on major infrastructure projects to study how changes in market access affect economic development. A number of studies (Datta, 2012; Ghani et al., 2016; Asturias et al., 2019; Abeberese and Chen, 2022; Baragwanath Vogel et al., 2024) have examined India's NHDP. In contrast to these studies, the present paper focuses on how a major infrastructure project affects economic development through the quality upgrading channel. In doing so, I underscore the fact that transportation infrastructure not only alter the size of the demand that a location faces but also the demand for quality that it faces.

Finally, the results in this paper emphasize the importance of understanding the impediments to technology adoption by suggesting that they may have prevented most Indian firms from realizing productivity gains through quality upgrading. A number of recent studies have taken a closer look at the impediments to technology adoption such as re-organization costs (Bloom et al., 2013; Atkin et al., 2017a; Hansman et al., 2020; Juhász et al., 2024b) and lack of access to knowledge (Juhász et al., 2024a). Since adjustment costs that firms face are not directly observed in this study, it is not possible to link them to any of these potential sources. Making this link appears crucial, especially since the analysis in this paper suggests that lowering these costs will make increasing access to narrowly higher-income markets more effective at generating productivity gains through incentivizing quality upgrading.

2 Theoretical Framework

This section develops a theoretical framework that links market access, quality upgrading, and measured productivity. The model gives a precise definition on a location's "access to higher-income markets" and shows how changes in trade costs affect firm decisions on how to upgrade their output quality. It also formalizes why large and small increases in access to higher-income markets will generate different responses from the firms.

2.1 Demand

There is a fixed set of locations (districts) indexed by $d \in D$ in the economy. There are iceberg trade costs between any two districts $o, d \in D$ so that $\tau_{od} > 1$ units of a good need to be shipped in order for a unit to arrive. Trade is free within each district so that $\tau_{dd} = 1$ for all $d \in D$. Each location is inhabited by consumers of heterogeneous types indexed by $i \in I$. A type-i consumer in location d has expenditure E_{id} and their population is denoted by N_{id}. Consumers have non-homothetic CES preferences (Comin et al., 2021) over the consumption of varieties of a differentiated good. The aggregate demand from i-consumers in location d for each variety v is given by:

$$C_{vid} = p_{vd}^{-\sigma} \varphi_{vi}^{\sigma-1} E_{id}^{\sigma} N_{id}, \tag{1}$$

where $\sigma > 1$ is the elasticity of substitution between the varieties, p_{vd} is the price of variety v in location d, and φ_{vi} is its appeal to i-consumers that is determined endogenously in the model. The variety's appeal to i-consumers is given by $\varphi_{vi} = U_i^{-1/q_v}$, where $q_v > 0$ is the quality of variety v and $U_i > 0$ is the utility of type-i consumers.[2] Note that the variety's appeal is increasing in quality provided $U_i > 1$.

To see the heterogeneity in consumers' preference for quality, observe that the (logged) relative expenditure shares on varieties $q_{v'} > q_v$ are given by

$$\log\left(\frac{p_{vd} C_{v'id}}{p_{vd} C_{vid}}\right) = (1-\sigma)\log\left(\frac{p_{v'd}}{p_{vd}}\right) + (\sigma - 1)\left(\frac{q_{v'} - q_v}{q_{v'} q_v}\right)\log U_i, \tag{2}$$

which is increasing in U_i. Thus, higher-utility consumers spend a greater share of their expenditure on higher-quality varieties. As we will see shortly, this heterogeneity in consumers' preferences for quality combined with trade costs will generate differences in firms' incentives to upgrade quality across space.

In Appendix Section B.1, I provide a sketch of a general equilibrium environment in which non-homothetic CES utility studied by Comin et al. (2021) gives rise to the demand function described here. In this paper, I take a partial equilibrium view in the sense that

[2] If consumer preferences for quality are assumed away (i.e., $q_v = 1$ for all v), then the demand function in equation (1) reduces to $C_{vid} = p_{vd}^{-\sigma} U_i^{1-\sigma} E_{id}^{\sigma} N_{id}$. Let $P_d = (\sum_v p_{vd}^{1-\sigma})^{1/(1-\sigma)}$ denote the standard CES price index in location d. Then using the fact that $U_i = E_{id}/P_d$, we get the standard expression for CES demand, $C_{vid} = p_{vd}^{-\sigma} P_d^{\sigma-1} E_{id} N_{id}$.

the feedback effects of firm decisions on the non-homotheticity of consumers are ignored.[3] In other words, U_i, E_{id}, and N_{id} are all taken to be exogenous variables in this model and I instead focus on firm responses to changes in trade costs with high- and low-income locations. For convenience of exposition, I refer to consumers with higher U_i as higher-income consumers going forward.

2.2 Technology

I assume that firms are able to upgrade the quality of their output in two different ways: increasing input intensity (x) and increasing sophistication (s). It will be useful to have a guiding example in mind. For this purpose, consider a rug maker like those studied by Atkin et al. (2017b). In their study, Egyptian rug makers produced higher-quality rugs for richer buyers abroad while producing simplistic rugs for the local market. The higher-quality rugs differed from the local rugs in several aspects. In one dimension, they had higher thread counts. In another, they had more "modern" designs and had better craftsmanship. I distinguish these dimensions of quality upgrading by considering whether they simply require more quantities of inputs (increasing input intensity) or require more skilled workers (increasing sophistication). Increasing thread counts would fall under increasing input intensity while creating more modern designs or providing better craftsmanship would fall under increasing sophistication.

To formalize these ideas, I assume that output quality q_v is determined by input intensity (x_v) and sophistication (s_v) according to a CES technology:

$$q_v = \left(x_v^\theta + s_v^\theta\right)^{1/\theta}, \tag{3}$$

where $\theta < 0$ is the elasticity of substitution between the two ways of upgrading quality that I assume are complementary. I also assume that more sophisticated production requires higher-quality inputs so that s_v denotes not only the firm's level of sophistication but the quality of its material inputs and the skill of its workers. The rising costs of purchasing higher-quality materials and more skilled labor are modelled as log-linear functions of sophistication with an elasticity $\eta > 0$:

$$p_{Mv} = p_{Mo} s_v^\eta \quad \text{and} \quad w_v = w_o s_v^\eta, \tag{4}$$

where p_{Mo} and w_o are local prices of materials and wages that are taken as given.

In turn, to produce physical quantities of output y_v, the firm has to combine materials M_v and labor L_v using a Cobb-Douglas production function:

$$y_v = \frac{z_v}{x_v} M_v^\alpha L_v^{1-\alpha}, \tag{5}$$

where z_v is the firm's exogenous efficiency in production and $\alpha \in (0,1)$ is the output elasticity on materials. Unlike in standard production functions, the firm's choice of input intensity x_v features directly in equation (5) and has the effect of lowering the quantity of output that the firm produces with a given quantity of input (i.e., lowers the quantity productivity).

Finally, I assume that firms inherit some initial level of sophistication s_{0v} that they take as exogenous. In order for firms to deviate away from this level of sophistication, they need to pay a fixed adjustment cost $F > 0$ to reorganize their production process. In contrast, they can always adjust input intensity freely. Moreover, firms are assumed to be myopic and do not consider the implications of their choices of sophistication on future profits.

[3] For example, Bastos et al. (2018) also consider a similar non-homothetic demand function but hold the income of consumers fixed.

2.3 Profit Maximization

Each firm v produces a unique variety and competes with other firms under monopolistic competition. Consider a firm v in location $o \in D$. It first chooses whether to pay the adjustment cost or not. If it does ($a_v = 1$), the firm then chooses both its input intensity (x_v) and sophistication (s_v). Otherwise ($a_v = 0$), the firm only chooses its input intensity (x_v) and is stuck at its initial level of sophistication ($s_v = s_v^0$) that it takes as given. Subsequently, the firm chooses the prices ($\{p_{vd}\}_{d \in D}$) and input usages (M_v, L_v). The firm makes these choices to maximize profits given by:

$$\pi_v = \sum_{d \in D} \sum_{i \in I} p_{vd} C_{vid} - p_{Mo} s_v^{\eta} M_v - w_o s_v^{\eta} L_v - F a_v. \tag{6}$$

Minimizing the cost function for a unit output, the marginal cost of production equals

$$mc_v = \kappa \frac{x_v}{z_v} p_{Mo}^{\alpha} w_o^{1-\alpha}, \tag{7}$$

where $\kappa = \eta^{\eta/\theta}(1+\eta)^{-(1+\eta)/\theta} \alpha^{-\alpha}(1-\alpha)^{\alpha-1}$ is a constant of model parameters. Now, let us consider minimizing this marginal cost of production to produce an output of quality q. If the firm has paid the adjustment cost, it is able to adjust both input intensity and sophistication to achieve the target quality q. Otherwise, it can only adjust its input intensity. The resulting minimized marginal costs under adjustment and without are

$$mc_{1v}(q) = \frac{\kappa_1}{z_v} q^{1+\eta} p_{Mo}^{\alpha} w_o^{1-\alpha} \text{ and} \tag{8}$$

$$mc_{0v}(q) = \frac{\kappa_0}{z_v} \left(q - s_{0v}^{\theta}\right)^{1/\theta} (s_{0v})^{\eta} p_{Mo}^{\alpha} w_o^{1-\alpha}, \tag{9}$$

respectively. Note that $mc_{1v}(q) \leq mc_{0v}(q)$ for all q since the firm has strictly less margins of adjustment when it does not pay the adjustment cost. In fact, since $\theta < 0$, this inequality extends to the elasticity of marginal costs with respect of quality. That is, $\frac{\partial \ln mc_{1v}}{\partial \ln q} \leq \frac{\partial \ln mc_{0v}}{\partial \ln q}$ for all q. Intuitively, when input intensity and sophistication are complementary ($\theta < 0$), it becomes increasing costly to produce higher-quality output from adjusting input intensity alone, because it is less efficient at increasing quality at sub-optimal levels of sophistication.

As is standard, the firm optimally chooses to charge a constant markup over marginal cost for its price. The optimal factory-gate price is $p_v = \left(\frac{\sigma}{\sigma-1}\right) mc_v$ while the prices that consumers face are $p_{vd} = \tau_{od} p_v$ for all $d \in D$. Substituting these optimal prices into the demand equation in (1), the total profit of the firm can be written as

$$\pi_v = \left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}\right) mc_v^{1-\sigma} \Phi_v^{\sigma-1} \mathrm{MA}_o - F a_v. \tag{10}$$

Here, MA_o is the total market access that firms in location o face and is given by

$$\mathrm{MA}_o = \sum_{i \in I} \sum_{d \in D} \tau_{od}^{1-\sigma} E_{id}^{\sigma} N_{id}. \tag{11}$$

Like in previous studies of market access (e.g., Donaldson and Hornbeck (2016)), the market access term in equation (11) sums across each destination market's size (i.e., income and population) after discounting by trade costs.[4] However, it differs in that it distinguishes consumers of different types $i \in I$. To make this explicit, let us define location o's market access to type-i consumers as

$$\mathrm{MA}_{io} \equiv \sum_{d \in D} \tau_{od}^{1-\sigma} E_{id}^{\sigma} N_{id} \tag{12}$$

and note that the total market access is simply the sum of these consumer specific market access terms, $\mathrm{MA}_o = \sum_{i \in I} \mathrm{MA}_{io}$. The consumer-specific market access terms in equation (12) succinctly summarize firms' access to consumers with differing preferences for higher-quality varieties. For instance, having lower trade costs with a destination market that is inhabited mostly by lower-income consumers will not have a large impact on the market access to higher-income consumers. Next, the term Φ_v in equation (10) is the variety's (market-access weighted) average product appeal defined as

$$\Phi_v \equiv \left(\sum_{i \in I} \frac{\mathrm{MA}_{io}}{\mathrm{MA}_o} \varphi_{vi}^{\sigma-1} \right)^{\frac{1}{\sigma-1}}, \tag{13}$$

where I remind that φ_{vi} is the appeal of a variety with quality q_v to type-i consumers.

These expressions make clear the trade-off the firm faces when choosing its output quality to maximize profits in equation (10). On the one hand, increasing output quality allows the firm to capture more demand and increase revenue by increasing Φ_v. In particular, this incentive is stronger when a larger share of its demand is coming from higher-income consumers as equation (13) shows. On the other hand, producing higher-quality output is more costly, increasing mc_v. The optimal choice of output-quality balances these two effects and is determined by the following first order condition:

$$\sum_{i \in I} \frac{U_i^{(1-\sigma)/q_v} \mathrm{MA}_{io}}{\sum_{j \in I} U_j^{(1-\sigma)/q_v} \mathrm{MA}_{jo}} \left(\frac{\ln U_i}{q_v} \right) = \frac{\partial \ln mc_v}{\partial \ln q_v}. \tag{14}$$

This expression shows that the firm's optimal quality depends on its location's market access to different types of consumers and the elasticity of marginal cost with respect to quality. For instance, when the location has relatively greater access to higher-income consumers who value quality more, the firm will choose to produce higher-quality output. This demand-side incentives is countered by the costs of producing higher-quality output. In fact, since $\frac{\partial \ln mc_{1v}}{\partial \ln q} \leq \frac{\partial \ln mc_{0v}}{\partial \ln q}$ for all q, it follows that the optimal output quality will be greater under adjustment. That is, $q_{1v} \geq q_{0v}$ for all $\{\mathrm{MA}_{io}\}_{i \in I}$.

Finally, the firm chooses whether to pay the adjustment cost or not. It will pay the adjustment cost if and only if the additional profits arising from a more sophisticated production process covers the adjustment cost:

$$\left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}} \right) \left[\left(\frac{\Phi_{1v}}{mc_{1v}} \right)^{\sigma-1} - \left(\frac{\Phi_{0v}}{mc_{0v}} \right)^{\sigma-1} \right] \mathrm{MA}_o \geq F. \tag{15}$$

[4] The market access term is typically defined with a price index. For example, Donaldson and Hornbeck (2016) defines market access as $\sum_{d \in D} \tau_{od}^{1-\sigma} P_d^{\sigma-1} E_d N_d$ (suppressing consumer type), where P_d is the standard CES price index in location d. Since utility is given by $U = E_d/P_d$ in their work, this market access term can be rewritten as $U^{1-\sigma} \sum_{d \in D} \tau_{od}^{1-\sigma} E_d^{\sigma} N_d$, which gives the market access term as defined in this work.

Observe that the left-hand-side of the inequality is strictly increasing in output quality since marginal cost rises with a greater elasticity when the adjustment cost is not paid. This implies that there is a cutoff value of quality q^* above which the left-hand-side is greater and so the firm decides to adjust its level of sophistication. Below this cutoff value, the firm achieves higher quality only by adjusting its input intensity.

The model implies a relationship between labor productivity and sophistication as well. The Cobb-Douglas production function implies that firms spend a fixed share $(1-\alpha)$ of their cost on labor. Since revenue is proportional to variable costs, $R_v = (\frac{\sigma}{\sigma-1})(p_{Mv}M_v + w_v L_v)$. Combining these two results, I can show that revenue per labor is log-linear in sophistication:

$$\frac{R_v}{L_v} = \left(\frac{\sigma}{\sigma-1}\right)\left(\frac{1}{1-\alpha}\right) w_o s_v^{\eta}. \tag{16}$$

In contrast, when the level of sophistication is fixed, changes in input intensity has no implications on the firm's labor productivity.

2.4 Effects of Trade Cost Changes on Quality Upgrading and Productivity

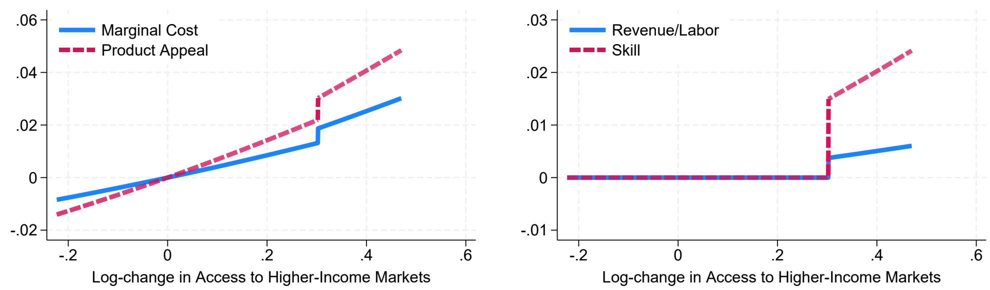
I now analyze how changes to bilateral trade costs affect firms' behavior on quality upgrading and obtain estimating equations that I can take to the data. The model I take to the data has two types of consumers: those with high ($i = H$) and low ($i = L$) income. With only two types of consumers, the optimal output quality chosen by the firm in equation (14) is strictly increasing in the relative market access to higher-income consumers ($\text{MA}_{Ho}/\text{MA}_{Lo}$) provided that $U_H > U_L$. Going forward, I will refer to this relative market access to higher-income consumers as "access to higher-income markets" (AHIM):

$$\text{AHIM}_o \equiv \frac{\text{MA}_{Ho}}{\text{MA}_{Lo}} = \frac{\sum_{d\in D} \tau_{od}^{1-\sigma} E_{Hd}^{\sigma} N_{Hd}}{\sum_{d\in D} \tau_{od}^{1-\sigma} E_{Ld}^{\sigma} N_{Ld}}. \tag{17}$$

This variable captures the demand-side incentives that firms in location o face to produce higher-quality output.

Given exogenous variables $\{w_d, p_{Md}, E_{id}, N_{id}, \tau_{od}, z_v, s_{0v}\}$, an equilibrium is defined as firms' choices of $\{a_v, x_v, s_v, p_{vd}, M_v, L_v\}$ that maximize profits as outlined in Section 2.3. Introducing time subscript t, consider an exogenous shift in trade costs from τ_{odt} to $\tau_{odt'}$. Denote the resulting access to higher-income markets as $\text{AHIM}_{ot'}$. The model predicts that increases in access to higher-income markets will cause local firms to upgrade their output quality. In turn, this higher output quality translates into higher marginal cost in equations (8) and (9) and higher product appeal in equation (13). These predictions are illustrated in a numerical example that is plotted in Figure 1. For clarity, this numerical example assumes that all firms were able to adjust to the optimal level of sophistication in the initial period t before the changes in access to higher-income markets materialized. It also abstracts away from changes in any other exogenous variables in the model. As the left panel shows, increases in access to higher-income markets lead to increased marginal cost and product appeal throughout the distribution of shocks. There is a discontinuous jump around 0.3, which is where the firms decide to pay the adjustment cost because the resulting loss in profits from only being able to adjust input intensity become too large. The marginal cost and product appeal continue to rise beyond the jump.

Figure 1: A Numerical Example of Model Predictions



In contrast, the model predicts that small increases in access to higher-income markets will not lead to changes in sophistication since the adjustment costs are too high. When sophistication remains the same, there are no changes in the skill of the workers and labor productivity as equation (16) shows. These predictions are illustrated in the flat portion of the right panel of Figure 1. Only when the firms decide to pay the adjustment cost around 0.3, we see a large jump in skill and labor productivity, which continue to increase afterwards. These are the main predictions of the model that I test in the data. Of course, changes in other exogenous variables that occur concurrently with the changes in access to higher-income markets mean that the resulting patterns in the data will be less precise. Nevertheless, I aim to test the broader predictions of the model that changes in access to higher-income markets lead to increases in marginal cost and product appeal regardless of the size of the shocks whereas increases in labor productivity and skill only show up for large shocks.

3 Data and Measurement

This section describes the different sources of data used for this study. First, I obtain population counts and various demographic information of each district from India's 2001 Census. In lieu of reliable district-level measures of income in 2001, I use the expenditure per capita of each district derived from the Household Consumer Expenditure Survey (HCES). Second, I construct three panels of data using the Annual Survey of Industries (ASI) data from 2000 to 2009 at the level of output varieties (unique plant and output pairs), input varieties (unique plant and input pairs), and plants. Third, I obtain a list of road upgrades that were part of the National Highway Development Project (NHDP) with information on their completion dates from India's National Highways Authority of India (NHAI), the entity in charge of the project. Fourth, I digitize scanned road maps of India to approximate the road network in 2000 and combine it with the list of upgraded roads to calculate travel times and trade costs between districts in 2000 and 2009. Finally, I specify how I use these data to measure the changes in each district's access to higher-income markets and each variety's product appeal. The assembled data allow me to compare variety-level or plant-level cumulative changes from 2000 to 2009 against the concurrent changes in access to higher-income markets due to the NHDP road upgrades (i.e., from March 2000 to March 2009).

3.1 District Demographics and Income

In the 2001 Census, India is divided into 35 states (including 7 union territories) that are further divided into 593 districts. For this study, I combine the nine districts of Delhi and the two districts of Mumbai to be consistent with other sources of data used in this study. The three island districts of Lakshadweep, Andamans, and Nicobars are excluded from the analysis. For each district, I obtain its total population, share of male population, share of urban population, and share of population with a high school degree or higher.

The 57th round of the National Sample Survey, which includes the HCES, was carried out from July 1, 2001 to June 30, 2002. Geographically, it covers all of India except (i) Leh and Kargil districts of Jammu & Kashmir, (ii) interior villages of Nagaland situated beyond 5 km of the bust route, and (iii) inaccessible villages in Andaman and Nicobar islands. The survey records each sampled household's size and expenditure on a number of goods and services for a reference period of 30 days.[5] These variables are first aggregated to calculate the national per-capita expenditure and then multiplied by 12 to get the annual figure.

Applying the exchange rate from 2001, the national per-capita expenditure was 146 US dollars, which is about a third of 450 US dollars—the World Bank's figure on India's GDP per capita in 2001. All households in the survey are then categorized as either high- or low-income, based on whether they are above or below this national figure. Within each district, the per-capita expenditure is calculated for these two groups of households to obtain measures of E_{idt} for $i \in \{H, L\}$ and $t = 2001$. On average, the district per-capita expenditure for high-income households are about 1.5 times higher than the average district per-capita expenditure. Next, for each district, I calculate the share of district population belonging to the two categories of households and then multiply by the district's total population from the Census to obtain measures of N_{idt} for $i \in \{H, L\}$ and $t = 2001$.

3.2 Plant-level and Variety-level Panels

The ASI surveys activities of registered manufacturing plants in India, those with 10 or more employees and using power and those with 20 or more employees without using power. Of these registered plants, the ASI in each year covers all plants with 100 or more employees and a rotating random sample of plants with 10 to 99 employees (20 to 99 employees for plants that do not use power).[6] Geographically, the ASI covers all states of India except Arunachal Pradesh (13 districts), Mizoram (8 districts), Sikkim (4 districts), and Lakshadweep (1 district). The location of each plant is identified down to the district-level.

An important feature of the ASI is that the value and physical quantities of output (input) are reported for each plant using narrowly defined product categories (ASI Commodity Classification, ASICC), which can then be used to derive their unit prices. The ASICC assigns five-digit codes to over 5,000 products and specifies a unit of measurement for each product, allowing for a meaningful comparison of quantities across plants and time.[7] There

[5] Covered items include food, pan, tobacco, intoxicants, fuel, light, clothing, footwear, durable goods, cooked meals, education, medical, rent, and taxes. Any purchases made towards the household's productive enterprises are excluded.

[6] All plants from the six less industrially developed states (Manipur, Meghalaya, Nagaland, Sikkim, Tripura, and Andaman & Nicobar Islands) are surveyed.

[7] The number of product codes are comparable to that of the six digit Harmonized Systems (HS) codes widely used for international trade data. As an example, there are ten categories for milk products. These are "fresh", "flavored, not frozen", "chilled or frozen", "skimmed or pasteurized, non-flavored", "condensed", "powder", "butter", "fat", and "sweet" milk. There is one residual category "milk products not elsewhere classified (n.e.c)".

are periodic changes to the ASICC, but there are no official concordances across time that I am aware of. In order to track products that are defined consistently over time, I only examine quantities and unit-prices of ASICC codes whose descriptions did not change during the entire sample period.[8] Using data on each (output and input) variety's (unique plant and 5-digit ASICC product pair) revenue (R_{vt}) and physical quantities (y_{vt}), I can readily obtain the unit prices of output and input varieties (p_{vt} and p_{Mvt}, respectively).

In addition, the ASI records each plant's total expenditure on material and labor inputs. Unfortunately, one shortcoming of the ASI is that the input usage and costs are not recorded separately for each output product but only as a whole for the plant. This issue is not unique to the ASI and is shared by most existing dataset on plant-level manufacturing activity. As a result, I make the following assumption to distribute input usage and costs across multiple output varieties within each plant. I assume that each product's input usage and costs are proportional to that product's share of revenue within the plant. That is,

$$\text{Material Cost}_{vt} = \frac{\text{Revenue}_{vt}}{\sum_{v \in V_{ft}} \text{Revenue}_{vt}} \text{Material Cost}_{ft}, \tag{18}$$

where V_{ft} is the set of products (5-digit ASICC) produced by plant f in year t. I apply the same method to distribute plant-level labor costs to each product it produces. Then, I obtain the marginal cost of production as the sum of material and labor costs divided by the total quantity of output:

$$mc_{vt} = \frac{\text{Material Cost}_{vt} + \text{Labor Cost}_{vt}}{y_{vt}}. \tag{19}$$

Technically speaking, what I refer to as marginal cost here is the average cost of production, which equals the marginal cost only under constant returns-to-scale production functions.

Taking advantage of the fact that I observe prices and quantities of material inputs that firms use, I construct a single index of material prices for each plant f following the approach of Bastos et al. (2018). I begin with variety-level (plant and input-product pair) input prices (p_{Mvt}), which are available whenever the plant reports both the purchase value and quantities. Then, I estimate the following regression equation:

$$\ln p_{Mvt} = \theta_{ft} + \theta_{gt} + \text{error}_{vt}, \tag{20}$$

where θ_{ft} are plant-year fixed effects and θ_{gt} are product-year fixed effects. Then, set plant-level material prices as $\ln p_{Mft} = \theta_{ft}$. The idea is that the product-year fixed effects θ_{gt} absorbs the price differences due to compositional differences in input products that each plant uses and so θ_{ft} captures the plant-specific component from the remaining variation. This allows me to measure quantities of material use by deflating the material cost by the plant-level material price, $M_{ft} = \text{Material Cost}_{ft}/p_{Mft}$. Using the within-plant revenue shares, I measure variety-level material use as $M_{vt} = \frac{\text{Revenue}_{vt}}{\sum_{v \in V_{ft}} \text{Revenue}_{vt}} M_{ft}$. Variety-level labor use is measured analogously.

Turning to the measurement of skill use, I use the ratio of non-production to production labor (measured in man-days) at each plant (s_{ft}). Non-production workers refer to persons

[8] The ASI surveys for 2000–2008 adopt essentially the same ASICC with less than 20 products being added during this period. However, the ASICC adopted in ASI 2008–2010 introduces a greater number of changes. For consistency over time, I only work with product codes that can be tracked across time based on their product description and unit of quantity.

engaged in supervisory, managerial, store keeping, sales, or purchasing roles while production workers refer to all other persons directly involved in the manufacturing process. Although crude, Verhoogen (2008) and Brambilla et al. (2012) among others have found this measure of skill use (in their data) to be informative. For example, Brambilla et al. (2012) found that the share of non-production workers increased when exporters gained better market access to higher-income countries. Other plant-level variables that I use from the ASI are total revenue (R_{ft}), quantities of total labor (L_{ft}, in man-days), wages (w_{ft} = Labor Cost_{ft}/L_{ft}), and book value of capital (K_{ft}). I measure district-level wages (w_{dt}) and material prices (p_{Md}) by taking the geometric average of plant-level wages and material prices in each district.[9]

I also obtain two measures of productivity. These are revenue per labor (R/L_{ft}) and revenue total-factor-productivity (TFP^R_{ft}), which is defined as

$$\ln \text{TFP}^R_{ft} = \ln R_{ft} - \alpha_M \cdot \ln M_{ft} - \alpha_L \cdot d \ln L_{ft} - \alpha_K \cdot \ln K_{ft}, \tag{21}$$

where the α-parameters are the output elasticities of a Cobb-Douglas production function. I estimate these output elasticities following the methodology of Levinsohn and Petrin (2003) and Ackerberg et al. (2015).[10] These measures of revenue productivity have been criticized (e.g., see Foster et al. (2008) and Syverson (2011)) for using revenues in place of physical quantities to measure output. However, Atkin et al. (2019) show that productivity measures using output quantities (e.g., TFPQ) actually perform worse than revenue productivity measures when product quality is changing. As quality changes are central in this paper, I opt to examine revenue productivity measures.

I use these ASI data from year 2000 to 2009 to construct panels of output varieties, input varieties, and plants.[11][12] The panel of plants allows me to examine plant-level changes in manufacturing outcomes from 2000 to 2009. Then, the panel of output varieties allows me to examine variety-level outcomes of these plants. Summary statistics for these panels are provided in Appendix A.

3.3 National Highway Development Project

The National Highway Development Project (NHDP) is a major highway project in India, the largest in the country's history at the time, that aimed at upgrading the national highway network with four/six-lane roads. Given the time span of the panel on manufacturing activities, this study focuses on the development of four/six-lane highway routes on the Golden Quadrilateral (GQ) and the North-South and East-West (NS-EW) networks, which were part of the first two phases of the project. I obtained a list of road segments that were part of Phases I and II of the NHDP from NHAI, the entity in charge of the project. The

[9]In practice, I take the average of their logged values.

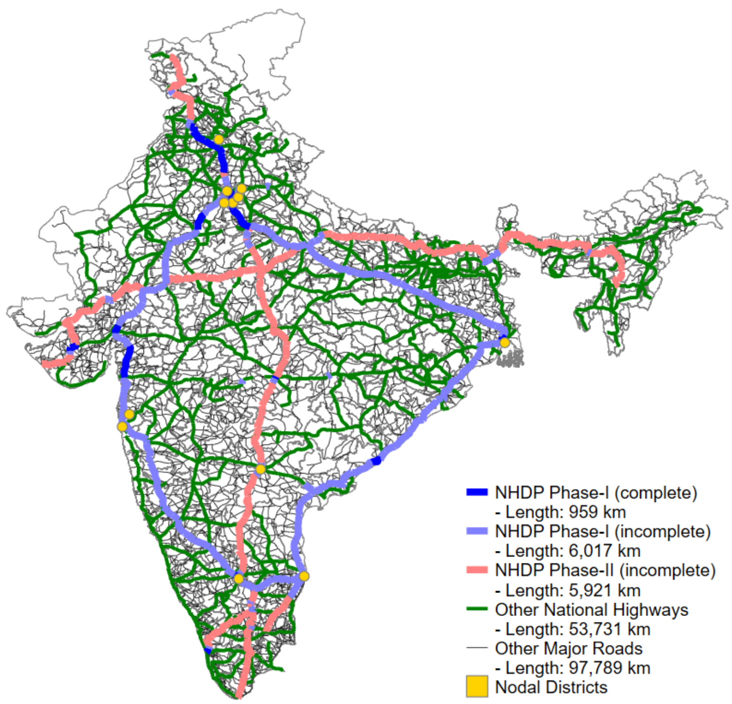
[10]After estimating the production function separately for each of the 64 product groups, I find that the means of the estimated elasticities are 0.41 for α_L, 0.26 for α_K, and 0.38 for α_M. The mean of estimated return-to-scale is 1.04.

[11]The ASI surveys manufacturing activities over the accounting year, which begins on April 1 and ends on March 31 of the following year. For example, the 2001–02 ASI covers activities from April 1, 2001 to March 31, 2002. For brevity, I refer to each year's ASI by its starting year since it takes up the majority of the survey period.

[12]The choice of the sample period is limited by two factors. First, lack of documentation by the data provider on the product and district codes prior to the survey in 2000 makes it difficult to reliably construct a panel of varieties using earlier years of data. Second, the ASI adopts a new product classification, the National Product Classification for Manufacturing Sectors (NPCMS), starting with the survey in 2010. There is an official concordance from the ASICC to the NPCMS, but the mappings are usually not one-to-one so that it is challenging to construct a panel of varieties going beyond the survey in 2009.

list contains information on the completion dates of NHDP road upgrades and is current as of May 31, 2017. I manually digitized these road segments, which are plotted in Figure 2 in blue (Phase I) and in red (Phase II). The GQ network, with a length of 5,846 km, connects four major metropolitan cities of Delhi, Mumbai, Chennai, and Kolkata (marked with yellow stars in the figure) while the NS-EW network, with a length of 7,142 km, connects the North-South ends (Srinagar to Kanyakumari) and the East-West ends (Silchar to Porbandar) of India.

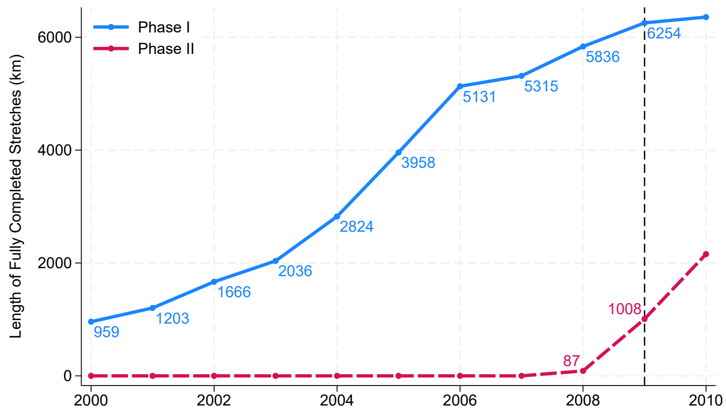
Figure 2: Digitized Road Network of India in 2000



Source: Author's calculation based on data from Survey of India and National Highways Authority of India
Notes: Completed NHDP road stretches (in dark blue) are those completed by March 2000 while incomplete stretches are those completed at a later date. Refer to Appendix Section A.5 for details on the digitization of this map.

Phase I of NHDP was officially approved by the government in December of 2000 and included the development of 7,498 km of four/six lane highway routes. It included 5,846 km of the GQ in its entirety, 981 km of the NS-EW, 356 km of Port Connectivity, and 315 km of other national highways. I exclude the 356 km of road pertaining to the Port Connectivity as these roads focus on access to ports, which is irrelevant for district-to-district road transportation. After this exclusion, the digitized road network in Figure 2 contains 6,976 km of Phase-I routes, which account for 97.7% of the remaining 7,142 km of roads. Phase II of NHDP was officially approved in December of 2003 and covered the development of 6,736 km of highway routes, consisting of the remaining 6,240 km of the NS-EW and 496 km of other national highways. The digitized road network in Figure 2 contains 5,921 km of Phase-II routes, which account for 94.9% of the proposed road length.

Figure 3: Completion Progress of NHDP (2000–2010)



Source: Author's calculation based on data from National Highways Authority of India
Notes: Each data point plots the length of fully completed stretches on a NHDP phase as of March in that year. The series for Phase I omits stretches relating to Port Connectivity projects. Partially completed stretches are not included.

The information on the completion date of the road upgrades allows me to track their progress over time. The completion progress of NHDP road upgrades from 2000 to 2010 is plotted in Figure 3. Given the time span of the available data on manufacturing, the primary focus of this study will be on changes in trade costs resulting from roads upgraded between March of 2000 and 2009. Figure 3 shows that a total of 6,303 km of roads were upgraded between March of 2000 and 2009. This covers 5,295 km of roads under Phase I and 1,008 km of road under Phase II. Note that the data source and hence these numbers I report only account for road stretches that have been fully completed and opened. Other official progress reports typically include partial progress on incomplete stretches of road upgrades into their statistics for completed road lengths and are higher. For the purposes of calculating trade costs, it makes sense to discard only partially completed road upgrades since they would not be open to transportation. Information on these road upgrades have been included into the digitized road network in 2001 from earlier. For example, Figure 2 depicts in dark blue 1,203 km of Phase-I road upgrades that were already completed by March of 2001, which is the end date of the first year of the ASI survey I use (i.e., the 2000–01 ASI).

Although subsequent phases of the project were approved during the sample period of this study, there were little to no progress made on them by March of 2009. Only 157 km of road stretches under Phase III have been fully completed by March of 2009 while work on later phases of the project had not begun yet. Given the timeline of their progress, these phases of the NHDP are not considered in this study.

3.4 Trade Costs between Districts

For this study, I focus on changes in trade costs resulting from NHDP road upgrades between March 2000 and 2009. The earliest manufacturing data I have cover activities from April

2000 to March 2001 (ASI 2000-01) and the latest data cover activities from April 2009 to March 2010 (ASI 2009-10). Thus, the road upgrades that occurred between March 2000 and 2009 capture concurrent changes in trade costs. In order to measure changes in trade costs between these periods, I first digitized a scanned road map of India published in 2011 by Survey of India, India's national mapping agency under the Department of Science and Technology. Although the map is published in 2011, it depicts India's road network around 2000, which can be confirmed by comparing the total length and list of national highways in the map against official sources. Further details on the digitization and verification of the road map are provided in the Appendix Section A.5.

The digitized road network of India in 2000 is plotted in Figure 2 and contains five categories of roads: (1) 959 km of NHDP Phase-I roads completed as of March 2000; (2) 6,017 km of NHDP Phase-I roads yet to be completed as of March 2000; (3) 5,921 km of incomplete NHDP Phase-II roads; (4) 53,731 km of non-NHDP national highways; and (5) 97,789 km of other roads of major importance. The 54,934 km of digitized national highways in 2001 (1,203 km of NHDP roads plus 53,731 km of other national highways) account for 95.1% of the 57,737 km of total length of national highways as of March 31, 2001 reported by the government (NHAI, 2001).

To calculate minimum travel times between district centroids, I assume road speeds of 55 km/h for upgraded NHDP roads, 35 km/h for national highways that are not upgraded, 25 km/h for other major roads, and 10 km/h for out-of-network travel. The choice of 55 km/h is based on an evaluation report on a section of the GQ network, which was upgraded from 2-lane to 4-lane roads (Miyazaki, 2006). The report found that average vehicle speed on the section increased from 35 km/h to 55 km/h. These numbers are somewhat lower than to those used by Alder (2016), who also examined driving speeds on the same Indian roads. Note that almost all NHDP roads were national highways prior to the upgrades so the NHDP increased road speeds by more than half in most cases.

Applying these road speeds to the digitized road network in 2000, I calculate the travel times between every pair of district centroids using an implementation of Dijkstra's algorithm in QGIS for finding shortest paths (Raffler, 2018). This is done for each year $t \in [2000, 2009]$ after setting the road speeds on NHDP roads that have been upgraded by March of year t to 70 km/h. These travel times are then converted into ad-valorem trade costs between districts o and d in year t using the following equation:

$$\tau_{odt} = \left(\zeta_1 \{\text{travel time between district (hours)}\}_{odt}\right)^{\zeta_0}. \tag{22}$$

Following Redding and Rossi-Hansberg (2017), the exponent $\zeta_0 = 0.375$ is chosen so that $-(\sigma - 1)\zeta_0 = -1.5$, which is the elasticity of trade flows with respect to distance in inter-regional data when $\sigma = 5$.[13] I assume a within-district travel time of 30 minutes and set $\zeta_1 = 2$ so that within-district trade costs are $\tau_{oo} = 1$ for all districts.

3.5 Measuring Market Access and Product Appeal

First, following equation (12), I use data on consumer expenditure (E_{idt}), populations (N_{idt}), and bilateral trade costs (τ_{odt}) with an assumed value of $\sigma = 5$ to measure the market access

[13]See Allen and Arkolakis (2022) and Baragwanath Vogel et al. (2024) for similar applications on converting travel times into trade costs.

to consumer type-i as:

$$\text{MA}^E_{iot} = \sum_{d \in D} \tau^{1-\sigma}_{odt} E^{\sigma}_{id,2001} N_{id,2001}. \tag{23}$$

Note that I hold consumer expenditures and populations fixed at their 2001 levels to avoid introducing simultaneity biases when I use the variable in a regression. Ideally, I would have used levels of these variables from 2000 but data were not available. In my sample, higher-income consumers are about 1.5 times richer than the national average while lower-income consumers are about 0.6 times poorer than the average. Thus, we can interpret market access to higher-income consumers (MA^E_{Hot}) as access to a market that is about 1.5 times richer than the average. Similarly, we can interpret market access to lower-income consumers (MA^E_{Lot}) as access to a market that is about 0.6 times poorer.

Then, using these market access terms and equation (17), I measure each district's total market access and access to higher-income markets (AHIM) as

$$\text{MA}^E_{ot} = \sum_{i \in \{H,L\}} \text{MA}^E_{iot} \tag{24}$$

$$\text{AHIM}^E_{ot} = \text{MA}_{Hot}/\text{MA}_{Lot}. \tag{25}$$

Dingel (2017) uses a similar measure of access to higher-income markets that instead relies on average district incomes, but his measure is not theoretically consistent with the model in Section 2 so I do not use it here.

Next, I describe the measurement of product appeal (Φ_{vt}) defined in equation (13). First, note that we can write firm revenue as $R_{vt} = p^{1-\sigma}_{vt} \Phi^{\sigma-1}_{vt} \text{MA}_{dt}$. Taking logged differences across time and using the empirical measure of total market access from above, I measure log-changes in product appeal as

$$d \ln \Phi_{vt} = \left(\frac{1}{\sigma - 1}\right) d \ln R_{vt} + d \ln p_{vt} - \left(\frac{1}{\sigma - 1}\right) d \ln \text{MA}^E_{ot}. \tag{26}$$

4 Empirical Analysis

4.1 Identification Strategy

A goal of this paper is to determine if increased access to domestic higher-income markets causes local firms to upgrade the quality of their output. As discussed in Section 2.4, I make inferences about firms' quality upgrading by examining changes in the marginal cost and product appeal of their outputs. Consider regressing variety-level changes in marginal cost (or product appeal) on the changes in access to higher-income markets:

$$d \ln Y_{v,09-00} = \delta_0 + \delta_1 \cdot d \ln \text{AHIM}_{d,09-00} + \delta_2 \cdot \ln \text{AHIM}_{d,2000} + \text{error}_{v,09-00}, \tag{27}$$

where $dY_{t'-t} = Y_{t'} - Y_t$ for any variable Y_t and I control for the initial level of logged access to higher-income markets. Controlling for the initial level ensures that identification of δ_1 comes from variations over time. An unbiased estimate of δ_1 then gives the effect of changes in access to higher-income markets on the variety-level changes in marginal cost over 10 years. Note that the regression specification in equation (27) is difference-in-differences (DID) in nature as it compares the variety-level changes in marginal cost in districts that saw greater access to higher-income markets to those in districts that saw only small changes. In this context, the changes in access to higher-income districts ($d \ln \mathrm{AHIM}_{d,09-00}$) from the NHDP road upgrades is the "treatment". Accordingly, to obtain an unbiased estimate of δ_1, the treatment variable has to be uncorrelated with any omitted variables that can affect the outcome variable.

A typical concern over studies that examine the effects of infrastructure projects on regional outcomes is that locations far away from the infrastructure are not valid controls for locations close to the infrastructure due to the non-random placement of infrastructures. Planners of highway projects typically design the roads to pass through specific locations in order to achieve certain economic or political goals. In such cases, locations close to the infrastructure likely possess characteristics that differ from those in distant locations. Indeed, in Appendix Section C.1, I show that districts closer to the NHDP network are larger, more educated, and more urbanized. To the extent that these characteristics also affect the outcome variables of interest, they will bias the OLS estimate of δ_1 in equation (27).

To overcome this issue, I limit the sample of districts used for estimation to those intersecting with NHDP roads upgraded by March 2009 while excluding the nodal districts.[14] I follow Datta (2012) and Ghani et al. (2016) in defining the nodal districts as the four targeted districts (Delhi, Mumbai, Chennai, and Kolkata) and additional eight districts (Gurgaon, Faridabad, Ghaziabad, Gautam Buddha Nagar [Noida], Thane, Chandigarh, Hyderabad, and Bangalore) that they argue were likely part of the NHDP by design. This reduces the number of districts in the sample from 570 to 117, which is about a fifth. The identification assumption is then that changes in access to higher-income markets within this estimation sample are uncorrelated with other factors that can affect firms' output quality or productivity. The idea behind this identification strategy is that the direction of change in a location's access to higher-income markets depends on the relative positions of the districts and their incomes not merely on being close to the road. The upgraded roads in a location may very well connect it to low-income regions of the country. The usefulness of this strategy is evident in Table 1, which shows that the existing correlations between changes in access to higher-income markets (2000-09) and observable district characteristics from 2001 in the full sample are gone within the estimation sample.

Note that this identification approach is possible only because I am interested in the effects of access to *higher-income* markets in particular and not the overall effects of the road upgrades. For the most part, existing studies of the NHDP have been interested with the general effects of the upgraded roads. For example, a number of papers look at the distance to the NHDP network as the measure of treatment (Datta, 2012; Ghani et al., 2016; Asturias et al., 2019; Abeberese and Chen, 2022) while others look at the changes in market access (Alder, 2016; Baragwanath Vogel et al., 2024). My identification strategy is difficult to implement with these measures of treatment, because there will be little to no variation left in the distance to the NHDP network or changes in market access within the subset of districts where roads were upgraded. In comparison, proximity to the upgraded roads is not by itself informative about the location's access to higher-income markets. If the upgraded roads lower trade costs with poorer regions, the location's access to higher-income markets will decrease. This is why there remains substantial variation in the changes in access to higher-income markets even within the subset of districts with upgraded roads. In Appendix Section C.1, I also discuss why I do not rely on hypothetical least-cost paths

[14]To be precise, a district intersects with NHDP roads upgraded by March 2009 if the ellipsoidal distance from the district's boundary (or any interior point) to an upgraded stretch is zero.

Table 1: Correlation Between Changes in Access to Higher-Income Markets and District Characteristics

	Full Sample				Estimation Sample			
	(1) Exp	(2) Pop	(3) HS	(4) Urb	(5) Exp	(6) Pop	(7) HS	(8) Urb
$d \ln \text{AHIM}^E_{d,09-00}$	-0.75***	2.87***	0.61	1.63**	-0.18	-0.38	0.34	0.30
	(0.28)	(0.93)	(0.40)	(0.65)	(0.33)	(0.85)	(0.48)	(0.82)
$\ln \text{AHIM}^E_{d,2000}$	0.27***	0.14*	0.34***	0.57***	0.28***	-0.18**	0.32***	0.44***
	(0.02)	(0.07)	(0.03)	(0.05)	(0.03)	(0.09)	(0.05)	(0.09)
R^2	0.33	0.02	0.20	0.21	0.41	0.04	0.27	0.20
N	570	570	570	563	117	117	117	117

Notes: The outcome variables are district-level (d) logged per-capita expenditure (Exp), population (Pop), share of population with high school degrees (HS), and share of urban population (Urb) in 2001. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

to obtain identification (Faber, 2014). In short, I find that the distance to least-cost paths that connect the four targeted cities are correlated with a number of district characteristics.

Figure 4 plots the spatial distribution of changes in access to higher-income markets from 2000 to 2009 for the districts in the estimation sample along with the locations of NHDP roads upgraded by March 2009. In the top quartile, districts saw 10% to 27% increases in their access to higher-income markets as a result of the road upgrades. In the bottom quartile, districts saw their access to higher-income markets decline by up to 7%. This is a consequence of seeing relatively larger reductions in trade costs with lower-income districts. To obtain a better sense of the magnitudes of these shocks, I use equation (25) to consider the change in access to higher-income markets arising from a hypothetical uniform change in trade costs on the higher-income markets:

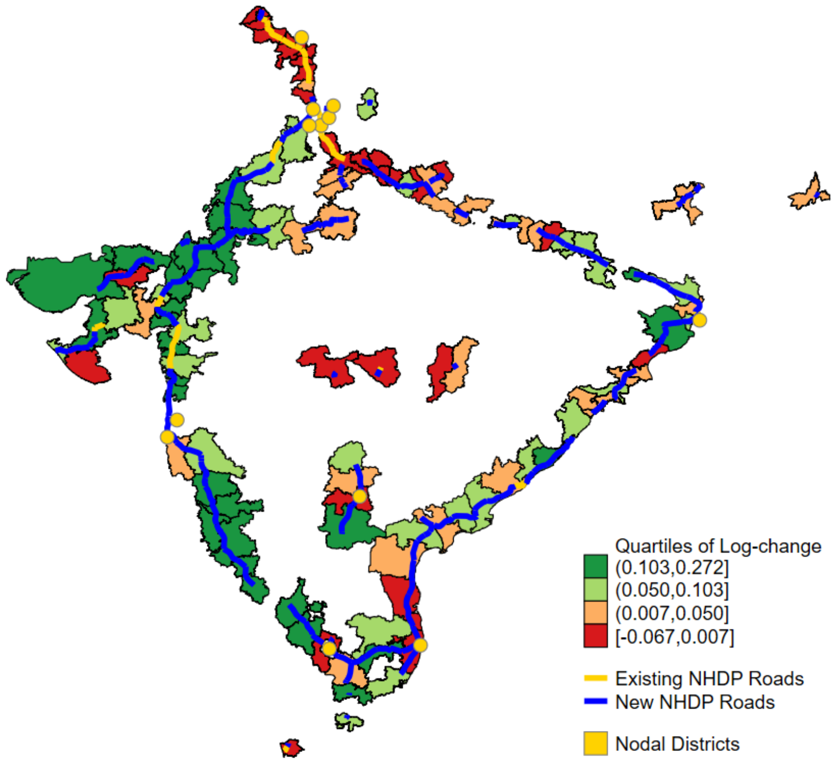
$$d \ln \text{AHIM}_{ot}(\lambda) = \ln \left(\frac{\lambda^{1-\sigma} \text{MA}_{Hot}}{\text{MA}_{Lot}} \right) - \ln \left(\frac{\text{MA}_{Hot}}{\text{MA}_{Hot}} \right) = (1-\sigma) \ln \lambda, \tag{28}$$

where the trade cost with higher-income markets are now $\lambda \tau_{odt}$. Rearranging the equation, I can calculate the equivalent change in trade costs with higher-income markets for any observed change in access to higher-income markets as $\lambda = \exp\left(\frac{d \ln \text{AHIM}_{ot}}{1-\sigma}\right)$. For example, this equation translates the average increase in access to higher-income markets (0.06) in the estimation sample into a 1.5% reduction in trade costs with higher-income markets. The largest shock in Figure 4 (0.27) is equivalent to a 6.5% reduction in trade costs with higher-income markets. So, there are sizable variations in the treatment variable even within the subset of districts intersecting with upgraded roads.

4.2 Effects on Marginal Cost and Product Appeal

Using this estimation sample, I first look at how variety-level changes in marginal cost and product appeal vary with the changes in access to higher-income markets. Figure 5 presents the local polynomial smoothed plots on how variety-level changes in marginal cost (blue) and product appeal (green) relate to the district-level changes in access to higher income markets. The figure shows positive linear relationships for both marginal cost and product

Figure 4: Spatial Distribution of Changes in Access to Higher-Income Markets

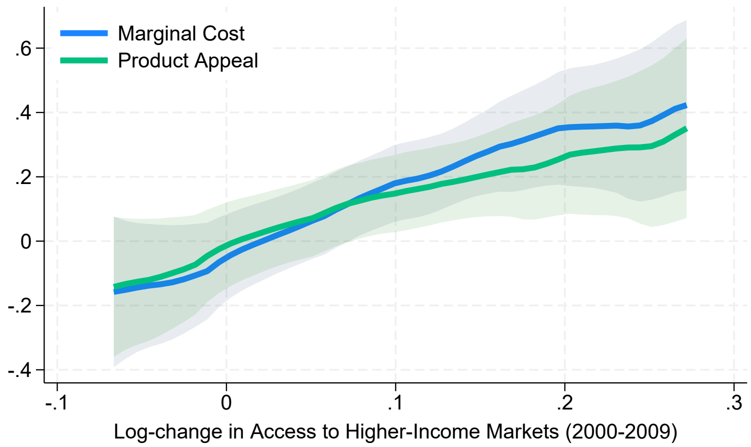


Notes: This figure plots the spatial distribution of log-changes in access to higher-income markets from March 2000 to 2009 as defined in equation (25) for the 117 districts in the estimation sample. The estimation sample consists of districts that intersect with NHDP roads upgraded by March 2009, which are depicted in yellow (those completed by March 2000) and blue lines (those completed between March 2000 and 2009), excluding the twelve nodal districts that are marked with yellow circles (Delhi, Mumbai, Chennai, Kolkata, Gurgaon, Faridabad, Ghaziabad, Gautam Buddha Nagar [Noida], Thane, Chandigarh, Hyderabad, and Bangalore).

appeal. In particular, there are no patterns that suggest firms responded differently based on the size of the shocks. I now proceed to estimating the regression equation in (27) through OLS and obtain the effects of access to higher-income markets on marginal cost (mc_{vt}), materials per output (M/y_{vt}), labor per output (L/y_{vt}), and product appeal (Φ_{vt}).

Table 2 reports the estimates for the baseline specification in equation (27). The first column estimates that a percent increase in access to higher-income markets caused local firms to spend 2.8% more on materials and labor on each unit of output relative to firms in unaffected locations. A standard deviation increase in access to higher-income markets (0.07) translates to 20% higher increases in marginal costs. Recall that a standard deviation increase is equivalent to a 1.7% reduction in trade costs with higher-income markets. The second and the third columns look at the effects of access to higher-income markets on per-unit material and labor use, respectively. I estimate that a percent increase in access to higher-income markets caused local firms to use 2.5% more materials and 2.4% more labor on each unit of output relative to firms in unaffected locations. Thus, one way these firms

Figure 5: Changes in Marginal Cost and Product Appeal Against Changes in Access to Higher-Income Markets



Notes: The series are the local polynomial smoothed plots of variety-level log-changes (2000 to 2009) in marginal cost (blue) and product appeal (green) against the concurrent district-level log-changes in access to higher-income markets (see equation (25)). The shaded regions are the 95% confidence intervals.

spent more on materials and labor was by using greater quantities of inputs on each unit of output, or by increasing the input intensity as in the model. The fourth column looks at changes in product appeal and finds a relative effect of 2.2% from a percent increase in access to higher-income markets. All estimates are statistically significant.

The results in Table 2 indicate that increased access to higher-income markets caused local firms to increase their expenditure on the production of each unit of output (column one) and that these outputs grew more appealing to consumers (column four) relative to the varieties produced in unaffected locations. These patterns are consistent with the model's prediction that increased access to higher-income markets will lead to increases in quality, which in turn will raise marginal costs and product appeal. Of course, there could be alternative explanations to these patterns. For example, increases in marginal cost can be rationalized by decreases in productivity (Garcia-Marin and Voigtländer, 2019) instead. However, if increases in access to higher-income markets are actually decreasing production efficiency without increasing quality, it would not explain why product appeal is increasing. The key idea here is that the fact that both changes in marginal cost and product appeal point to the same conclusion makes a compelling case that quality upgrading took place.

Columns two and three then indicate that one way in which these firms upgraded output quality was by using more quantities of materials and labor for each unit of output, or by increasing input intensity to use the terminology in the model. Atkin et al. (2017b) had similarly found that production of higher-quality rugs required longer hours of labor. Although it is not the main focus of this study, the results here show that this pattern generalizes to many other manufactured output as well. The fact that production of higher-quality output requires more quantities of input for a broad set of products provides further support for Atkin et al. (2019) and Verhoogen (2023) in their argument that quantity based measures of productivity (e.g., TFPQ) are not valid solutions to the shortcomings of revenue-

Table 2: Effects of Access to Higher-Income Markets on Marginal Cost, Input Quantities, and Product Appeal

	Log-change (2000-09):			
	(1)	(2)	(3)	(4)
	mc_v	M/y_v	L/y_v	Φ_v
$d\ln \mathrm{AHIM}^E_{d,09-00}$	2.80***	2.46***	2.36***	2.24***
	(0.70)	(0.72)	(0.79)	(0.75)
$\ln \mathrm{AHIM}^E_{d,2000}$	0.11*	0.04	0.16*	0.10
	(0.07)	(0.07)	(0.08)	(0.07)
N	1104	1104	1104	1104
R^2	0.01	0.01	0.01	0.01

Notes: The variables in the table include variety-level (v) log-changes in marginal cost (mc_{vt}), material per output (M/y_{vt}), labor per output (L/y_{vt}), product appeal (Φ_{vt}), and district-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}). Refer to Section 3 for more details on these variables. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

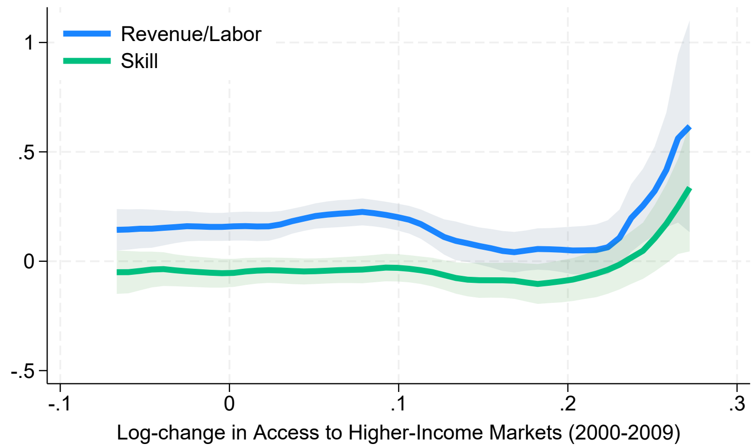
based productivity measures when there are quality differences. A few authors (De Loecker et al., 2016; de Roux et al., 2021) have begun working on incorporating output and input quality into existing methods for estimating production functions and productivity, and the results here indicate that their solutions are relevant for a wide set of products.

In Appendix Section C.2, I test the robustness of the results on marginal cost by using an alternative measure of marginal cost that is recovered using unit prices and markups as in Garcia-Marin and Voigtländer (2019). I find that unit prices show a similar pattern to the marginal costs presented in Table 2 while markups show a negative response to increased access to higher-income markets although at a much smaller magnitude. As a result, the alternative measure of marginal cost I consider show the same pattern as in Table 2. I also test the sensitivity of results on product appeal to the value of the elasticity of substitution, which I assume to be 5 based on the previous literature. I consider values ranging from 3 to 7 and find that the results are robust to these values of the elasticity of substitution.

4.3 Effects on Productivity, Skill, Capital, and Material Prices

I now turn my attention to the effects of access to higher-income markets on productivity. The main measure of productivity that I focus on here is revenue per labor and revenue TFP. In Figure 6, I begin by examining the local polynomial smoothed plots on how plant-level changes in labor productivity (blue, revenue per labor) and skill (green, ratio of non-production to production labor) relate to the district-level changes in access to higher income markets. Unlike in Figure 5, the relationships do not appear linear. When the shocks are small in magnitude, say below 0.2, the relationship is mostly flat for both labor productivity and skill use. Only when the shocks are large, we start to observe relative gains in labor productivity and greater skill use in response to increased access to higher-income markets. The fact that the same shocks identified precise linear relationships in Figure 5 should dismiss doubts that there is insufficient variations in the shocks or that these shocks are imprecisely measured.

Figure 6: Changes in Labor Productivity and Skill Against Changes in Access to Higher-Income Markets



Notes: The series are the local polynomial smoothed plots of plant-level log-changes (2000 to 2009) in revenue per labor (blue) and skill (green, ratio of non-production to production labor) against the concurrent district-level log-changes in access to higher-income markets (see equation (25)). The shaded regions are the 95% confidence intervals.

Motivated by the patterns in Figure 6, I estimate the following regression equation to formally test the differential responses in firm-level outcomes to shocks that are large and small:

$$\begin{aligned} d\ln Y_{f,09-00} = &\ \delta_1^{\text{large}} \cdot d\ln \text{AHIM}_{d,09-00} \times 1\{\text{Large}_d = 1\} \\ &+ \delta_1^{\text{small}} \cdot d\ln \text{AHIM}_{d,09-00} \times 1\{\text{Large}_d = 0\} \\ &+ \delta_0^{\text{small}} + \delta_0^{\text{large}} \cdot \text{Large}_d + \delta_2 \cdot \ln \text{AHIM}_{d,2000} + \text{error}_{f,09-00}, \end{aligned} \tag{29}$$

where $d\ln Y_{f,09-00}$ is the plant-level change (2000-2009) in an outcome variable of interest and Large_d is a dummy variable that equals one if $d\ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. The cutoff for large shocks is determined by minimizing the sum of squared residuals in equation (29) using plant-level changes in revenue per labor as the outcome variable. This cutoff value corresponds to a 4.4% decline in trade costs with higher-income markets. The regression specification in equation (29) aims to estimate separate linear relationships between the outcome variable and the access to higher-income markets when shocks are small ($\text{Large}_d = 0$) and when shocks are large ($\text{Large}_d = 1$). The difference in the coefficients ($\delta_1^{\text{large}} - \delta_1^{\text{small}}$) then tells us if there was a differential response to the shocks based on their sizes.

Using the regression specification in equation (29), I first look at plant-level changes in revenue per labor (R/L_{ft}) and report the results under the first column of Table 3. I estimate a zero effect for small shocks but estimate that each percentage point increase in access to higher-income markets beyond a 18% increase leads local firms to generate 9.5% higher revenue per labor compared to unaffected firms. The difference in effects for large and

Table 3: Effects of Access to Higher-Income Markets on Plant Productivity

	Log-change (2000-09):			
	(1)	(2)	(3)	(4)
	R/L_f	R_f	L_f	TFP_f^R
Large_d=0 × $d \ln \text{AHIM}^E_{d,09-00}$	-0.02	-0.03	-0.01	-0.19
	(0.56)	(0.87)	(0.54)	(0.35)
Large_d=1 × $d \ln \text{AHIM}^E_{d,09-00}$	9.48***	-0.71	-10.19***	3.25***
	(1.56)	(1.69)	(0.90)	(1.03)
$\ln \text{AHIM}^E_{d,2000}$	-0.01	0.03	0.04	0.02
	(0.04)	(0.06)	(0.04)	(0.03)
Large_d	-2.03***	0.08	2.11***	-0.69***
	(0.32)	(0.34)	(0.19)	(0.20)
$\delta_1^{\text{large}} - \delta_1^{\text{small}}$	9.5***	-.68	-10.18***	3.44***
	(1.52)	(1.42)	(.72)	(1.05)
N	1452	1452	1452	1452
R^2	0.01	0.00	0.01	0.00

Notes: The variables in the table include plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue (R_{ft}), total labor (L_{ft}), revenue TFP (TFP^R_{ft}), and district-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}). Large_d is a dummy variable that equals one if $d \ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Refer to Section 3 for more details on these variables. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

small shocks ($\delta_1^{\text{large}} - \delta_1^{\text{small}}$) is 9.5, which is statistically significant. For the largest shock in my data (0.27), these estimates imply 54% higher growth in revenue per labor. Recall that the largest shock is comparable to a 6.5% reduction in trade costs with the higher-income markets. In the second and third columns, I decompose these effects on revenue per labor into effects on revenue and on labor, respectively. For both small and large shocks, the decomposition reveals effects on local firms' revenue that are close to zero. The third column shows that almost all of the effects on labor productivity are coming from the local firms' ability to reduce total labor while retaining their revenues. The fourth column estimates a similar pattern for revenue TFP, but the magnitude of the effects are about third of that on labor productivity. For instance, the largest shock translates to a 19% higher growth in revenue TFP.

To better ascertain the sources of these productivity gains, I now estimate how large and small increases in access to higher-income markets affected local firms' skill use (ratio of non-production to production labor), value of capital, and material input prices. The first column of Table 4 estimates that each percentage point increase in access to higher-income markets beyond a 18% increase leads local firms to use 7.2% more non-production labor to production labor relative to firms in unaffected locations. The difference in effects for large and small shocks ($\delta_1^{\text{large}} - \delta_1^{\text{small}}$) is 7.45, which is statistically significant. Based on this estimate, the largest shock in my data implies a 58% relative increase in skill use. These effects on skill upgrading are consistent with previous findings of Verhoogen (2008) and Brambilla et al. (2012). The estimates in the second column indicate that the effects on the value of capital were of similar magnitudes. Together with the observation that the total labor decreased with increased access to higher-income markets in the third column of Table 3, the result on capital implies that production at these firms became more capital intensive. Finally, the third column estimates that the largest increase in access to higher-income markets in the data caused local firms to use 20% more expensive materials relative to unaffected firms. These results on material input prices are consistent with those in Bastos et al. (2018) and Hansman et al. (2020), providing at least suggestive evidence that these firms used higher-quality materials. For small shocks, the estimated effects are all close to zero.

Table 4: Effects of Access to Higher-Income Markets on Skill, Capital, and Material Prices

	Log-change (2000-09):		
	(1)	(2)	(3)
	s_f	K_f	p_{Mf}
Large_d=0 $\times$ $d \ln \text{AHIM}^E_{d,09-00}$	-0.22	0.02	-0.06
	(0.50)	(0.56)	(0.40)
Large_d=1 $\times$ $d \ln \text{AHIM}^E_{d,09-00}$	7.23***	7.87**	2.78***
	(1.65)	(3.64)	(0.94)
$\ln \text{AHIM}^E_{d,2000}$	-0.00	-0.03	0.03
	(0.03)	(0.03)	(0.02)
Large_d	-1.43***	-1.75**	-0.55***
	(0.31)	(0.70)	(0.20)
$\delta_1^{\text{large}} - \delta_1^{\text{small}}$	7.45***	7.85**	2.84***
	(1.66)	(3.63)	(1.03)
N	1452	1452	1452
R^2	0.00	0.01	0.00

Notes: The variables in the table include plant-level (f) log-changes in skill (s_{ft}, ratio of non-production to production labor), value of capital (K_{ft}), material prices (p_{Mft}), and district-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}). Large_d is a dummy variable that equals one if $d \ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Refer to Section 3 for more details on these variables. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

With large increases in access to higher-income markets, the results in Tables 3 and 4 indicate that local firms experienced relatively higher growth in productivity, which was achieved through increased uses of skilled labor, capital, and more expensive materials. As for small increases in access to higher-income markets, local firms did not experience any productivity effects and there are no visible signs that they introduced any changes to their production processes. So, together with the results in Table 2, these results suggest that local firms upgraded their output quality only through using more quantities of inputs when shocks were small.

These empirical findings are broadly consistent with the predictions of the model in Section 2. When the shocks are small, firms find it too costly to reorganize their production process by paying the adjustment cost. Instead, they resort to adjusting their input intensity to achieve their desired output quality. However, increasing input intensity has no implications on labor productivity and skill use. This explains why we continue to observe increases in marginal cost and product appeal even for small shocks but do not observe any responses in labor productivity and skill use. In comparison, when shocks are large, firms find it profitable to pay the adjustment cost so that they can better optimize their production process. In this case, higher output quality is achieved both through increasing input intensity and sophistication. As a result, we observe increases in labor productivity and skill use as well as increases in marginal cost and product appeal in response to increased access to higher-income markets.

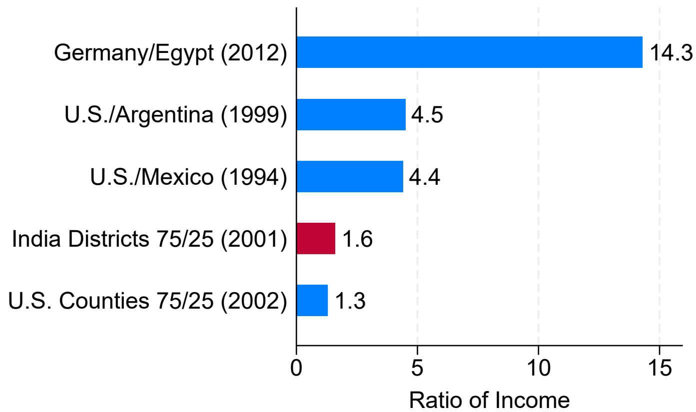
4.4 Discussion of Findings and Policy Implications

My first finding is that increases in access to higher-income markets led to relatively higher increases in marginal cost and product appeal, which I interpret as signs that output quality increased. This finding is consistent with a growing body of evidence (see Verhoogen (2023) for a more extensive review) that increased access to higher-income markets causes quality upgrading. At the same time, a key difference in the current setting from those in previous studies is that I examine much narrower income differences within India. For the most part, previous studies have focused on how variations in access to higher-income countries affect local quality upgrading. For example, Verhoogen (2008), Brambilla et al. (2012), and Bastos et al. (2018) all take advantage of exchange rate shocks while Atkin et al. (2017b) examined sales from Egypt to buyers in developed countries like Germany. Figure 7 plots selected examples of income differences from a number of these studies as well as the regional income differences within India and the U.S. These ratios confirm that international income differences are many times larger than income differences found within the borders of a country. Thus, the fact that local firms upgraded their output quality in response to increased access to domestic higher-income markets in India suggests that developing countries do not need to rely solely on countries with income that are many times higher than theirs to promote local quality upgrading.

My second finding is that local firms responded differently to small and large shocks in access to higher-income markets. Large shocks induced local firms to use more skilled labor, acquire more capital, and use more expensive materials, which then showed up as higher labor productivity and revenue TFP growth. The identified effects for large shocks are consistent with past findings that increased access to higher-income markets cause skill upgrading (Verhoogen, 2008; Brambilla et al., 2012), use of more expensive material inputs (Bastos et al., 2018; Hansman et al., 2020), and productivity gains (Atkin et al., 2017b). Nevertheless, for most Indian firms that faced smaller shocks, their quality upgrading did not lead to gains in productivity nor increased uses of skilled labor, capital, and more expensive materials. Instead, these firms responded by increasing their output quality through using more quantities of input per output.

Seen through the lens of model, adjustment costs associated with adopting more sophisticated production processes led most Indian firms to upgrade quality through increasing input intensity as opposed to sophistication. As a result, they did not experience the productivity gains associated with quality upgrading through increasing sophistication. To the extent that productivity gains are the end goal of promoting quality upgrading, my empirical findings suggest that small reductions in trade costs with only narrowly higher-income markets are insufficient. On the one hand, these results justify the emphasis on access to

Figure 7: Comparison of International and Domestic Regional Income Differences



Notes: This plot is a bar chart of selected ratios of income. For countries, income is measured by the per-capita GDP from the World Bank. For Indian Districts, income is measured by the per-capita expenditure from the Household Consumer Expenditure Survey. For U.S. counties, income is measured with the per-capita personal income from the Bureau of Economic Analysis. "75/25" denotes the ratio of the 75th and 25th percentiles of income.

high-income countries in promoting local economic development as the wider income difference can generate sufficiently strong incentives for local firms to upgrade output quality through increasing sophistication. On the other hand, these results also present possibilities for developing countries to look outside high-income countries to achieve local economic development through quality upgrading. Specifically, the analysis in this paper suggests that there should either be a large reduction in trade costs with narrowly higher-income markets or a lowering of the costs that firms face in adopting more sophisticated production processes in order to achieve productivity gains associated with quality upgrading.

5 Additional Results and Robustness Tests

In this section, I begin with an analysis of the time dynamics in the estimated effects from the previous section. By doing so, I assess the timing of the effects against the completion progress of the NHDP road upgrades in Figure 3 and check for signs of pre-existing trends. Next, I address a number of remaining concerns over the main empirical specification. Section 5.2 confirms that the results in Section 4 are driven by changes in access to higher-income markets as opposed to larger markets. Section 5.3 considers the possibility that changes in access to higher-income markets are correlated with changes in access to higher-quality materials and skilled labor. Section 5.4 assesses the impact of competition on the results found in this paper. Finally, Section 5.5 examines the difference in access to in-state and out-of-state higher-income markets in driving the results.

5.1 Time Dynamics in Estimated Effects

I examine the time dynamics in the estimated effects of access to higher-income markets on the main outcome variables from Section 4. The main goal behind this exercise is to gain confidence over the identification strategy by checking for pre-existing trends and confirming that the timing of the estimated effects align with the progress of the NHDP in Figure 3. Recall that the identification assumption I rely on is that across districts intersecting with the upgraded roads (excluding the nodal districts), changes in access to higher-income markets are uncorrelated with unobserved variables that can affect local firms' output quality or productivity. For this assumption to hold, it should be the case that pre-existing trends in the outcome variables are not correlated with the changes in access to higher-income markets. Since I am examining road upgrades that took place from 2000 to 2009, an assessment of pre-existing trends would typically require data on the outcome variables prior to 2000 (e.g., variety-level changes in marginal cost from 1998 to 1999). Unfortunately, the earliest year of data I have is from 2000. Nevertheless, I can have an indirect assessment of pre-existing trends by taking advantage of the fact that relatively little road upgrades were completed during the first few years of the NHDP as shown in Figure 3.

Specifically, I estimate the baseline specification in equation (27) separately for changes in the outcome variables from 2000 to each year $t \in [2001, 2009]$:

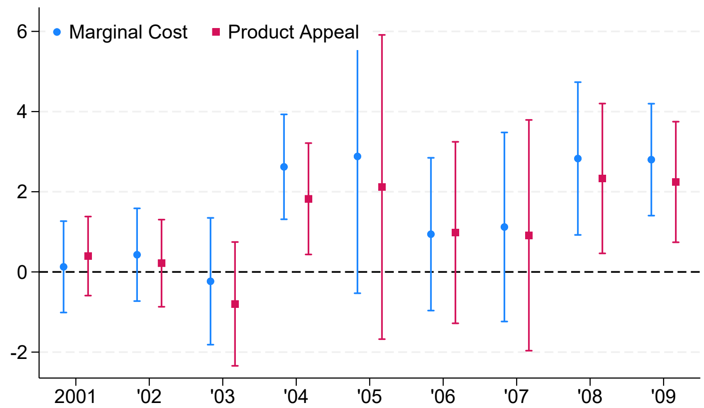
$$d \ln Y_{v,t-00} = \delta_{0t} + \delta_{1t} \cdot d \ln \mathrm{AHIM}_{d,09-00} + \delta_{2t} \cdot \ln \mathrm{AHIM}_{d,2000} + \mathrm{error}_{v,t-00}, \qquad (30)$$

where Y_{vt} is either the marginal cost or product appeal. For example, the estimate of $\delta_{1,2001}$ then captures the effects of changes in access to higher-income markets from 2000 to 2009 on changes in the outcome variable (e.g., marginal cost) from 2000 to 2001. Figure 3 shows that only a few hundred kilometers of roads were upgraded during this time and so I do not expect to see large effects here. If the regression specification is picking up pre-existing trends, we might see large effects early on or sustained increases in effects that do not align with the progress of the road upgrades. Figure 3 shows that a majority of the road upgrades analyzed in this paper were completed around 2004–2005 so I expect the effects to show primarily around that time and get to the numbers reported in Section 4 in 2009.

Figure 8 plots the estimates of δ_{1t} for each year $t \in [2001, 2009]$ and for changes in marginal cost (blue) and product appeal (red). Note that the coefficients for 2009 are the same as those in Table 2. In accordance with the timing of the NHDP progress in Figure 3, both the estimates for marginal cost and product appeal indicate null effects until 2003. There are statistically significant positive effects in 2004 but the estimates become imprecise, only fully showing up in 2008 and 2009. These patterns confirm that the empirical specification did not pick up pre-trends in the outcome variables and that the estimates become more precise towards the later years of the data, which lends support to the validity of the identification assumption for these two variables.

Next, I turn my attention to the dynamics in the effects of access to higher-income markets on productivity, skill, capital, and material prices. With these variables, I continue to estimate differential effects for large and small shocks. That is, I estimate the baseline specification in equation (29) separately for changes in the outcome variables from 2000 to

Figure 8: Cumulative Estimated Effects on Marginal Cost and Product Appeal



Notes: This graph plots the estimates of δ_{1t} in equation (30) for years $t \in [2001, 2009]$. The outcome variables are variety-level changes in marginal cost (blue) and product appeal (red). The bars denote the 95% confidence intervals based on standard errors clustered at the district-level.

each year $t \in [2001, 2009]$:

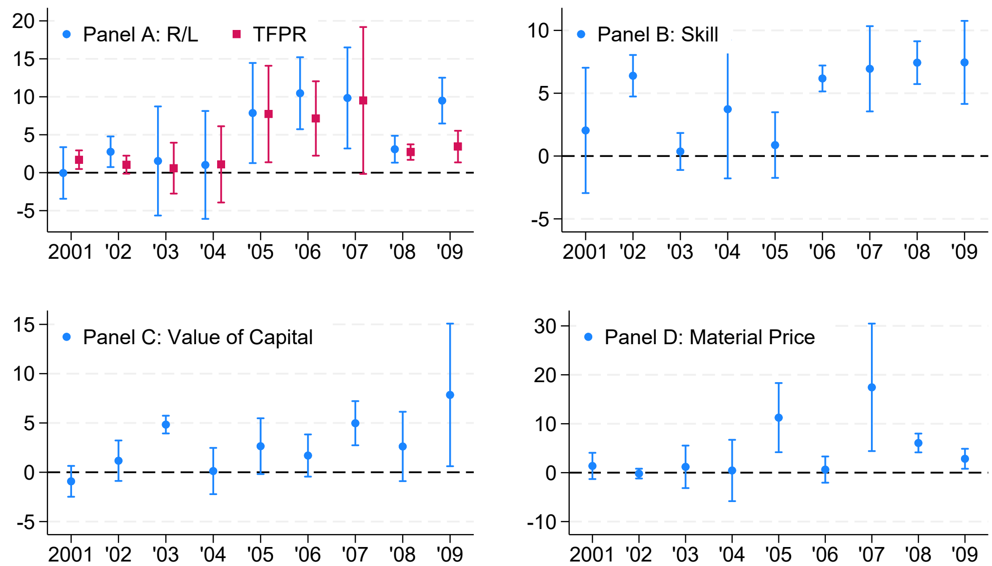
$$\begin{aligned} d\ln Y_{f,t-00} = & \ \delta_{1t}^{\text{large}} \cdot d\ln \text{AHIM}_{d,09-00} \times 1\{\text{Large}_d = 1\} \\ & + \delta_{1t}^{\text{small}} \cdot d\ln \text{AHIM}_{d,09-00} \times 1\{\text{Large}_d = 0\} \\ & + \delta_{0t}^{\text{small}} + \delta_{0t}^{\text{large}} \cdot \text{Large}_d + \delta_{2t} \cdot \ln \text{AHIM}_{d,2000} + \text{error}_{f,t-00}, \end{aligned} \tag{31}$$

where Y_{ft} is either revenue per labor, revenue TFP, skill, value of capital, or material price. The estimates of the difference $\delta_{1t}^{\text{large}} - \delta_{1t}^{\text{small}}$ for each year $t \in [2001, 2009]$ are reported in 9. Panel A reports the estimates for revenue per labor (blue) and revenue TFP (red). Panels B, C, and D report the estimates for skill, value of capital, and material prices, respectively.

The estimates are rather noisy as a result of there being considerably fewer firms experiencing large increases in access to higher-income markets by 2009. However, patterns similar to those in Figure 8 emerge. Panel A shows that effects on revenue per labor shows no visible changes until 2004 and show up starting from 2005. As for revenue TFP, there in an initial deviation away from zero in 2001 but it quickly disappears until 2004 and the effects really starting showing up starting in 2005. Compared to revenue per labor, it shows a decline in 2008–2009, which reflects the decline in material prices in Panel D in those years. Panel B shows a mostly flat effect on skill use until 2005, with a large one-time deviation in 2002, and show a large increase in 2006 and smaller increases onward. Panel C shows a less visible increase in value of capital starting in 2005. Panel D shows a flat response until 2004 and increases in later years but the patterns are especially noisy.

Overall, it does not appear that the regression specifications picked up on pre-existing trends in the outcome variables as the dynamics reveal that the estimated effects are mostly flat in the earlier years of the data, only showing up once significant shares of the road upgrades have been completed. At the same time, the dynamics for capital and material prices in Panels C and D appear less sharply aligned with the progress of the NHDP road upgrades. So, the estimates on these variable may have to be read with more caution.

Figure 9: Cumulative Estimated Effects on Productivity, Skill, Capital, and Material Price



Notes: This graph plots the estimates of $\delta_{1t}^{\text{large}} - \delta_{1t}^{\text{small}}$ in equation (31) for years $t \in [2001, 2009]$. The outcome variables are plant-level changes in revenue per labor (Panel A, blue), revenue TFP (Panel A, red), skill (Panel B), value of capital (Panel C), and material prices (Panel D). The bars denote the 95% confidence intervals based on standard errors clustered at the district-level.

5.2 Controlling for Total Market Access Changes

The NHDP road upgrades caused reductions in trade costs across India and led to changes in access to higher-income markets. Of course, it also led to changes in total market access. There is a large literature (Allen and Arkolakis, 2014; Donaldson and Hornbeck, 2016; Donaldson, 2018) that studies the effects of market access on various economic outcomes by taking advantage of large infrastructure projects such as the NHDP. In fact, a number of them have studied the NHDP in India (Ghani et al., 2016; Baragwanath Vogel et al., 2024). The focus of this paper is not on access to larger markets but on the access to *higher-income* markets and how it affects the quality upgrading behavior of firms. The model in Section 2 shows that what induces firms to upgrade quality is greater demand from high-income consumers who have stronger preferences for quality not larger demand in general. To show that changes in access to higher-income market are driving the results in this paper as opposed to the changes in general market access, I include the changes in total market access as an additional control to the baseline specifications in equations (27) and (29). The resulting estimates are reported in Table 5.

The point estimates on $d \ln \text{AHIM}_{d,09-00}^{E}$ in Table 5 are similar to those in Section 4, which show that the results in this paper are driven by changes in access to higher-income markets as opposed to larger markets. The estimated effects of changes in total market access are statistically indistinguishable from zero in all but one outcome variable, changes in capital. For changes in capital, I estimate that changes in total market access had a

Table 5: Comparison of Effects of Market Access and Access to Higher-Income Markets

	Log-change (2000-09):						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	mc_v	Φ_v	R/L_f	TFP_f^R	s_f	K_f	p_{Mf}
$d\ln \text{AHIM}_{d,09-00}^E$	3.92***	3.59***					
	(0.98)	(1.06)					
$\ln \text{AHIM}_{d,2000}^E$	0.12*	0.10*	-0.01	0.02	-0.00	-0.03	0.03
	(0.06)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
$d\ln \text{MA}_{d,09-00}^E$	-1.01	-1.22	-0.71	-0.24	0.00	-0.67*	0.16
	(1.02)	(1.11)	(0.48)	(0.37)	(0.53)	(0.38)	(0.44)
$\text{Large}_d{=}0 \times d\ln \text{AHIM}_{d,09-00}^E$			0.76	0.07	-0.23	0.76	-0.24
			(0.79)	(0.57)	(0.90)	(0.63)	(0.70)
$\text{Large}_d{=}1 \times d\ln \text{AHIM}_{d,09-00}^E$			8.13***	2.80**	7.23***	6.61**	3.08**
			(1.08)	(1.13)	(1.85)	(2.59)	(1.28)
Large_d			-1.55***	-0.54*	-1.43***	-1.31**	-0.65*
			(0.33)	(0.28)	(0.46)	(0.52)	(0.36)
$\delta_1^{\text{large}} - \delta_1^{\text{small}}$			7.37***	2.73*	7.46***	5.85**	3.31*
			(1.52)	(1.39)	(2.32)	(2.65)	(1.73)
N	1104	1104	1452	1452	1452	1452	1452
R^2	0.02	0.01	0.01	0.00	0.00	0.01	0.00

Notes: The outcome variables are variety-level (v) log-changes in marginal cost (mc_{vt}) and product appeal (Φ_{vt}); plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue TFP (TFP_{ft}^R), skill (s_{ft}, ratio of non-production to production labor), value of capital (K_{ft}), and material prices (p_{Mft}) from year 2000 to 2009. Other variables include district-level (d) log-changes in access to higher-income markets (AHIM_{dt}^E) and total market access (MA_{dt}^E), and a dummy variable (Large_d) that equals one if $d\ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Refer to Section 3 for more details on these variables. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

small negative effect that is only statistically significant at the 10% level. One reason that changes in total market access may be playing such a small quantitative role here is that the estimation sample is restricted to districts where NHDP roads were built. As a result, all of these locations are benefiting from the reduced trade costs and so there may not be enough variation to identify the effects of total market access within this sample. This is in contrast with other papers that study the effects of market access, which typically compare locations close to the infrastructure to distant locations.

5.3 Consideration of Changes in Access to Inputs

One remaining concern is that the changes in access to higher-income markets used in this study may be correlated with changes in access to higher-quality material inputs or skilled labor. Since road upgrades imply bilateral reductions in trade costs, having lower trade costs with higher-income markets may not only mean better access to high-income consumers but also higher-quality material inputs produced there. Indeed, I report in Appendix Section C.4 that higher-income districts produce more expensive varieties within narrowly defined products, which are likely of higher quality. If higher-quality materials and skilled labor are complementary, then increased access to higher-quality materials could provide an alternative explanation to the findings in Section 4. The productivity benefits of higher-quality material inputs are also well-documented in the literature that examined tariff reductions on imported inputs following trade liberalization episodes (Amiti and Konings, 2007; Amiti and Khandelwal, 2013). Reduced travel times may also impact commuting patterns to nearby districts and affect the local availability of skilled labor. For instance, increased access to nearby higher-income markets could be correlated with increased access to skilled labor in these locations through commuting.

Table 6: Effects of Access to Higher-Income Markets with Plant-level Controls

	Log-change (2000-09):					
	(1) mc_v	(2) Φ_v	(3) R/L_f	(4) TFP_f^R	(5) s_f	(6) K_f
$d \ln \text{AHIM}^E_{d,09-00}$	2.68***	2.16***				
	(0.69)	(0.75)				
$\ln \text{AHIM}^E_{d,2000}$	0.15**	0.12*	-0.03	-0.00	-0.03	-0.04
	(0.07)	(0.07)	(0.04)	(0.03)	(0.03)	(0.03)
$d \ln w_{f,09-00}$	0.00	0.06	0.56***	0.22***	0.63***	0.17***
	(0.11)	(0.11)	(0.06)	(0.04)	(0.05)	(0.06)
$d \ln p_{Mf,09-00}$	0.03	0.02	0.02	0.39***	0.02	0.06
	(0.09)	(0.08)	(0.03)	(0.02)	(0.04)	(0.04)
Importer (0 or 1)$_{ft}$	-0.24	-0.23	-0.01	0.01	-0.09**	-0.00
	(0.17)	(0.17)	(0.05)	(0.03)	(0.04)	(0.05)
Exporter (0 or 1)$_{ft}$	-0.12	-0.07	-0.13*	-0.06	0.04	0.00
	(0.10)	(0.14)	(0.07)	(0.04)	(0.07)	(0.07)
Large_d=0 × $d \ln \text{AHIM}^E_{d,09-00}$			-0.18	-0.25	-0.22	0.00
			(0.53)	(0.28)	(0.51)	(0.56)
Large_d=1 × $d \ln \text{AHIM}^E_{d,09-00}$			9.56***	2.18***	7.18***	7.64**
			(1.69)	(0.77)	(1.59)	(3.62)
Large_d			-2.02***	-0.47***	-1.37***	-1.70**
			(0.35)	(0.15)	(0.30)	(0.70)
$\delta_1^{\text{large}} - \delta_1^{\text{small}}$			9.74***	2.43***	7.4***	7.64**
			(1.62)	(.76)	(1.59)	(3.63)
N	1104	1104	1452	1452	1452	1452
R^2	0.02	0.01	0.12	0.25	0.12	0.01

Notes: The variables in the table include variety-level (v) log-changes in marginal cost (mc_{vt}) and product appeal (Φ_{vt}); plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue TFP (TFP_{ft}^R), skill (s_{ft}, ratio of non-production to production labor), value of capital (K_{ft}), wages (w_{ft}), and material prices (p_{Mft}) from year 2000 to 2009. Other variables include district-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}) and a dummy variable (Large_d) that equals one if $d \ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

To address these concerns, I include plant-level changes in factor prices (wages and material prices) as additional controls to the baseline specifications. To the extent that changes in access to higher-quality materials or skilled labor are reflected in the factor prices that plants pay, they will allow me to control for these alternative mechanisms. I also include dummy variables that indicate whether a plant exports or imports. A plant is an exporter if it derived any of its sales that year from exporting. It is an importer if it purchased any imported material in that year. These controls try to address the possibility that the road upgrades may have impacted international trade costs. The estimated results are reported in Table 6. While the inclusion of these additional control variables absorb sizable amount of the variation in the outcome variables, which is reflected in higher values of R^2, they have no impact on the results obtained in Section 4. Therefore, although it is possible that the NHDP road upgrades affected Indian districts' access to higher-quality materials and skilled labor, the results here suggest that they may not be correlated with the changes in access to higher-income markets used in this study.

5.4 Consideration of Changes in Degree of Competition

Another concern is the role of competition that comes with reductions in bilateral trade costs. It is conceivable that changes in access to higher-income markets are correlated with lowering of local prices of higher-quality products that are produced there. For example, in Appendix Section C.2, I find that increased access to higher-income markets led to lowering of markups, which may be reflecting increased competition. To gauge how changes in competition may be affecting the estimates in this paper, I make use of the fact that I observe the production locations of manufactured output. For each 5-digit product g, let $\mathcal{D}_g$ denote the set of districts where product g is produced based on the locations of plants that are producing product g. Then for each firm f, let $\mathcal{G}_f^Y$ denote the set of products g that the firm produces. The set of districts that produce outputs that a firm f produces can then be determined as $\mathcal{D}_f^Y = \cup_{g \in \mathcal{G}_f^Y} \mathcal{D}_g$.

Using this information, I can define a firm-specific measure of access to higher-income markets that do not produce the outputs that the firm produces. The idea is that changes in trade costs to markets that do not produce the same product as the firm does not affect the local degree of competition for the firm. Specifically, consider the following measure of access to higher-income markets to "non-competing" districts:

$$\text{AHIM}_{ft}^N \equiv \frac{\sum_{d \notin \mathcal{D}_f^Y} \tau_{odt}^{1-\sigma} E_{Hd,2001}^{\sigma} N_{Hd,2001}}{\sum_{d \notin \mathcal{D}_f^Y} \tau_{odt}^{1-\sigma} E_{Ld,2001}^{\sigma} N_{Ld,2001}}. \tag{32}$$

The R^2 from regressing $d \ln \text{AHIM}_{d,09-00}^E$ on $d \ln \text{AHIM}_{f,09-00}^N$ using the firms in the estimation sample is 0.30. So, a significant share of the variation in $d \ln \text{AHIM}_{d,09-00}^E$ comes from locations where firms may be facing competition from. Denote the remaining variation in the original changes in access to higher-income markets after subtracting out the changes in access to non-competing higher-income markets as

$$d \ln \text{AHIM}_{ft}^{E-N} \equiv d \ln \text{AHIM}_{dt}^E - d \ln \text{AHIM}_{ft}^N. \tag{33}$$

Then, I rerun the main analysis replacing the original changes in access to higher-income markets with the above two measures of changes in access to higher-income markets to assess the role of competition in the current setting. The results are presented in Table 7. The estimates in the table reveal no visible differences between the variations in AHIM_{ft}^N and AHIM_{ft}^{E-N} in identifying the effects of access to higher-income markets. This suggests that the changes in competition resulting from the NHDP road upgrades did have a large influence on the estimates in Section 4.

5.5 Access to In-state and Out-of-state Higher-Income Markets

Here, I assess whether the effects of access to higher-income markets are coming from locations within or outside state borders. Van Leemput (2021) documents substantial trade

Table 7: Effects of Access to Non-Competing Higher-Income Markets

	Log-change (2000-09):						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	mc_v	Φ_v	R/L_f	TFP_f^R	s_f	K_f	p_{Mf}
$d\ln \text{AHIM}^N_{f,09-00}$	2.91***	2.29***					
	(0.72)	(0.76)					
$d\ln \text{AHIM}^{E-N}_{f,09-00}$	2.41**	2.07*					
	(1.00)	(1.08)					
$\ln \text{AHIM}^E_{d,2000}$	0.09	0.09	-0.01	0.01	-0.00	-0.01	0.02
	(0.07)	(0.07)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)
Large_d=0 × $d\ln \text{AHIM}^N_{f,09-00}$			-0.03	-0.11	-0.24	-0.12	0.08
			(0.56)	(0.34)	(0.52)	(0.64)	(0.40)
Large_d=1 × $d\ln \text{AHIM}^N_{f,09-00}$			9.28***	2.88**	7.40***	7.93**	2.65***
			(1.54)	(1.13)	(1.67)	(3.91)	(1.00)
Large_d=0 × $d\ln \text{AHIM}^{E-N}_{f,09-00}$			-0.10	-0.48	-0.19	0.39	-0.45
			(0.68)	(0.46)	(0.57)	(0.61)	(0.49)
Large_d=1 × $d\ln \text{AHIM}^{E-N}_{f,09-00}$			9.65***	3.28***	7.00***	8.53**	2.12**
			(1.65)	(1.00)	(1.81)	(3.63)	(0.81)
Large_d			-2.02***	-0.65***	-1.43***	-1.82**	-0.47**
			(0.33)	(0.21)	(0.32)	(0.73)	(0.18)
$\delta^{\text{large}}_{1,N} - \delta^{\text{large}}_{1,E-N}$	.49	.23	-.37	-.4	.41	-.6	.54
	(.91)	(.92)	(.34)	(.35)	(.45)	(.86)	(.45)
N	1104	1104	1449	1449	1449	1449	1449
R^2	0.01	0.01	0.01	0.00	0.00	0.01	0.00

Notes: The outcome variables are variety-level (v) log-changes in marginal cost (mc_{vt}) and product appeal (Φ_{vt}); plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue TFP (TFP^R_{ft}), skill (s_{ft}, ratio of non-production to production labor), value of capital (K_{ft}), and material prices (p_{Mft}) from year 2000 to 2009. Regressors include plant-level log-changes in access to higher-income markets where firm outputs are not produced (AHIM^N_{ft}, equation (32)) and the remaining variation in changes in access to higher-income markets (AHIM^{E-N}_{ft}, equation (33)), each interacted with a dummy variable Large_d that equals one if $d\ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

barriers across Indian state borders. As a result, changes in access to higher-income markets outside state borders may have less influence on local firms. To do this, I first identify for each district $d \in D$, the set of districts D^S_d that are in the same state. Then, I measure each district's access to higher-income markets by leaving out districts in other states:

$$\text{AHIM}^S_{dt} \equiv \frac{\sum_{o\in D^S_d} \tau^{1-\sigma}_{dot} E^\sigma_{Ho,2001} N_{Ho,2001}}{\sum_{o\in D^S_d} \tau^{1-\sigma}_{dot} E^\sigma_{Lo,2001} N_{Lo,2001}}. \tag{34}$$

Denote the remaining variation in the original changes in access to higher-income markets after subtracting out the changes in access to higher-income markets in the same state as

$$d\ln \text{AHIM}^{E-S}_{dt} \equiv d\ln \text{AHIM}^E_{dt} - d\ln \text{AHIM}^S_{dt}. \tag{35}$$

I then replace the original measure of changes in access to higher-income markets with these two measures to assess where the effects are coming from. Table 8 reports the estimates.

Table 8: Effects of Access to In-state and Out-of-state Higher-Income Markets

	Log-change (2000-09):						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	mc_v	Φ_v	R/L_f	TFP_f^R	s_f	K_f	p_{Mf}
$d\ln \text{AHIM}_{d,09-00}^{S}$	2.94***	2.41***					
	(0.64)	(0.65)					
$d\ln \text{AHIM}_{d,09-00}^{E-S}$	2.35**	1.69					
	(1.02)	(1.12)					
$\ln \text{AHIM}_{d,2000}^{E}$	0.11*	0.10	-0.01	0.02	-0.01	-0.02	0.04*
	(0.06)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Large_d=0 × $d\ln \text{AHIM}_{d,09-00}^{S}$			0.30	0.05	-0.39	0.01	0.21
			(0.53)	(0.36)	(0.58)	(0.62)	(0.45)
Large_d=1 × $d\ln \text{AHIM}_{d,09-00}^{S}$			9.08***	2.73**	7.55***	7.78**	2.20**
			(0.80)	(1.37)	(2.16)	(3.66)	(1.00)
Large_d=0 × $d\ln \text{AHIM}_{d,09-00}^{E-S}$			-0.36	-0.42	-0.06	0.04	-0.33
			(0.57)	(0.36)	(0.51)	(0.65)	(0.38)
Large_d=1 × $d\ln \text{AHIM}_{d,09-00}^{E-S}$			8.65***	2.16	7.91***	7.68*	1.56
			(0.70)	(1.43)	(2.25)	(3.90)	(1.05)
Large_d			-1.91***	-0.53*	-1.53***	-1.73**	-0.37*
			(0.16)	(0.27)	(0.42)	(0.72)	(0.20)
$\delta_{1,S}^{\text{large}} - \delta_{1,E-S}^{\text{large}}$	.59	.72	.43**	.57***	-.36	.1	.64***
	(.63)	(.69)	(.17)	(.17)	(.37)	(.4)	(.13)
N	1104	1104	1452	1452	1452	1452	1452
R^2	0.02	0.01	0.01	0.01	0.00	0.01	0.01

Notes: The outcome variables are variety-level (v) log-changes in marginal cost (mc_{vt}) and product appeal (Φ_{vt}); plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue TFP (TFP_{ft}^R), skill (s_{ft}, ratio of non-production to production labor), capital (K_{ft}), and material prices (p_{Mft}) from year 2000 to 2009. Regressors include district-level (d) log-changes in access to higher-income markets within the same state (AHIM_{dt}^S, equation (34)) and the remaining variation in changes in access to higher-income markets (AHIM_{dt}^{E-S}, equation (35)), each interacted with a dummy variable Large_d that equals one if $d\ln \text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In general, the estimates on $d\ln \text{AHIM}_{dt}^S$ are somewhat larger than those on $d\ln \text{AHIM}_{dt}^{E-S}$ in accordance with the sizable trade barriers across Indian states. However, it is not the case that changes in access to higher-income markets out of state have no impact. For marginal cost, revenue labor, skill use, and capital, changes in access to higher-income markets out of state show comparable magnitudes of effects that are statistically significant. On the other hand, the point estimates for product appeal, revenue TFP, and material prices are smaller and less precisely estimated. Overall, the results in Table 8 indicate that changes in access to both in-state and out-of-state higher-income markets were important in driving firms' quality upgrading and productivity gains.

6 Conclusion

In this paper, I have presented new evidence that relatively small income differences found within India are sufficient in incentivizing quality upgrading among local firms. I further showed that these quality upgrading lead to productivity gains when increases in access to higher-income markets are large. These results highlight that demand from high-income consumers outside the high-income countries can play an important role in promoting productivity gains for firms in developing countries through quality upgrading.

However, there are some key limitations to this study. One is that it takes a partial equilibrium view and focuses on the responses of individual firms to the changes in access to higher-income markets. It is conceivable that firm decisions to produce higher-quality outputs could have generated additional demand for local skilled workers and have led to changes in the migration flows of skilled workers. Overall, the lack of a general equilibrium analysis makes it difficult to understand the aggregate implications of these results.

Another important finding in this paper is that the model suggests that firms may be facing substantial adjustment costs to reorganize their production into more sophisticated ones. The fact that firms face such costs is hardly surprising, but this paper finds that they were large enough to prevent most Indian firms from experiencing productivity gains despite upgrading quality. Of course, having no direct measure of such adjustment costs, this study has little to say on how these costs may be lowered. Understanding the elements of these adjustment costs appears to be an important area of future research.

References

A. B. Abeberese and M. Chen. Intranational trade costs, product scope and productivity: Evidence from India's Golden Quadrilateral project. *Journal of Development Economics*, 156:102791, May 2022. ISSN 0304-3878. doi: 10.1016/j.jdeveco.2021.102791. URL `https://www.sciencedirect.com/science/article/pii/S0304387821001486`.

D. A. Ackerberg, K. Caves, and G. Frazer. Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451, 2015. ISSN 0012-9682. URL `https://www.jstor.org.ezproxy.cul.columbia.edu/stable/43866416`. Publisher: [Wiley, The Econometric Society].

S. Alder. Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development, Oct. 2016. URL `https://papers.ssrn.com/abstract=2856050`.

T. Allen and C. Arkolakis. Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3):1085–1140, 2014. ISSN 0033-5533. URL `https://www.jstor.org/stable/26372570`. Publisher: Oxford University Press.

T. Allen and C. Arkolakis. The Welfare Effects of Transportation Infrastructure Improvements. *The Review of Economic Studies*, 89(6):2911–2957, Nov. 2022. ISSN 0034-6527. doi: 10.1093/restud/rdac001. URL `https://doi.org/10.1093/restud/rdac001`.

M. Amiti and A. K. Khandelwal. Import Competition and Quality Upgrading. *The Review of Economics and Statistics*, 95(2):476–490, 2013. ISSN 0034-6535. URL `https://www.jstor.org/stable/43554399`. Publisher: The MIT Press.

M. Amiti and J. Konings. Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *The American Economic Review*, 97(5):1611–1638, 2007. ISSN 0002-8282. URL `https://www.jstor.org/stable/30034578`. Publisher: American Economic Association.

J. Asturias, M. García-Santana, and R. Ramos. Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. *Journal of the European Economic Association*, 17(6):1881–1940, Dec. 2019. ISSN 1542-4766. doi: 10.1093/jeea/jvy039. URL https://doi.org/10.1093/jeea/jvy039.

D. Atkin, A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen. Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. *The Quarterly Journal of Economics*, 132(3):1101–1164, 2017a. ISSN 0033-5533. URL https://www.jstor.org/stable/26372700. Publisher: Oxford University Press.

D. Atkin, A. K. Khandelwal, and A. Osman. Exporting and Firm Performance: Evidence from a Randomized Experiment. *The Quarterly Journal of Economics*, 132(2):551–615, 2017b. ISSN 0033-5533. URL https://www.jstor.org/stable/26495145. Publisher: Oxford University Press.

D. Atkin, A. K. Khandelwal, and A. Osman. Measuring Productivity: Lessons from Tailored Surveys and Productivity Benchmarking. *AEA Papers and Proceedings*, 109:444–449, 2019. ISSN 2574-0768. URL https://www.jstor.org/stable/26723987. Publisher: American Economic Association.

K. Baragwanath Vogel, G. H. Hanson, A. Khandelwal, C. Liu, and H. Park. Using Satellite Imagery to Detect the Impacts of New Highways: An Application to India, Jan. 2024. URL https://www.nber.org/papers/w32047.

P. Bastos and J. Silva. The quality of a firm's exports: Where you export to matters. *Journal of International Economics*, 82(2):99–111, Nov. 2010. ISSN 0022-1996. doi: 10.1016/j.jinteco.2010.07.001. URL https://www.sciencedirect.com/science/article/pii/S0022199610000632.

P. Bastos, J. Silva, and E. Verhoogen. Export Destinations and Input Prices. *The American Economic Review*, 108(2):353–392, 2018. ISSN 0002-8282. URL https://www.jstor.org/stable/26527908. Publisher: American Economic Association.

N. Bloom, B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. Does Management Matter? Evidence from India *. *The Quarterly Journal of Economics*, 128(1):1–51, Feb. 2013. ISSN 0033-5533. doi: 10.1093/qje/qjs044. URL https://doi.org/10.1093/qje/qjs044.

I. Brambilla, D. Lederman, and G. Porto. Exports, Export Destinations, and Skills. *The American Economic Review*, 102(7):3406–3438, 2012. ISSN 0002-8282. URL https://www.jstor.org/stable/41724639. Publisher: American Economic Association.

D. Comin, D. Lashkari, and M. Mestieri. Structural Change with Long-Run Income and Price Effects. *Econometrica*, 89(1):311–374, 2021. ISSN 0012-9682. URL https://www-jstor-org.ezproxy.cul.columbia.edu/stable/48628854. Publisher: [Wiley, The Econometric Society].

S. Datta. The impact of improved highways on Indian firms. *Journal of Development Economics*, 99(1):46–57, Sept. 2012. ISSN 0304-3878. doi: 10.1016/j.jdeveco.2011.08.005. URL https://www.sciencedirect.com/science/article/pii/S0304387811000861.

J. De Loecker and F. Warzynski. Markups and Firm-Level Export Status. *The American Economic Review*, 102(6):2437–2471, 2012. ISSN 0002-8282. URL https://www.jstor.org/stable/41724661. Publisher: American Economic Association.

J. De Loecker, P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik. Prices, Markups, and Trade Reform. *Econometrica*, 84(2):445–510, 2016. ISSN 0012-9682. URL https://www.jstor.org/stable/43866442. Publisher: [Wiley, The Econometric Society].

N. de Roux, M. Eslava, S. Franco, and E. Verhoogen. Estimating Production Functions in Differentiated-Product Industries with Quantity Information and External Instruments, Jan. 2021. URL https://www.nber.org/papers/w28323.

J. I. Dingel. The Determinants of Quality Specialization. *The Review of Economic Studies*, 84(4 (301)):1551–1582, 2017. ISSN 0034-6527. URL https://www.jstor.org/stable/26543865. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

D. Donaldson. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *The American Economic Review*, 108(4-5):899–934, 2018. ISSN 0002-8282. URL https://www.jstor.org/stable/26527994. Publisher: American Economic Association.

D. Donaldson and R. Hornbeck. Railroads and American Economic Growth: A “Market Access” Approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016. ISSN 0033-5533. URL https://www.jstor.org/stable/26372653. Publisher: Oxford University Press.

B. Faber. Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System. *The Review of Economic Studies*, 81(3 (288)):1046–1070, 2014. ISSN 0034-6527. URL https://www.jstor.org/stable/43551619. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

L. Foster, J. Haltiwanger, and C. Syverson. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *The American Economic Review*, 98(1):394–425, 2008. ISSN 0002-8282. URL https://www.jstor.org/stable/29729976. Publisher: American Economic Association.

A. Garcia-Marin and N. Voigtländer. Exporting and Plant-Level Efficiency Gains: It’s in the Measure. *Journal of Political Economy*, 127(4):1777–1825, Aug. 2019. ISSN 0022-3808. doi: 10.1086/701607. URL https://www.journals.uchicago.edu/doi/abs/10.1086/701607. Publisher: The University of Chicago Press.

E. Ghani, A. G. Goswami, and W. R. Kerr. Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *The Economic Journal*, 126(591):317–357, 2016. ISSN 0013-0133. URL https://www.jstor.org/stable/24738167. Publisher: [Royal Economic Society, Wiley].

J. C. Hallak. Product quality and the direction of trade. *Journal of International Economics*, 68(1):238–265, Jan. 2006. ISSN 0022-1996. doi: 10.1016/j.jinteco.2005.04.001. URL https://www.sciencedirect.com/science/article/pii/S0022199605000516.

J. C. Hallak and P. K. Schott. Estimating Cross-Country Differences in Product Quality. *The Quarterly Journal of Economics*, 126(1):417–474, 2011. ISSN 0033-5533. URL https://www.jstor.org/stable/23015671. Publisher: Oxford University Press.

C. Hansman, J. Hjort, G. León-Ciliotta, and M. Teachout. Vertical Integration, Supplier Behavior, and Quality Upgrading among Exporters. *Journal of Political Economy*, 128(9):3570–3625, Sept. 2020. ISSN 0022-3808. doi: 10.1086/708818. URL https://www.journals.uchicago.edu/doi/abs/10.1086/708818. Publisher: The University of Chicago Press.

C. J. Hottman, S. J. Redding, and D. E. Weinstein. Quantifying the Sources of Firm Heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364, 2016. ISSN 0033-5533. URL https://www.jstor.org/stable/26372665. Publisher: Oxford University Press.

R. Juhász, S. Sakabe, and D. Weinstein. Codification, Technology Absorption, and the Globalization of the Industrial Revolution, July 2024a. URL https://www.nber.org/papers/w32667.

R. Juhász, M. P. Squicciarini, and N. Voigtländer. Technology Adoption and Productivity Growth: Evidence from Industrialization in France. *Journal of Political Economy*, pages 000–000, Feb. 2024b. ISSN 0022-3808. doi: 10.1086/730205. URL https://www.journals.uchicago.edu/doi/full/10.1086/730205. Publisher: The University of Chicago Press.

A. Khandelwal. The Long and Short (of) Quality Ladders. *The Review of Economic Studies*, 77(4):1450–1476, 2010. ISSN 0034-6527. URL https://www.jstor.org/stable/40836654. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

M. Kugler and E. Verhoogen. Prices, Plant Size, and Product Quality. *The Review of Economic Studies*, 79(1):307–339, 2012. ISSN 0034-6527. URL https://www.jstor.org/stable/41407052. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

J. Levinsohn and A. Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341, 2003. ISSN 0034-6527. URL https://www.jstor.org.ezproxy.cul.columbia.edu/stable/3648636. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

K. Manova and Z. Zhang. Export Prices Across Firms and Destinations. *The Quarterly Journal of Economics*, 127(1):379–436, 2012. ISSN 0033-5533. URL https://www.jstor.org/stable/41337213. Publisher: Oxford University Press.

K. Miyazaki. National Highway-2 Improvement Project, Nov. 2006. URL https://www.jica.go.jp/Resource/english/our_work/evaluation/oda_loan/post/2007/pdf/project27_full.pdf.

NHAI. NHAI Annual Report of 2000-2001. Technical report, National Highways Authority of India, New Delhi, Sept. 2001. URL https://nhai.gov.in/nhai/sites/default/files/2019-03/report_2000_01_English.pdf.

C. Raffler. *QNEAT3 - Qgis Network Analysis Toolbox 3*. May 2018. doi: 10.13140/RG.2.2.13042.02248.

S. J. Redding and E. Rossi-Hansberg. Quantitative Spatial Economics. *Annual Review of Economics*, 9:21–58, 2017. ISSN 1941-1383. URL https://www.jstor.org/stable/26774426. Publisher: Annual Reviews.

S. Subba Rao. Road Map of India, 2011.

C. Syverson. What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365, 2011. ISSN 0022-0515. URL https://www.jstor.org/stable/23071619. Publisher: American Economic Association.

E. Van Leemput. A passage to India: Quantifying internal and external barriers to trade. *Journal of International Economics*, 131:103473, July 2021. ISSN 0022-1996. doi: 10.1016/j.jinteco.2021.103473. URL https://www.sciencedirect.com/science/article/pii/S0022199621000507.

E. Verhoogen. Firm-Level Upgrading in Developing Countries. *Journal of Economic Literature*, 61(4):1410–1464, Dec. 2023. ISSN 0022-0515. doi: 10.1257/jel.20221633. URL https://www.aeaweb.org/articles?id=10.1257/jel.20221633.

E. A. Verhoogen. Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector. *The Quarterly Journal of Economics*, 123(2):489–530, 2008. ISSN 0033-5533. URL https://www.jstor.org/stable/25098908. Publisher: Oxford University Press.

A Data Appendix

A.1 District Concordance

The different sources of data used in this study span across periods of time in which India has introduced changes to its administrative districts. For instance, the Indian Census reports 466 districts in 1991 but 593 districts in 2001. Furthermore, some data sources aggregate a number of districts when reporting. For example, the ASI often aggregates all districts of Delhi when reporting the location of each plant. For these reasons, there is a need to create a district concordance that consistently maps districts across data sources and over time. I begin with the 593 districts in 2001 and introduce the following changes. All nine districts of Delhi are combined. The two districts Mumbai and Suburban Mumbai are combined. The three island districts of Lakshadweep, Andamans, and Nicobars are excluded from the analysis. This reduces the number of districts from 593 to 581.

A.2 District-level Variables

In Table 9, I report the summary statistics on the district-level variables used in this study separately for the estimation and full sample. These include log-changes in access to higher-income markets ($d \ln \text{AHIM}_{d,09-00}$), local wages ($d \ln w_{d,09-00}$), and local price of materials ($d \ln P_{Md,09-00}$). The measurements of AHIM, local wages, and local price of materials are described in Section 3 of the main text. Note that the logged change in local wages has fewer observations compared to other variables, because it requires plant-level wages for its construction and some districts are not covered by the ASI sample. Additional variables include the logged values of access to higher-income markets ($\ln \text{AHIM}_{d,2001}$), total district population ($\ln \text{Population}_{d,2001}$), share of district population with high school degrees (Pop. Share: High School$_{d,2001}$), and share of district population living in urban places (Pop. Share: Urban$_{d,2001}$) in 2001.

Table 9: Summary Statistics on District-level Variables

	Estimation					**Full**				
	Obs	Mean	SD	p10	p90	Obs	Mean	SD	p10	p90
$d \ln \text{AHIM}^E_{d,09-00}$	117	0.06	0.07	-0.03	0.15	570	0.03	0.05	-0.02	0.09
$\ln \text{AHIM}^E_{d,2001}$	117	4.90	0.66	4.15	5.69	570	4.81	0.66	4.03	5.53
$E_{d,2001}$	117	174.33	54.28	113.80	253.34	570	171.72	66.82	105.69	254.96
$E_{Ad,2001}$	117	255.34	68.19	194.58	333.31	570	253.18	66.20	189.03	334.98
$E_{Bd,2001}$	117	109.83	15.42	88.13	129.42	570	109.00	18.21	85.77	131.74
$\ln \text{Population}_{d,2001}$	117	14.57	0.61	13.83	15.25	570	14.06	0.97	12.84	15.05
Pop Shr: High School$_{d,2001}$	117	0.16	0.06	0.09	0.25	570	0.14	0.07	0.07	0.24
Pop Shr: Urban$_{d,2001}$	117	0.28	0.16	0.10	0.53	570	0.23	0.18	0.06	0.46
$d \ln w_{d,09-00}$	83	0.13	0.19	-0.09	0.36	340	0.11	0.27	-0.17	0.40
$d \ln p_{Md,09-00}$	83	-0.28	0.25	-0.55	0.00	337	-0.23	0.36	-0.57	0.18

Notes: "Estimation" refers to the sample of districts that are directly intersecting with an upgraded NHDP road as of March 2009 while "Full" refers to all districts. Reported statistics are the number of observations (Obs), mean, standard deviation (SD), 10th percentile (p10), and 90th percentile (p90). See Section A.2 for descriptions of these variables.

I obtain district-level population counts, broken down by educational level and urban/rural categorization, for 2001 from the Indian Census (Table B-3). This Census table reports each district's urban and rural population counts for six educational levels: (1) illiterate; (2) literate but below matric/secondary; (3) matric/secondary but below graduate; (4) technical diploma or certificate not equal to degree; (5) graduate and above other than technical degree; and (6) technical degree or diploma equal to degree or post-graduate degree. In this paper, educational levels (3) and above are considered to be "high school or higher".

A.3 Variety-level Variables

In Table 10, I report the summary statistics on the output and input variety-level variables used in this study separately for the estimation and full sample. These are log-changes in marginal cost ($d \ln mc_{v,09-00}$), per-unit cost of materials ($d \ln umc_{v,09-00}$), per-unit cost of labor ($d \ln ulc_{v,09-00}$), revenue share ($d \ln \text{RevShr}_{v,09-00}$), output price ($d \ln p_{v,09-00}$), output quantity ($d \ln y_{v,09-00}$), average product appeal ($d \ln \Phi_{v,09-00}$), and input price ($d \ln p_{Mv,09-00}$). The measurement of marginal cost, per-unit cost of materials, per-unit cost of labor, and revenue shares are defined in Section 3 of the main text. Output price is obtained by dividing the total revenue of each variety by its total quantity. Finally, input price is obtained by dividing the value of each input variety by its total quantity. Note that since these variables measure changes from 2000 to 2009, the same variety (i.e., unique plant-product pair) needs to be observed in both periods in order to appear in this sample.

Table 10: Summary Statistics on Variety-level Variables

	Estimation					**Full**				
	Obs	Mean	SD	p10	p90	Obs	Mean	SD	p10	p90
$d \ln mc_{v,09-00}$	1113	0.02	1.66	-0.97	1.42	3393	0.02	1.62	-0.97	1.37
$d \ln umc_{v,09-00}$	1113	0.02	1.66	-0.93	1.40	3393	0.01	1.62	-0.97	1.36
$d \ln ulc_{v,09-00}$	1113	0.04	1.75	-1.22	1.54	3393	0.02	1.71	-1.32	1.59
$d \ln \Phi_{v,09-00}$	1113	-0.00	1.71	-1.14	1.49	3393	0.00	1.69	-1.18	1.45
$d \ln R_{v,09-00}$	1113	0.15	1.57	-1.33	1.76	3393	0.15	1.54	-1.35	1.67
$d \ln p_{v,09-00}$	1113	0.00	1.64	-0.89	1.31	3393	-0.01	1.59	-0.96	1.24
$d \ln y_{v,09-00}$	1113	0.15	2.18	-1.75	2.12	3393	0.16	2.05	-1.67	2.00
$d \ln \text{RevShr}_{v,09-00}$	1113	0.06	1.39	-1.19	1.42	3393	0.08	1.34	-0.97	1.24
	Obs	Mean	SD	p10	p90	Obs	Mean	SD	p10	p90
$d \ln p_{Mv,09-00}$	928	-0.21	1.50	-1.18	0.62	2784	-0.21	1.43	-1.09	0.60

Notes: "Estimation" refers to the sample of observations locating in districts that are directly intersecting with an upgraded NHDP road as of March 2009 while "Full" refers to all observations. Reported statistics are the number of observations, mean, standard deviation, 10th percentile, and 90th percentile.

A.4 Plant-level Variables

In Table 11, I report the summary statistics on the plant-level variables used in this study separately for the estimation and full sample. These are log-changes in total revenue ($d \ln R_{f,09-00}$), skill intensity ($d \ln \text{SI}_{f,09-00}$), skill premium ($d \ln \text{SP}_{f,09-00}$), wage ($d \ln w_{f,09-00}$), value of capital ($d \ln K_{f,09-00}$), revenue TFP ($d \ln \text{TFP}^{R}_{f,09-00}$), value-added TFP ($d \ln \text{TFP}^{\text{VA}}_{f,09-00}$), revenue per labor ($d\ln \mathrm{R/L}_{f,09-00}$), and value-added per labor ($d\ln \mathrm{VA/L}_{f,09-00}$). Skill intensity is measured as the ratio of non-production to production labor while skill premium is measured as the ratio of non-production to production wages. Plant wages are computed by dividing the total expenditure on labor by the total quantity of labor (reported in man-days). The value of capital is measured as the reported book value of fixed assets. Value-added is the difference between total revenue and total expenditure on materials. Sometimes, plants had negative values for value-added, which explains the lower number of observations for variables that take logarithms of value-added in the table.

Table 11: Summary Statistics on Plant-level Variables

	Estimation					**Full**				
	Obs	Mean	SD	p10	p90	Obs	Mean	SD	p10	p90
$d\ln \mathrm{R/L}_{f,09-00}$	1465	0.08	0.80	-0.86	0.99	4291	0.08	0.77	-0.82	0.96
$d\ln \mathrm{VA/L}_{f,09-00}$	1287	0.04	1.12	-1.23	1.30	3783	0.06	1.08	-1.14	1.25
$d\ln \mathrm{TFP}^{R}_{f,09-00}$	1465	0.30	0.61	-0.40	1.05	4291	0.31	0.59	-0.37	0.98
$d\ln \mathrm{TFP}^{\mathrm{VA}}_{f,09-00}$	1287	0.06	1.10	-1.20	1.28	3783	0.08	1.07	-1.13	1.24
$d\ln s_{f,09-00}$	1465	-0.04	0.82	-1.07	1.02	4291	-0.03	0.78	-1.01	0.98
$d\ln w_{f,09-00}$	1465	0.15	0.40	-0.35	0.67	4291	0.11	0.38	-0.35	0.61
$d\ln K_{f,09-00}$	1465	-0.01	1.02	-1.20	1.28	4291	-0.02	1.04	-1.25	1.32
$d\ln R_{f,09-00}$	1465	0.19	0.97	-0.90	1.34	4291	0.15	0.93	-0.90	1.26
$d\ln \mu_{f,09-00}$	1465	0.00	0.15	-0.12	0.12	4291	0.00	0.15	-0.13	0.13

Notes: "Estimation" refers to the sample of observations locating in districts that are directly intersecting with an upgraded NHDP road as of March 2009 while "Full" refers to all observations. Reported statistics are the number of observations, mean, standard deviation, 10th percentile, and 90th percentile. See Section A.4 for a description of the variables.

A.5 Construction of Road Network

In this section, I describe in greater detail how I digitize India's road network in 2001 as shown in Figure 2. As a first step, I obtained a scanned road map of India published by Survey of India, which is India's national mapping agency under the Department of Science and Technology (Subba Rao, 2011). The scanned map plots four types of roads: national highways with number, all weather motorable (differentiated according to importance by two levels of thickness), motorable in fair weather, and where delay may occur. Of these, I manually digitized all national highways and major (thicker of the two levels) all weather motorable roads using QGIS. The digitized map consists of 59,088 km of national highways and 92,718 km of major roads.

Although this map is published in 2011, comparison of the depicted national highways to official records reveals that the map depicts India's road network from around 2001. For example, India's official report on the total length of national highways lists 57,737 km as of March 31, 2001 (NHAI, 2001). This number rises to 70,934 km by 2011. Therefore, the total length of the digitized national highways comes much closer to the official numbers in 2001 than in 2011. I was able to obtain two lists of national highways from the National Highways Authority of India dated November 30 of 2001 and March 6 of 2007.[15] By comparing the

[15]The official website of NHAI underwent major changes and older records no longer appear

numbered national highways from the scanned map with these lists further confirm that the map depicts the road network from around 2001 since it is missing many of the national highways that appear in the list from 2007. I make use of the national highway list from 2001 to determine which of the digitized national highways existed in 2001. This process results in 54,934 km of digitized national highways in 2001, which account for 95.1% of 57,737 km of national highways from the official records for 2001. Figure 2 plots these digitized roads for 2001.

Next, I obtained detailed status updates on the NHDP road upgrades from NHAI's website dated May 31, 2017.[16] These status updates list the name, length, associated national highway number, and date of start/completion for each stretch (203 in Phase I and 166 in Phase II). These information allow me to identify which stretches of the NHDP roads were completed by each year, which are plotted in Figure 3. To match each of these stretches with the digitized roads from earlier, I make use of the information on each stretch's description, its length, and associated national highway number. Only the roads pertaining to the Port Connectivity projects are omitted from this process.

B Theoretical Appendix

B.1 Derivation of Demand

This sections describes an environment that gives rise to the demand function in equation (1). Consider a consumer of type i in location d, who obtains utility over the consumption of varieties of a differentiated good. Her utility U_{id} takes a non-homothetic CES form studied by Comin et al. (2021) and is implicitly defined by the constraint

$$\sum_v \left(\frac{c_{vid}}{g_v(U_{id})} \right)^{\frac{\sigma-1}{\sigma}} = 1, \tag{36}$$

where $\sigma > 1$ is the constant elasticity of substitution between varieties, c_{vid} is the quantity of consumption of variety v, and $g_v(\cdot)$ is a positive valued, continuously differentiable, and monotonically increasing function for each variety. I choose the following functional form:

$$g_v(U) = U^{1/q_v}, \tag{37}$$

where q_v is the quality of variety v.

The budget constraint is $E_{id} = \sum_v p_{vd} c_{vid}$, where E_{id} is her expenditure and p_{vd} is the price of variety v that consumers face in location d. The demand for variety v is given by

$$c_{vid} = p_{vd}^{-\sigma} U_{id}^{(1-\sigma)/q_v} E_{id}^{\sigma}, \tag{38}$$

where the level of utility U_{id} is implicitly defined as a function of expenditure, prices, and qualities that satisfy

$$E_{id} = \left[\sum_v p_{vd}^{1-\sigma} U_{id}^{(1-\sigma)/q_v} \right]^{\frac{1}{1-\sigma}}. \tag{39}$$

there. I made use of Wayback Machine to find these lists from archived web pages. See https://web.archive.org/web/20020614225057/http://www.nhai.org:80/statewise2.htm for example. I was not able to obtain a list of national highways from 2009 and resort to the list from 2007 instead.

[16] Again, I rely on Wayback Machine to retrieve these lists from the archived web pages of NHAI's old website.

Suppose in equilibrium, the utility of each type of consumer is equalized across space through a free mobility condition so that $U_{id} = U_i$ for all $d \in D$, which then delivers the demand in equation (1).

C Empirical Appendix

C.1 Discussion on Least-Cost Networks

Following Faber (2014), a popular identification strategy among studies of infrastructure projects relies on hypothetical least-cost paths that connect the nodes of the actual network. To implement this strategy here, I would first calculate a hypothetical least-cost network that connects the four targeted districts of Delhi, Mumbai, Chennai, and Kolkata, which has recently been done by Baragwanath Vogel et al. (2024). Then, hypothetical changes in access to higher-income markets would be calculated based on travel times calculated on this least-cost network, which would then be used to instrument for the observed changes in access to higher-income markets. The exclusion restrict in this case will require that the hypothetical changes in access to higher-income markets from the least-cost network be uncorrelated with any omitted variables that can affect the outcome variables of interest. As with the NHDP network, another way of looking at is to ask whether locations distant from the least-cost network are valid controls for districts close to the network. There are reasons to believe that this is unlikely to be the case with India's NHDP.

Unlike the Chinese highway project studied by Faber (2014), least-cost paths are much more constrained in its shape because it has to connect the four targeted districts of Delhi, Mumbai, Chennai, and Kolkata. In comparison, China's highway system had the goal of connecting all provincial capitals and cities with an urban population above 50,000 (Faber, 2014). Given the locations of these four cities, it is difficult for least-cost paths to deviate much from the actual network. The least-cost paths constructed by Baragwanath Vogel et al. (2024) (Figure B.1) indeed show much overlap with the actual network. For example, their network with six and nine nodal cities show only small deviations that are unlikely to affect travel times between districts. Their network with four nodal cities show larger deviations, especially on Delhi-Mumbai and Mumbai-Chennai routes. In fact, this network appear quite similar to the straight-line network used by Ghani et al. (2016), a hypothetical network composed of five straight lines that connect the four targeted districts of Delhi, Mumbai, Chennai, and Kolkata (two lines are used to connect Chennai and Kolkata so that the network does not cross over the ocean). This straight-line network is relatively easy to construct, so I can check if proximity to this network correlates with observable district characteristics.

In Table 12, I report the results from regressing a number of district characteristics (logged per-capita expenditure, logged population, logged share of population with high school degree, and logged share of population in urban places) in 2001 on the minimum distance to the NHDP and the straight-line network. The regression results show with statistical significance that districts close to the NHDP are larger, more educated, and more urbanized on average. In comparison, districts close to the straight-line network are poorer, larger, and less educated on average. It is unsurprising that districts close to the NHDP network are richer given the historical importance of the four cities, which would have naturally led to economic development along the paths connecting these locations. Although the signs are switched in some cases, distance to the straight-line network show significant correlation with observable district characteristics, casting doubt on their "exogeneity" as instrumental

Table 12: Regression of District Characteristics on Minimum Distance to the NHDP and Straight-line Networks

	NHDP				Straight-line Network			
	(1) Exp	(2) Pop	(3) HS	(4) Urb	(5) Exp	(6) Pop	(7) HS	(8) Urb
Distance to NHDP	0.02	-0.64***	-0.08***	-0.13***				
	(0.02)	(0.05)	(0.02)	(0.04)				
Distance to SL					0.05***	-0.26***	0.02**	-0.02
					(0.01)	(0.02)	(0.01)	(0.02)
R^2	0.00	0.28	0.02	0.02	0.07	0.28	0.01	0.00
N	570	570	570	563	570	570	570	563

Notes: The dependent variables are districts' logged per-capita expenditure (Exp), population (Pop), share of population with high school degree (HS), and share of population in urban places (Urb) in 2001. All distances are measured in 100 km. Heteroskedasticity robust standard errors in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

variables. Since Baragwanath Vogel et al. (2024) work with much finer geographic units (markets defined as sets of contiguous pixels that contain built-up economic activity), these small deviations may suffice for them. However, as I look at travel times between districts, these deviations are likely too small to avoid being correlated with district characteristics. Indeed, the correlation between the least-cost paths and local economic characteristics was a concern in Faber (2014) as well. He addressed it accordingly through the use of various controls. This concern seems to be particularly strong in the context of India's NHDP, because the stated goal was to connect the four historically important cities of India. As their historical importance persist to the present day, it is difficult for least-cost networks connecting these districts to be "exogenous".

C.2 Alternative Measures of Marginal Cost and Product Appeal

In this section, I test the robustness of the estimated effects on variety-level changes in marginal cost and product appeal reported in Table 2 with respect to alternative measures of these variables.

First, I come up with an alternative measure of marginal cost following the approach of Garcia-Marin and Voigtländer (2019). A by-product of estimating the production function in equation (21) is a measure of markups. De Loecker and Warzynski (2012) shows that for a flexibly variable input, which I assume that material inputs are, the first-order condition of the cost-minimization problem implies that markups should satisfy:

$$\mu_{vt} \equiv \frac{p_{vt}}{\tilde{mc}_{vt}} = \alpha_M \left(\frac{p_{Mv} M_{vt}}{p_{vt} y_{vt}} \right)^{-1}, \tag{40}$$

where $\tilde{mc}_{vt}$ is an alternative measure of marginal cost. Given an estimate of the output elasticity on materials and data on the cost share of materials, I can measure the markup in equation (40). Then, the marginal cost can be obtained as the ratio of output price and markup, $\tilde{mc}_{vt} = p_{vt}/\mu_{vt}$.

The first three columns in Table 13 report the estimated effects on variety-level changes in unit price, markups, and the alternative measure of marginal cost, respectively. Output

Table 13: Effects of Access to Higher-Income Markets on Alternative Measures of Marginal Cost and Product Appeal

	Log-change (2000-09):						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	p_v	μ_v	$\tilde{mc}_v$	$\Phi_v^{\sigma=3}$	$\Phi_v^{\sigma=4}$	$\Phi_v^{\sigma=6}$	$\Phi_v^{\sigma=7}$
$d \ln \text{AHIM}^E_{d,09-00}$	2.47***	-0.15**	2.62***	2.02**	2.17***	2.29***	2.32***
	(0.71)	(0.06)	(0.71)	(0.84)	(0.78)	(0.74)	(0.73)
$\ln \text{AHIM}^E_{d,2000}$	0.09	-0.00	0.10	0.10	0.10	0.10	0.10
	(0.07)	(0.01)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
N	1104	1104	1104	1104	1104	1104	1104
R^2	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Notes: The variables in the table include variety-level (v) log-changes in output price (p_{vt}), markup (μ_{vt}), marginal cost ($\tilde{mc}_{vt} = p_{vt}/\mu_{vt}$), and product appeal ($\Phi_{vt}^{\sigma=x}$) for $x \in \{3, 4, 6, 7\}$. District-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}) are also included. Refer to Section 3 for more details on these variables. Estimates of the constant are omitted. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

prices increased as one would expect from the effects on marginal cost identified in Table 2. Markups are decreasing with increased access to higher-income markets, which likely reflects the increased competition that local firms face from the bilateral reduction in trade costs. However, the magnitude is much smaller in comparison to the estimated effect on output prices. As a result, the third column identifies an effect on the alternative measure of marginal cost that is similar in magnitude to the estimated effect on output prices as well as that on marginal costs from Table 2.

Next, for product appeal, I consider alternative values of the elasticity of substitution σ. In the baseline specification, the assumed value was 5. Here, I consider

$$d \ln \Phi_{vt}^{\sigma=x} = \left(\frac{1}{\sigma-1}\right) d \ln R_{vt} + d \ln p_{vt} - \left(\frac{1}{\sigma-1}\right) d \ln \text{MA}^E_{ot} \tag{41}$$

for values $x \in \{3, 4, 6, 7\}$. The results are shown in the last four columns of Table 13. Although the estimated effects shrink with lower values of the elasticity of substitution, they are not very sensitive and retain their statistical significance for all tested values.

C.3 Robustness Test: State Fixed Effects

Table 14 re-estimated the main results in Section 4 after including state fixed effects. Note that the state fixed effects here absorb state-level trends in the outcome variables. The estimated effects on marginal cost, product appeal, revenue per labor, revenue TFP, and skill use are slightly larger than the baseline estimates but remain qualitatively similar. In comparison, the estimate on value of capital in the sixth column is much smaller in magnitude and is no longer statistically significant while the estimate on material prices is also less precise. Despite some differences, the conclusions on quality upgrading and productivity gains in the main analysis of the paper are robust to the inclusion of state fixed effects.

Table 14: Effects of Access to High-Income Markets with State Fixed Effects

	Log-change (2000-09):						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	mc_v	Φ_v	R/L_f	TFP^R_f	s_f	K_f	p_{Mf}
$d\ln\text{AHIM}^E_{d,09-00}$	4.37***	3.64**					
	(1.45)	(1.44)					
$\ln\text{AHIM}^E_{d,2000}$	0.19*	0.19	-0.04	0.01	-0.04	-0.03	0.02
	(0.11)	(0.12)	(0.05)	(0.03)	(0.04)	(0.06)	(0.03)
$\text{Large}_d{=}0 \times d\ln\text{AHIM}^E_{d,09-00}$			-0.99	-0.49	-0.95	-0.33	0.21
			(0.61)	(0.35)	(0.61)	(0.75)	(0.43)
$\text{Large}_d{=}1 \times d\ln\text{AHIM}^E_{d,09-00}$			8.76***	3.45**	7.29***	3.33	2.19*
			(1.34)	(1.35)	(2.73)	(3.04)	(1.26)
Large_d			-2.00***	-0.78**	-1.55**	-0.76	-0.40
			(0.36)	(0.33)	(0.64)	(0.78)	(0.30)
$\delta_1^{\text{large}} - \delta_1^{\text{small}}$			9.75***	3.94***	8.24***	3.66	1.98
			(1.64)	(1.44)	(2.89)	(3.28)	(1.36)
N	1103	1103	1450	1450	1450	1450	1450
R^2	0.03	0.03	0.04	0.03	0.01	0.01	0.02

Notes: All columns are estimated with state fixed effects. The outcome variables are variety-level (v) log-changes in marginal cost (mc_{vt}) and product appeal (Φ_{vt}); plant-level (f) log-changes in revenue per labor (R/L_{ft}), revenue TFP (TFP^R_{ft}), skill (s_{ft}, ratio of non-production to production labor), value of capital (K_{ft}), and material prices (p_{Mft}) from year 2000 to 2009. Other variables include district-level (d) log-changes in access to higher-income markets (AHIM^E_{dt}) and a dummy variable (Large_d) that equals one if $d\ln\text{AHIM}_{d,09-00} \geq 0.18$ and zero otherwise. Refer to Section 3 for more details on these variables. Standard errors clustered by district in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.4 Cross-sectional Patterns

In this section, I establish a number of cross-sectional patterns relating to product quality documented in the existing literature using the data used in this study. Being able to replicate previously documented patterns is critical in assuring that the same economic forces present in the data used in the past are also present in India. Without this evidence, there will be questions over the external validity of my findings.

I confirm the cross-sectional relationship between district income and the following four variables: output price, skill use, plant wages, and material input prices. Specifically, I regress the logarithm of each of these four variables on the logarithm of district per-capita expenditure from 2001 with product (5-digit ASICC) fixed effects. Table 15 reports the results, which confirm the patterns found by previous researchers in their respective data. The results in the table indicate that higher-income locations produce more expensive varieties, use more skilled workers and pay higher wages, and use more expensive material inputs. Together, these results suggest that the same economic forces present in the data used by previous researchers are very much present in the Indian data.

Table 15: Cross-sectional Patterns between Access to Higher-Income Markets and Output Prices, Skill, Wages, and Material Prices

Log value in 2000:	(1) p_v	(2) s_f	(3) w_f	(4) p_{Mf}
$\ln$ Per-Capita Expenditure$_{d,2001}$	0.094**	0.140***	0.128***	0.052**
	(0.037)	(0.034)	(0.039)	(0.021)
N	26458	26458	26458	26310
R^2	0.89	0.49	0.63	0.40

Notes: All columns are estimated with product fixed effects. The outcome variables are logged output price, plant skill (ratio of non-production to production labor), plant wages, and material price. Standard errors clustered by district in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.