Infrastructure Investment and Labor Monopsony Power[*]

Wyatt J. Brooks [†1], Joseph P. Kaboski [‡2], Illenin O. Kondo [§3], Yao Amber
Li [¶4], and Wei Qian [‖5]

[1]*Arizona State University*
[2]*University of Notre Dame, CEPR, and NBER*
[3]*Federal Reserve Bank of Minneapolis*
[4]*Hong Kong University of Science and Technology*
[5]*Shanghai University of Finance and Economics*

June 16, 2021

Infrastructure Investment and Labor Monopsony Power

Abstract

In this paper we study whether or not transportation infrastructure disrupts local monopsony power in labor markets using an expansion of the national highway system in India. Using panel data on manufacturing firms, we find that monopsony power in labor markets is reduced among firms near newly constructed highways relative to firms that remain far from highways. We estimate that the highways reduce labor markdowns significantly. We use changes in the composition of inputs to identify these effects separately from the reduction of output markups that occurs simultaneously. The impacts of highway construction are therefore pro-competitive in both output and input markets, and act to increase the share of income that labor receives by 1.8–2.3 percentage points.

1 Introduction

Investments in transportation infrastructure are key to promoting economic growth and spreading growth geographically by lowering trade costs (e.g., Donaldson, 2018), inducing firm growth (e.g., Lu, 2020), increasing competition (e.g., Asturias et al., 2019), and allowing workers to commute (e.g., Asher and Novosad, 2020). Yet the effect that transportation infrastructure has on wages is potentially ambiguous as connecting markets causes many things to change simultaneously: workers may work in other markets (raising local wages), firms may attract workers from new markets where wages are lower (lowering local wages), and local production may increase or decrease based on changes in output competition across markets.

In this paper, we study an aspect of how wages are affected by transportation infrastructure that has received less attention: to what extent does transportation infrastructure affect the ability of firms to exercise monopsony power in local labor markets? In principle, easier labor commutes may increase the mobility of a labor force that is otherwise captive to the local labor market. This is especially important in rural areas that are otherwise unconnected by roads. We address this question in the context of the India's Golden Quadrilateral (GQ) expressway expansion initiative. The Golden Quadrilateral initiative is of interest for several reasons. First, it is one of the largest highway expansions in the world. Second, in developing countries monopsony power may be particularly strong. India has traditionally been known for having spatially segmented markets where the potential for monopsony power is high (see, for example, Brooks et al., 2021; Binswanger and Mark R. Rosenzweig, 1984; Braverman and Stiglitz, 1982, 1986). Finally, the GQ was built quite rapidly, expanding from only 5 to 95 percent complete between 2001 and 2006.

We focus on labor "markdowns" and their impact on labor's share of aggregate income. A markdown is the ratio of the value of the marginal product of labor to the wage that is above and beyond what is explained by a markup in the output market. As an example, suppose a firm uses oil and labor as inputs. We observe that the value of the marginal product of oil exceeds its input price by 10%, while the value of the marginal product of labor exceeds the wage by 32%. By assuming that the market for materials (oil, in this example) is competitive, we can apply the methods of Brooks et al. (2021) to disentangle the firm's markup on its output from the markdown on labor. In this example, since oil is competitive, the 10% gap between the value of the marginal product of oil and its price is equal to the markup on its output. Then the markdown on its labor is simply 20% (i.e., 1.32/1.1-1). Applying these methods to detailed geographic data, we find substantial pre-existing labor markdowns in the data. We then show that for firms near the newly constructed expressway, average

markdowns are reduced substantially.

Figure 1 illustrates the pattern for three different measures of the markdown. It shows the time path of labor markdowns by tercile of how close a firm's district is to the GQ, which is all but completed in 2006. The paths of the three terciles show roughly identical increases until 2006, at which point the markdowns in the most remote locations continue to rise, while the two terciles closest to the expressway flatten off.

Figure 1: Impact of Highway on Markdowns by Proximity Tercile
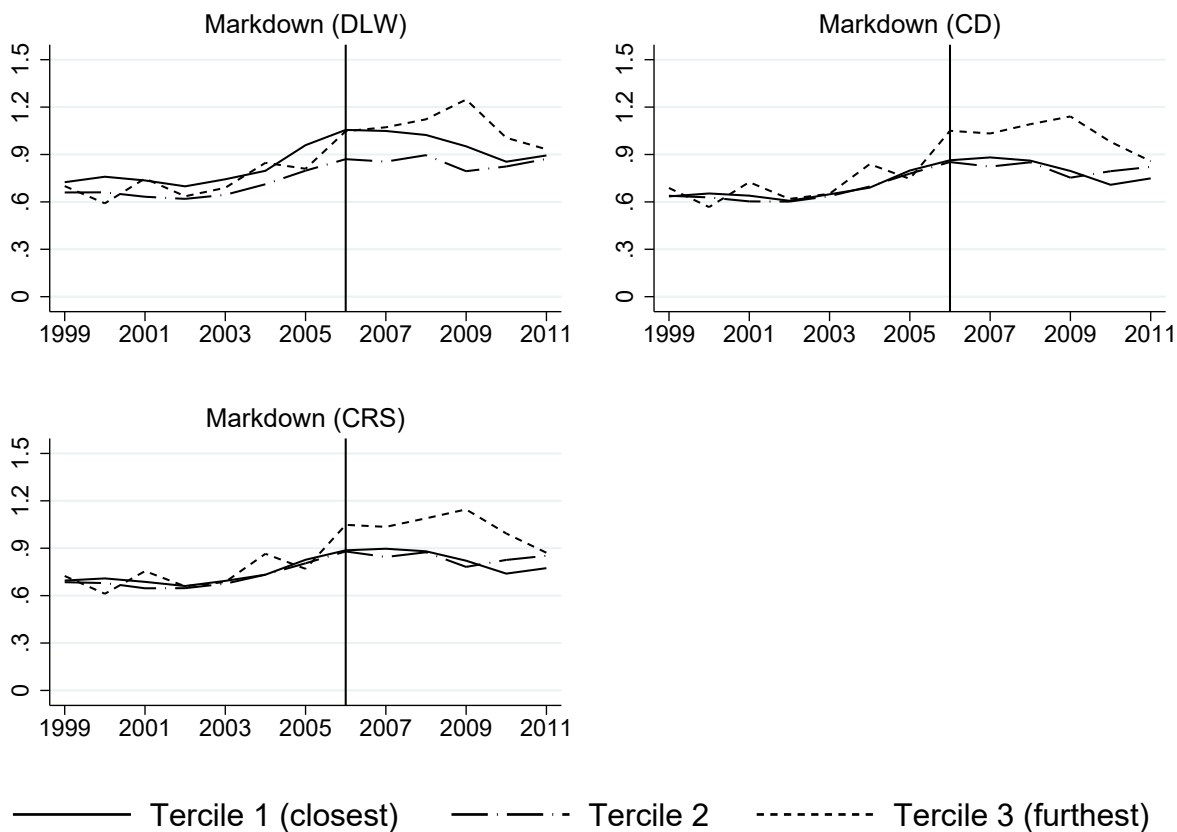
Note: This graph shows the time path of labor markdowns by tercile of how close a firm's district is to the Golden Quadrilateral. The reported markdown is the weighted average of firms' markdowns, where the weight is the product of the firms' average labor compensation and the survey-provided sampling weight. DLW, CD and CRS are the three methods we use to measure markdowns, and their construction is explained in Section 3.3. The vertical line indicates the year 2006 when GQ is mostly completed.

We delve into the causes of the lower markdowns for connected districts. First, we show that even though the impact of the GQ connection on markups is perhaps negative, it cannot account for the lower markdown. Instead we measure relative increases in labor compensation and labor's share among connected firms. While we cannot isolate the precise way

2

in which these lower markdowns are manifested, we rule out several possible explanations. In particular, we do not find strong evidence that the highway-induced reduction in labor markdowns is correlated with the firm labor share. Neither do we find evidence of increased labor supply elasticity or changes in skill premia in integrated labor markets.

Nevertheless, the markdown patterns we uncover are economically significant. Markdowns induce an aggregate labor's share that is between 3 and 7 percentage points lower than it would be in the absence of markdowns. The introduction of the GQ reduces markdowns and leads to an increase in the aggregate labor's share of aggregate output by about 1.8 to 2.3 percentage points.

The rest of this paper is organized as follows. We review the relevant literature in the remaining introduction. Section 2 reviews a general model of firm with monopsony power, and uses it to derive our formula for markdowns. In Section 3, we give background on the Golden Quadrilateral initiative, our data, and our practical measurement of markups and markdowns. Section 4 then presents the empirical analysis, and Section 5 concludes.

1.1   Related Literature

This paper contributes to multiple ongoing areas of research. There is a growing literature on labor monopsony, especially in the United States (Card et al., 2018; Gouin-Bonenfant, 2018; Lamadon et al., 2019; Berger et al., 2018; Hershbein et al., 2020). The first three examine the sharing of rents in labor markets with search and matching frictions. Berger et al. (2018) study a similar labor market using mergers for identification. Hershbein et al. (2020) study classical monopoly power, documenting time series patterns in markdowns in the U.S., including a sharp increase after 2000. Most directly, we borrow our measure for markdowns on labor from Brooks et al. (2021), which also focuses on labor markdown in India. This paper is novel in looking at the impact of large-scale infrastructure investment on firms' monopsony power, however.

Another series of papers have analyzed the impacts of infrastructure investments on firms and labor markets. Most closely related is Asturias et al. (2019), which uses a trade model to quantify the impact of increased competition in the product market from the expansion of the Golden Quadrilateral highway system in India. We utilize their data and complement their findings by empirically assessing the impact of the GQ on labor market competition. Although they find impacts on markups, the techniques developed here are robust to variation in product market markups, and those markups do not affect our estimation of input markdowns. Other work has studied the impact of roads on labor migration. For example, Asher and Novosad (2020) show that the primary impact of a large national

rural road construction program connecting villages in India is an occupational move away from agriculture to wage income, and suggest that this is driven by labor opportunities from outside the villages. Similarly, Brooks and Donovan (2020) show how bridge infrastructure investment in Nicaragua substantially changed labor patterns by allowing rural workers to access new labor markets. We complement this literature by showing how changes in labor market opportunities and competition for workers impacts firms' ability to markdown wages monopsonistically.[1]

Several other papers have looked at the impact of the National Trunk Highway System, another major highway expansion but in China. Using a before-after approach and region level data, Faber (2014) finds that it lowered the industrial output and growth of newly connected areas. In contrast, using more continuous variation and firm-level data, Lu (2020) finds that the same project promotes growth of firms in newly connected areas using variations in the timing of highway segment construction. Alder and Kondo (2020) show how the highway planning in China was driven by political economy considerations, and they solve for the optimal highway system devoid of such consideration. An advantage of our India setting is that we do not suffer from the challenge of disentangling multiple reforms—including relaxation of the Hukou system that allows for more migration and unilateral trade liberalization–which could have a first-order impact on the labor markets and confound interpretation.

2   Model

Following Brooks et al. (2021), this section derives our measure of the markdown and links it to labor's share of income.

2.1   Monopsonistic Firm's Problem

We consider the problem of a firm who has market power in both the product market, the local labor market, and potentially other input markets. We index the firm and its output by $n = 1...N$, its industry by $i = 1...I$, and its location by $k = 1...K$. Firms use two inputs: labor and materials. Importantly, the firm is a price-taker in one input market. In our application, materials will be the input for which the firm is a price-taker.[2]

Letting x_{nki}^M denote the quantities of materials and x_{nki}^L denote the labor employed by

---

[1]See also Alder (2016) for a study of the aggregate and regional effects of the GQ and counterfactual network designs using a Ricardian spatial trade model.

[2]Brooks et al. (2021) develop this model with an arbitrary number of inputs. Our results extend to this case so long as the firm is a price-taker in at least one of them.

firm n located in k and operating in industry i, we express the industry-specific production function quite generally as:

$$y_{nki} = F_i(x_{nki}^L, x_{nki}^M; Z_{nki}) \tag{1}$$

where Z_{nki} is a set of firm-level characteristics, including productivity but also any other potential firm or location-specific factors that might affect the level or shape of technology.

For labor input, the firm faces an inverse supply function that depends on the aggregate labor supply X_{ki}^L in location k and industry i:

$$w_{ki}^L = G_i^L\left(X_{ki}^L\right), \tag{2}$$

where the aggregate quantity equals, by market clearing, the total input demanded across all N_{ki} firms in the industry and location:

$$X_{ki}^L = \sum_{n=1}^{N_{ki}} x_{nki}^L. \tag{3}$$

Again, the supply of materials is perfectly elastic at a given factor price, w_{ki}^M. The most natural interpretation of this assumption is that materials are traded both within country and internationally so that local firms have no market power to exert over their prices.[3]

Likewise, the firm faces an inverse demand for its output, given the output of all other goods, which we denote $\{y_{jki}\}_{j\neq i}$:

$$p_{nki} = H_i\left(y_{nki}; \{y_{jki}\}_{j\neq i}\right) \tag{4}$$

Here we assume that firms face a single demand function. This would be the case if they sell to a single national market and if there are no transportation costs across locations.[4]

The firm's profit maximization problem is therefore:

$$\max_{\{y_{nki}, x_{nki}^M, x_{nki}^L\}} p_{nki}y_{nki} - w_{ki}^M x_{nki}^M - w_{ki}^L x_{nki}^L \tag{5}$$

subject to:

$$y_{nki} = F_i(x_{nki}^L, x_{nki}^M; Z_{nki})$$

---

[3]We believe this assumption is reasonable among the Indian manufacturing firms being studied in this paper, but we admit it may not be appropriate in every context. See Morlacco (2019) for a discussion of market power among large French importers.

[4]Alternatively, the model could be extended to include a competitive wholesale sector in each location that buys the local output from the firms, then sells it to many locations based on transportation costs and variation in local demand. The firms internalize that their choice of output affects wholesaler demand so that there are still markups for their output. This case is not presented here for ease of exposition.

$$p_{nki} = H_i\left(y_{nki}; \{y_{jki}\}_{j \neq i}\right)$$

$$w_{ki}^L = G_i^L\left(X_{ki}^L\right).$$

The fact that p_{nki} and w_{ki}^L are both functions in the constraints emphasizes that firms internalize their effect on both output prices and input prices. In particular, by producing more output, they reduce the price of their own output, and by choosing to use more labor, firms internalize the effect of higher wages.

Using λ_{nki} as the Lagrange multiplier on the production function, the firm's first-order conditions are:

$$p_{nki} + \frac{\partial p_{nki}}{\partial y_{nki}} y_{nki} = \lambda_{nki} \tag{6}$$

$$w_{ki}^L + \frac{\partial w_{ki}^L}{\partial x_{nki}^L} x_{nki}^L = \lambda_{nki} \frac{\partial F_i}{\partial x_{nki}^L} \tag{7}$$

$$w_{ki}^M = \lambda_{nki} \frac{\partial F_i}{\partial x_{nki}^M} \tag{8}$$

Notice that equations (6), (7) and (8) can be rewritten, respectively, as:

$$\frac{\lambda_{nki}}{p_{nki}} = 1 + \frac{\partial \log(p_{nki})}{\partial \log(y_{nki})} \tag{9}$$

$$\lambda_{nki} \frac{y_{nki} \frac{\partial \log(F_i)}{\partial \log(x_{nki}^L)}}{w_{ki}^L x_{nki}^L} = 1 + \frac{\partial \log(w_{ki}^L)}{\partial \log(x_{nki}^L)} \tag{10}$$

$$\lambda_{nki} \frac{y_{nki} \frac{\partial \log(F_i)}{\partial \log(x_{nki}^M)}}{w_{ki}^M x_{nki}^M} = 1 \tag{11}$$

2.2 Markups

We define a markup as the ratio of output price to marginal cost. A common measure for markups from de Loecker and Warzynski (2012) is the ratio of the the elasticity of an input to its cost share. However, that ratio which we define as μ_{nki}^L for labor input is not, in general, equal to the markup:

$$\mu_{nki}^L \equiv \frac{\frac{\partial \log(F_i)}{\partial \log(x_{nki}^L)}}{\frac{w_{ki}^L x_{nki}^L}{p_{nki} y_{nki}}}. \tag{12}$$

In the absence of monopsony power, using any input implies the same measured markup. However, this is no longer true with monopsony power. If the firm is not price-taking in labor, for example, μ_{nki}^L could exceed one for two reasons: market power in the output market, or monopsonistic market power in the labor market. Here our assumption that the material

input M is perfectly elastically supplied is helpful, since it provides a way of measuring markups in output prices without being confounded by the presence of monopsonistic market power on other inputs.

Manipulating the equations above, we can solve for the markup μ_{nki}^M making use of the price-taking input as follows:

$$\frac{\frac{w_{ki}^M x_{nki}^M}{p_{nki} y_{nki}}}{\frac{\partial \log(F_i)}{\partial \log(x_{nki}^M)}} = \frac{1}{\mu_{nki}^M} = 1 + \frac{\partial \log(p_{nki})}{\partial \log(y_{nki})}. \tag{13}$$

2.3 Markdowns

Taking the ratio of equations (7) and (8), we can then isolate monopsony power in the market for labor relative to the markup:

$$\forall m, \frac{\mu_{nki}^L}{\mu_{nki}^M} = 1 + \frac{\partial \log(w_{ki}^L)}{\partial \log(x_{nki}^L)} \tag{14}$$

In other words, the left-hand side is a properly normalized measure of the exercise of classical monopsonistic market power, i.e., what the literature refers to as a "markdown". This gives us a clear way of measuring the markdown. At times, we refer to the markdown as the "markdown ratio" to emphasize that the markdown is calculated as the ratio of two markup measures.

Substituting in the definition of the aggregate supply $X_{ki}^L = \sum_{n=1}^{N_{ki}} x_{nki}^L$, and manipulating shows how this equals the ratio of the share of firm $s_{nki}^L = \frac{w_{ki}^L x_{nki}^L}{\sum_l w_{ki}^L x_{lki}^L}$ to the elasticity of labor supply $\epsilon_L = \frac{\partial \log(X_{ki}^L)}{\partial \log(w_{ki}^L)}$. Then:

$$\frac{\partial \log(w_{ki}^L)}{\partial \log(x_{nki}^L)} = \frac{1}{\epsilon_L} s_{nki}^L. \tag{15}$$

and the markdown equation becomes:

$$\frac{\mu_{nki}^L}{\mu_{nki}^M} = 1 + \frac{1}{\epsilon_L} s_{nki}^L. \tag{16}$$

Thus, a markdown will be high when either the elasticity of labor supply is low or the firm has a large share of the market.[5]

---

[5]In principle, the ratio may also reflect other frictions that can distort the firm's labor input choice away from the price-taking allocation, but we emphasize that the component associated with market share reflects the exercise of monopsony power.

2.4 Labor's Share of Income

Again, following Brooks et al. (2021), we can derive the aggregate factor payment share for labor. Labor's share of value added in this economy is therefore:

$$\eta_L = \frac{\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} w_{ki}^L x_{nki}^L}{\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} (p_{nki}y_{nki} - w_{ki}^M x_{nki}^M)}. \tag{17}$$

Define the labor share of a given firm in the national labor pool as:

$$\omega_{nki}^L = \frac{w_{ki}^L x_{nki}^L}{\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} w_{ki}^L x_{nki}^L}. \tag{18}$$

Then notice by taking the reciprocal of the labor share, we can derive an expression that depends on firm-level labor shares of the national labor pool, and ratios of input expenditure to revenue:

$$\frac{1}{\eta_L} = \sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} \frac{p_{nki}y_{nki}}{w_{ki}^L x_{nki}^L}\omega_{nki}^L - \sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} \frac{w_{ki}^M x_{nki}^M}{w_{ki}^L x_{nki}^L}\omega_{nki}^L. \tag{19}$$

Finally, notice that the ratios of input expenditure to revenue appear in the definitions of the markups. That is:

$$\mu_{nki}^L \equiv \frac{\theta_{nki}^L}{\frac{w_{ki}^L x_{nki}^L}{p_{nki}y_{nki}}}, \qquad \mu_{nki}^M \equiv \frac{\theta_{nki}^M}{\frac{w_{ki}^M x_{nki}^M}{p_{nki}y_{nki}}} \tag{20}$$

where for any input m,

$$\theta_{nki}^m \equiv \frac{\partial \log(F_i)}{\partial \log(x_{nki}^m)}. \tag{21}$$

These imply that:

$$\frac{p_{nki}y_{nki}}{w_{ki}^L x_{nki}^L} = \frac{\mu_{nki}^L}{\theta_{nki}^L}, \qquad \frac{p_{nki}y_{nki}}{w_{ki}^M x_{nki}^M} = \frac{\mu_{nki}^M}{\theta_{nki}^M} \implies \frac{w_{ki}^M x_{nki}^M}{w_{ki}^L x_{nki}^L} = \frac{\mu_{nki}^L \theta_{nki}^M}{\mu_{nki}^M \theta_{nki}^L}.$$

Finally, this can be substituted into equation (19) to get:

$$\frac{1}{\eta_L} = \sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{n=1}^{N_{ki}} \left[\frac{\mu_{nki}^L}{\mu_{nki}^M}\frac{\mu_{nki}^M - \theta_{nki}^M}{\theta_{nki}^L}\omega_{nki}^L \right]. \tag{22}$$

We have only rearranged definitions, to put labor's share into a form where we can easily isolate the role of markdowns (and markups) by constructing and applying counterfactual series of μ_{nki}^L and μ_{nki}^M into the above formula. In particular, looking at equation (16) and setting $s_{nki}^L = 0$, we construct a counterfactual $\frac{\mu_{nki}^L}{\mu_{nki}^M} = 1$ that gives labor's share when monopsony power has been eliminated.

3   Policy Background, Data, and Measures

This section explains the policy initiative, data, and various measures for markdowns that we utilize.

3.1   The Golden Quadrilateral Initiative

The Golden Quadrilateral is a highway system in India that is so named because it connects the largest metropolitan areas across India: Delhi in the north, Calcutta in the east, Chennai in the south, and Mumbai in the west in a circuit. The Quadrilateral spans 5,846 km, making it the largest highway system in India, and among the largest in the world.

The initiative was implemented by the National Highways Authority of India. It connected the four urban areas with an expressway with four to six lanes across for the first time. Prior to its construction no expressway connected the cities: only two percent of national highways were four lanes, over a quarter of national highways were of "poor" road quality, and about a quarter of the roads were considered congested (World Bank, 2002). This system was intended to greatly reduce travel times. For example, the Delhi-Gurgaon Expressway, which is a part of Golden Quadrilateral Highway Project, has reduced the traveling time between Gurgaon and Delhi from 60 minutes to approximately 20 minutes.

It was announced in 1999 and construction started in 2001. The original estimated cost was 600 billion rupees (about \$12.8 billion in 2001), but the project was completed significantly under budget at 250 billion rupees (\$5.3 billion in 2001). Nevertheless, the total cost was sizable; it would have constituted 1 percent of GDP in 2001.

Asturias et al. (2019) geocoded the 127 stretches that comprise the highway. They show that it was built quite rapidly; while only 5% was complete in 2001, 95% was complete by 2006, the original target date.[6] A second stage that involved a cross across the four corners was only ten percent complete by 2006. Hence, we follow Asturias et al. (2019) and focus on the original quadrilateral, using 2006 as the key date. Figure 2 below reproduces

---

[6]The expressway was not fully completed until 2013.

their Figure 1 and shows the Golden Quadrilateral (GQ) completion on top of the previous (non-expressway) road network.

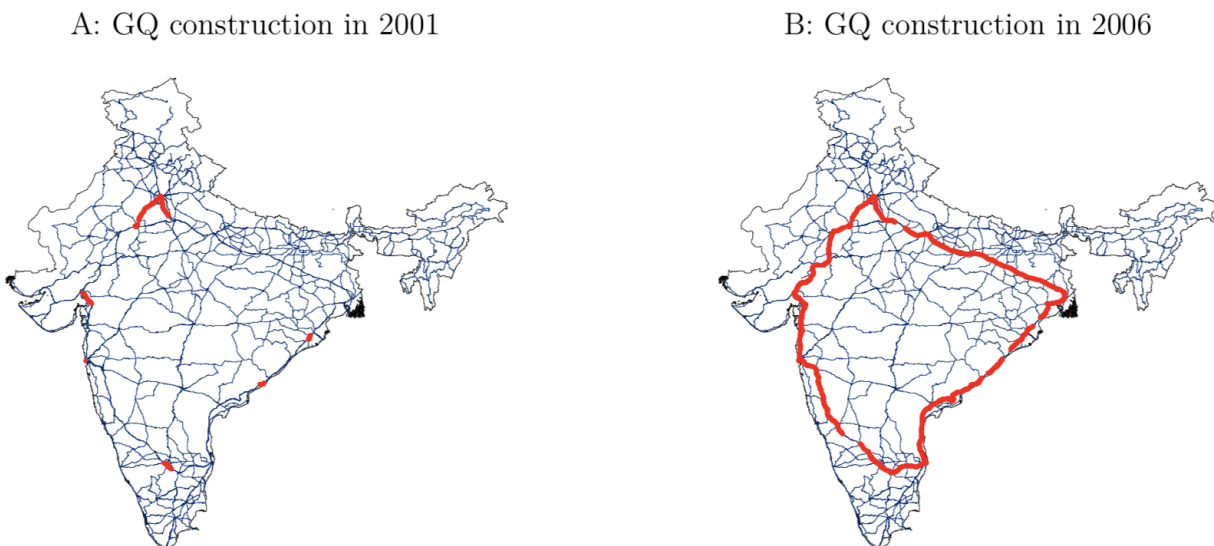A: GQ construction in 2001   B: GQ construction in 2006

Figure 2: Road Network in India and the Golden Quadrilateral (GQ)

Another important policy that impacted labor markets around the same time is the National Rural Employment Guarantee Act of 2005 (NREGA), which guaranteed 100 days a year of government-sponsored manual labor employment to any adult in rural areas willing to work. As this introduced an alternative source of employment, it is possible for this to have reduced monopsony power. However, since it is guaranteed in rural areas it is likely to have impacted those areas far from the GQ, rather than those closer to the GQ. Nevertheless, we consider NREGA in our robustness analyses.

3.2   Data

We utilize two datasets. The road and geospatial data come from the previously mentioned Asturias et al. (2019). Their data is based on geospatial data for all the National Highways of India supplied by ML Infomap, which they augment using information provided by the NHAI on the completion dates of various 127 stretches.

We link these data to the panel version of India's Annual Survey of Industries, which is collected by their Ministry of Statistics and Programme Implementation. These data are establishment level, so we have information on the actual location of production. Although not completely representative, the coverage is broad, containing all large plants (greater than 50 employees) and a sample of smaller plants that depends on the industry and the

number of plants within that industry and state. The approximate number of establishments contained in the sample varies from 23,000 to 44,000 over the years 1999 to 2011. We focus on manufacturing and use the narrow 4-digit industry classification. We utilize measures of output (the value of gross output), material expenditures (the total value of domestic and imported items purchased for production), labor payments (the sum of wage, bonus, and contribution to provident and other funds), and capital (the value of fixed assets, net of depreciation). Labor payment data is used to construct share of the labor market at the district and 4-digit industry level.

Geographically, the finest data we have is at the district level. There are over 600 districts in India, and the highway data measures the shortest straight-line distance between the highway and the most populous city in the district.[7] We then link these data to the firm district to yield estimates of firm distance to the highway. We start by dropping the areas within 50 km of the four major urban centers. We then construct terciles of districts based on their proximity to the completed expressway. Figure 3 shows the three terciles; red are closest (1st tercile) whereas yellow are furthest (3rd tercile). The tercile boundaries around the highway are large. The first tercile includes districts within 55 kilometers of the expressway, while the second includes districts within 191 kilometers. Clearly, these cannot be interpreted directly as commuting zones. However, in a spatial model of partially integrated labor markets, the spatial spillovers can extend far beyond the first order effects on directly connected labor markets.[8][9] By its nature, distance to the highway will be geographically clustered. To assuage concerns, we drop areas within 50 kilometers of the four metro centers, and we also show that our results are robust to separately controlling for coastal areas. Finally, we show that our results are robust to the use of distance to the minimum distance straight-line connection between metro centers, an identification robustness procedure used by Asturias et al. (2019) and Alder and Kondo (2020).

---

[7]The district classification follows the 2011 Census of India. The number of districts has increased from 640 in the year 2011 to 720 as of 2018.

[8]For example, Allen and Arkolakis (2014) find that the impacts of the U.S. interstate system extend far beyond the actual interstates. Similarly, Caliendo et al. (2019) find extensive indirect impacts of China trade shocks on labor markets throughout the United States.

[9]Moreover, in the appendix we show that our results are largely robust to the inclusion of state*time dummies, which indicates that the identification comes off of within-state, across-district variation.
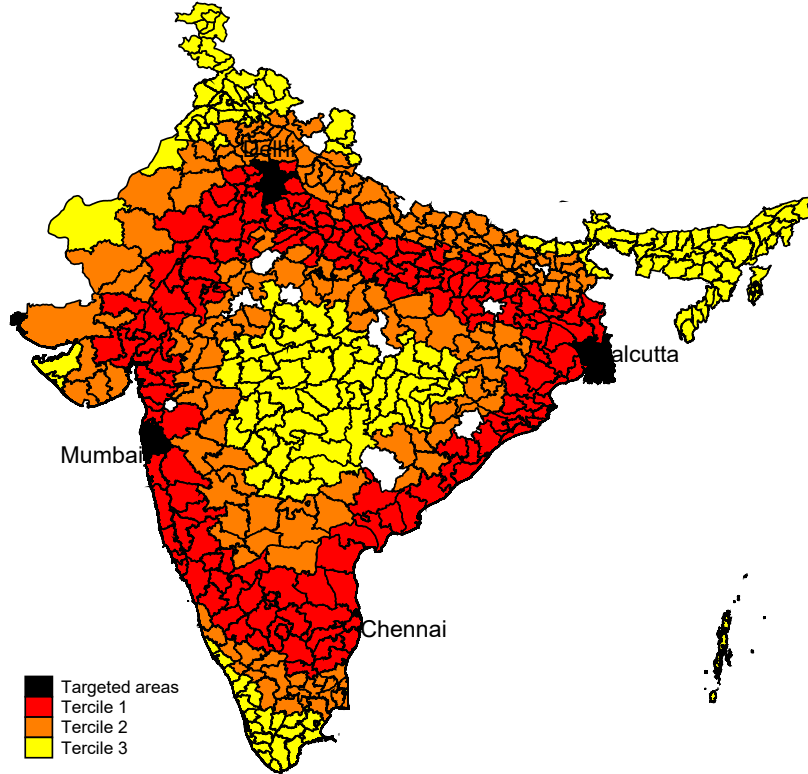
Golden Quadrilateral

Figure 3: Indian Districts by GQ Proximity Terciles

Note: This graph plots the terciles of districts based on their proximity to the completed expressway. The first tercile includes districts within 55 kilometers of the expressway, while the second includes districts within 191 kilometers. The targeted areas are districts within 50 kilometers radius from the center of the four "corners" of the Golden Quadrilateral: Delhi, Calcutta, Chennai, and Mumbai.

3.3 Measuring Markdowns and Markups

We measure markdowns according to equation (16), which requires measuring markups. This can be done in various ways, which we explore in this subsection. Following Brooks et al. (2021), we consider three alternative approaches to measuring markups, all of which they report as giving comparable results.

Using the formula in equation (13) requires an output elasticity, defined as $\theta_{i,t}^M = \frac{\partial \log(F_i)}{\partial \log(x_{nkit}^M)}$. The standard approach, used by de Loecker and Warzynski (2012) (referred to hereafter as DLW), is to estimate the production function by applying the methods of Ackerberg et al. (2015). They estimate translog production functions, which can then be used to easily solve for the output elasticities. Although most common, this approach has some important shortcomings, especially when used in conjunction with DLW to estimate markups. The main limitation is that the production is only identified for the case of either a value-added pro-

12

duction function or a gross output production function in which materials are Leontieff (see Ackerberg et al. (2015) and also Gandhi et al. (2020) for a full explanation). Either of these special cases preclude the estimation of the elasticity of output with respect to materials, the precise parameter necessary to apply the de Loecker and Warzynski (2012) formula.[10] Since this is the standard way of estimating markups (e.g., de Loecker and Warzynski (2012), Edmond et al. (2015), de Loecker et al. (2016), and Brooks et al. (2020)), we present this as one measure, but we allow for several alternatives. We label this markup method "DLW", since it most closely follows their implementation.

Our second method is to estimate markups as the gross profit margin. This is a valid method as long as the production function is constant returns to scale and the firm is price-taking in its inputs (i.e., there is no monopsony power). Again, the second condition is problematic given our focus on monopsony.[11] The precise formula we use is:

$$\mu_{nkit}^{M} = \frac{sales}{costs} = \frac{py}{RK + wL + qM}.\tag{23}$$

where we can measure sales (py), labor payments (wL), and materials expenditures (qM) directly from the data. For capital, we have the stock of capital (K) rather than the payments to capital (RK). The key therefore is to differentiate payments to capital from profits that stem from markups/market power. Notice that the reason this measure of markups is less appropriate in the presence of markdowns is because it attributes all profits (in excess of returns to capital) to markups (higher revenues per unit of output), while some actually would come from markdowns (lower costs per unit of output).

We discipline the return to capital using the cost of capital measured in the data using $R = r + \delta$. We look at the return on corporate bonds, which yields a value of $r = 0.08$. We assume a standard depreciation rate of $\delta = 0.05$ to yield R values of 0.13 in India. This yields an average markup of 1.21. We label this second markup measure as "CRS", which stands for the constant returns to scale assumption.

Our third method uses the markup formula in equation (13), but instead of estimating $\theta_{i,t}^{M}$, it simply assumes that the production function is Cobb-Douglas with respect to materials, i.e., $\theta_{i,t}^{M} = \theta^{M}$. This is a strong assumption on functional form, but it is internally consistent and is therefore our preferred method. This measure of the markup is merely the inverse of the share of materials normalized by a scalar, the elasticity of output with respect to materials, θ^{M}. This scalar cannot be identified from the factor payment share to materials

---

[10]An additional limitation is that it assumes a production function that is constant across firms (within an industry) and only differs by a factor-neutral productivity parameter.

[11]It is, however, less restrictive along other dimensions. It allows for firm-specific production functions that are time-varying, for example. In this sense, it also allows for more general forms of technological change, including factor augmenting technical change.

in the presence of market power. Hence, we need to simply assign a value. Specifically, we choose θ^M so that the average level of these markups equals the average measured using the CRS method.[12] Naturally, the qualitative patterns are independent of this scaling, but the magnitudes depend on it. We refer to this third markup measure as "CD", which stands for Cobb-Douglas. In each case, markups are clearly measured with substantial error. We therefore winsorize 3 percent in both sides of the tails of each 2-digit industry in each year.

In addition, we can measure the labor-based markup μ_{nkit}^L. We measure the labor-based markup again using the CD approach, assuming a constant θ^L. However, since we lack a solid target for markdowns (analogous to our the markup target used to assign θ^M), θ^L cannot be identified in the same way. Instead, we calibrate this elasticity by using a structural prediction of the model. Namely, equation (16) says that if the firm has no market power, then the ratio $\mu_{nkit}^L/\mu_{nkit}^M$ is equal to one. We choose the value of θ^L to satisfy this equation.[13] We again winsorize the 3% tails of the distribution in each 2-digit industry in each year. Notice in the CD case that the markdown becomes materials payments divided by labor payments multiplied by a constant equaling the ratio θ^L/θ^M.

Table 1 gives summary statistics for the most important measures coming from this analysis, including implied values of markups and markdowns. The first three columns give summary statistics based on sampling weighted data using the survey-provided sampling weights. Average markups range from 1.07 (in the CD formula) to 1.42 (in the DLW formula). The average values of the markdowns are 1.01 across all three formulas. The average firm is located 119 km from the completed GQ highway, and has a labor market share of 0.10 in the district-industry.

The last three columns give alternative summary statistics that additionally weight by labor-compensation, which lead to somewhat larger markdowns.[14] The regressions we report utilize survey-provided sampling weights alone. Theory gives us no direction in choosing between raw (sample-weighted) regressions and labor-compensation weighted, since these

---

[12]The reported averages are not equal because of the winsorizing process.

[13]To be more precise, we start with unnormalized values $\tilde{\mu}_{nkit}^L = p_{nkit}y_{nkit}/w_{kit}^L x_{nkit}^L$. We then regress $\tilde{\mu}_{nkit}^L/\mu_{nkit}^M$ on labor share and a constant:

$$\frac{\tilde{\mu}_{nkit}^L}{\mu_{nkit}^M} = \chi + \gamma s_{nkit}^L + u_{nkit}$$

We then set μ_{nkit}^L equal to $\tilde{\mu}_{nkit}^L/\hat{\chi}$, effectively assigning $\theta^L = 1/\hat{\chi}$. This guarantees that the predicted value of that ratio is equal to one when the labor share is zero.

[14]If $s_{i,t}$ is the sampling weight for firm i at time t and $c_{i,t}$ is the average labor compensation of firm i, then the new sampling weight is:

$$\text{weight}_{i,t} = \frac{s_{i,t}c_{i,t}}{\sum_j s_{j,t}c_{j,t}}$$

.

Table 1: Summary Statistics

	Simple Average			Labor-Comp. Weighted		
	Mean	Median	SD	Mean	Median	SD
Markup (DLW)	1.42	1.27	0.51	1.56	1.42	0.66
Markup (CD)	1.07	0.94	0.40	1.21	1.10	0.43
Markup (CRS)	1.21	1.15	0.28	1.34	1.27	0.36
Markdown (DLW)	1.01	0.51	1.43	1.13	0.54	1.78
Markdown (CD)	1.01	0.50	1.48	1.09	0.54	1.82
Markdown (CRS)	1.01	0.53	1.38	1.07	0.58	1.62
log(labor compensation)	14.0	13.8	1.7	18.4	18.3	2.4
Labor's share	0.44	0.31	0.51	0.37	0.24	0.49
Labor market share (state)	0.03	0.003	0.11	0.26	0.10	0.32
Labor market share (district)	0.10	0.02	0.20	0.46	0.33	0.40
Distance to highway (km)	119	72	137	92	22	125
Distance by groups						
Tercile 1	6	0	14	6	0	14
Tercile 2	128	135	39	129	134	38
Tercile 3	299	273	112	295	264	108

Notes: Market shares are computed using 4-digit industries. Labor compensation is in thousand Rupees (in real value). We winsorize the markup, markdown, and labor's share measures above and below the 3rd and 97th percentiles. Columns 1-3 report weighted summary statistics where the weight is the survey-provided sampling weights. Columns 4-6 report weighted summary statistics where the weight is the product of the firms' average labor compensation and the survey-provided sampling weight.

should be identical. Empirically, the magnitudes of coefficients are very similar, but standard errors are larger and so we lose significance when weighting by labor compensation.[15]

Examining these summary statistics, we must acknowledge that the distribution of (normalized) measured markdowns is not consistent with the distribution implied by the theory along a key dimension. In particular, the median markdown ratio is well below one, while the formula in equation (16) implies markdown ratios weakly greater than one. However, we are not distressed by this apparent conflict, since it would follow naturally from measurement error, which is clearly present in the data. Since markdown ratios have the markup in the denominator, classical measurement error in the denominator would lead to non-classical measurement error in markdowns. It would impute a rightward skew to the distribution, which we see, with the mean well above the median.

However, by Jensen's inequality, the average markdown in the raw distribution may be

---

[15]The coefficients on the triple interaction of the firm share with the highway implementation in Table 5 are not inconsistent with the results being weaker for the largest firms. However, if one is concerned that measurement for the smallest firms may be poor, our results are robust to dropping the smallest 20% of firms as shown in Table A.7 of the appendix.

then overstated, so that our rescaling factor may be too large. When we normalize θ^L using the regression constant as above, the normalization may push too many firms below a markdown ratio of one. Again, this scaling will affect the level of markdown ratios and therefore the magnitude of regression coefficients, but it has no effect at all on qualitative patterns nor the implications for the impact of labor market power on labor's share.[16] It will only affect the magnitudes of markdown ratios and the estimated regression coefficients.[17]

The validity and robustness of our markdown measures is an important concern. We can evaluate the validity of our measures on multiple fronts. First, we can examine areas where market power is unlikely. In these four urban areas, which we drop from our analyses, we do not measure positive markups on average, and, while there is still considerable variation across firms, the standard deviation is a third less than outside of these urban centers. Second, our markdown measurement relies on the assumption of Cobb-Douglas production functions, or at least a log linear local approximation to the production functions around the relevant production points for the small changes we measure. Looking more closely, we can examine the markdown of electricity, which is unlikely to be monopsonistic. The implied markdown is remarkably stable (or, equivalently, non existent once properly scaled), reflecting the fact that the average electricity share of expenditures is stable, fluctuating between 13 and 15 percent. Thus, Cobb-Douglas seems reasonable. Third, using the same methods, Brooks et al. (2021) find that the various measures are robust to estimating assumptions, and the results in Figure 1 and the remaining analysis confirm this.[18] Finally, we will also see that our markdowns are correlated with firm labor market share, consistent with the exercise of monopsony power.

---

[16]To see that this has no impact on aggregate labor's share, note that θ^L_{nki} in the denominator of the right-hand side in equation (22) cancels with the implicit θ^L in the numerator of μ^L_{nki}/μ^M_{nki}.

[17]For example, since the sampling-weighted median markdown in Table 1 is 0.5 for the CD benchmark, an alternative normalization in which we rescale so that the median markdown ratio is one would merely imply multiplying markdowns, regression coefficients, and standard errors that follow by 2. This alternative scaling would be consistent with an assumption that measurement error for the median firm is zero *and* the median firm has a negligible market share. The median market share is just 0.003, and this is nearly identical for firms with markdown ratios near the median.

[18]Moreover, the Brooks et al. (2021) results, and those of Hershbein et al. (2020) for the U.S., show that the markdowns are not capturing labor adjustment costs. Hershbein et al. (2020) use a number of other specifications that go beyond Cobb-Douglas, and find very similar results across their specifications.

4 Empirical Analysis

4.1 Main Results

In this section we present our empirical results. Our principle test is to evaluate how connection to the expressway impacts labor market markdowns. We therefore estimate the following equation:

$$\frac{\mu_{nkit}^{L}}{\mu_{nkit}^{M}} = \chi_n + \chi_t + \alpha \ln(py_{nkit}) + \beta_1 * \chi_{tercile=1} * \chi_{t>2006} + \beta_2 * \chi_{tercile=2} * \chi_{t>2006} + u_{nkit}$$

(24)

where χ_n and χ_t are firm- and time-fixed effects; $\chi_{tercile=1}$ and $\chi_{tercile=2}$ are indicators for whether the firm is located in one of the closest or middle terciles of locations from the expressway; and $\chi_{t>2006}$ is an indicator for whether the date is after the "completion" (i.e., 95%) date. We also control for the (log) level of output, py_{nkit}, since the ability to exercise monopsony power may be greater when firms are larger. We exclude those firms located within 50 km of the four major metropolitan "corners" of the Golden Quadrilateral to focus on newly connected areas (although this does not have important impacts on our results). Since the omitted group is the group that is furthest from the GQ, the estimates of interest are β_1 and β_2, the coefficient on the double interactions, as they are interpreted as the impact of connection relative to those firms that remain unconnected. (Since they would be subsumed by the plant and year-specific fixed effects, we do not include the tercile indicators or completion indicator as direct effects.)

The estimates are presented in Table 2. (Standard errors are clustered at the district level throughout the results we present.) As anticipated by theory, firm size is substantially associated with an ability to exercise monopsony power. More to the point, however, we see that relative to the unconnected firms in the third tercile, the firms in the first and second tercile show smaller markdowns after connection, and the β_1 and β_2 estimates are significant at the 1 percent level. The magnitude of the first and the second terciles are quantitatively similar and indeed, statistically indistinguishable. Moreover, the estimates are very similar across the three alternative measures of markdowns. The estimates are also economically significant: recalling that the average markdown was about 1.15 percent across these specifications, we cannot reject that connection to the GQ fully eliminates the exercise of monopsony power.

Table 2: Impact of Highway on Markdown

	(1)	(2)	(3)
	Markdown (DLW)	Markdown (CD)	Markdown (CRS)
$\mathbb{1}\{\text{tercile=1}\} \times \mathbb{1}\{\text{t>2006}\}$	-0.073***	-0.072***	-0.069***
	(0.021)	(0.021)	(0.020)
$\mathbb{1}\{\text{tercile=2}\} \times \mathbb{1}\{\text{t>2006}\}$	-0.088***	-0.087***	-0.082***
	(0.023)	(0.024)	(0.022)
log(output)	0.369***	0.365***	0.320***
	(0.015)	(0.015)	(0.014)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	238,456	238,456	238,456
Adjusted R-squared	0.845	0.850	0.848

Notes: Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

4.2 Robustness

Next we study the robustness of these results to additional controls or specifications. The results of these exercises are given in the appendix.

First, while we cannot allow for district-specific year-to-year variation, in Table A.1 we add state-time dummies as controls which allow for year-to-year variation to differ at a subnational level. For some states and years, these state-specific time patterns are significant, but our results are quite robust, which indicates that the identification is driven by within-state, across-district variation.

Second, Figure 3 shows that many coastal areas are close the to GQ, and perhaps coastal districts had a differential time pattern. We allow for this by creating a dummy for the coastal district, and interacting it with time dummies as well as shown in Table A.2. The time trend is insignificant, and the results are virtually unchanged.

Third, we have noted NREGA, which was enacted around the same time and may have disproportionately impacted rural districts relative to more urban districts. In Table A.3 we allow for two district-level controls: the number of job cards per capita, which measures the extensive margin or breadth of the program, and the total per capita labor expenditures, which measures the intensive margin or depth of the program. The former is ultimately significant, and associated with a larger markdown, but it does not impact our result for the impact of the GQ.

Finally, to account for potentially endogenous placement of roads, we also construct terciles using the distance to a straight line, minimum distance connection between the metro centers, and our results are again robust. In sum, the results are quite robust, and this is true for each table we present. These results are given in Table A.4.

To get a better sense of the year-to-year identification, we start by grouping the first two terciles together (since their magnitudes were similar). We then interact this group with every year in the sample (instead of relying on the post-2006 sample). That is, we estimate:

$$\frac{\mu^L_{nkit}}{\mu^M_{nkit}} = \chi_n + \chi_t + \alpha ln(py_{nkit}) + \sum \beta_t * \chi_{tercile=1,2} * \chi_t + u_{nkit} \qquad (25)$$

The estimates of β_t over time, along with their 95% confidence intervals, are presented visually in Figure 4. The omitted year is completion year of 2006, which is therefore normalized to zero. We see that prior to 2006, the estimates are not statistically different from zero. After 2006, however, the coefficients show a strong break becoming negative and significantly so. However, using a Chow test, we reject a single linear trend in the data at the 5% level, i.e., a structural break in 2006 exists in the data.

We therefore find strong evidence for the decline of markdowns after the road is connected. We now try to understand this finding in more depth.

One possible explanation for the declining markdown measure is that it fell because markups rose. Recall that the markdown is the ratio of two measures of "markups": the labor measured markup over the true markup. One reason that the markdown could fall is that the true markup rises. Another reason to look at markups is because the markdown reflects the ratio of materials payments to labor payments. A decrease in materials payments, due to access to cheaper inputs that the expressway opened up together with a low elasticity of substitution toward inputs, for example, could appear as an increase in markdowns. This would also show up as an increase in markups.

For both of these reasons, we examine the impact of the expressway on the markups themselves. That is, we run regressions analogous to those in Table 2, but where the dependent variable is the markup rather than the markdown.

The results are presented in Table 3. We find mixed evidence regarding the impact of highway on markups. The coefficients on the double interactions are small and statistically insignificant when using the CD and DLW measures. For the CRS estimates, the magnitude is larger (at least five times as large), and the impact on the first and second terciles is statistically significant. Recall, however, that a problem with the CRS measure of markups is that the gross profit margin can incorporate profitability that comes from markdowns. Hence, these larger results are not inconsistent with the results that markdowns fell in these

Figure 4: Impact of Highway Proximity on Markdowns over Time
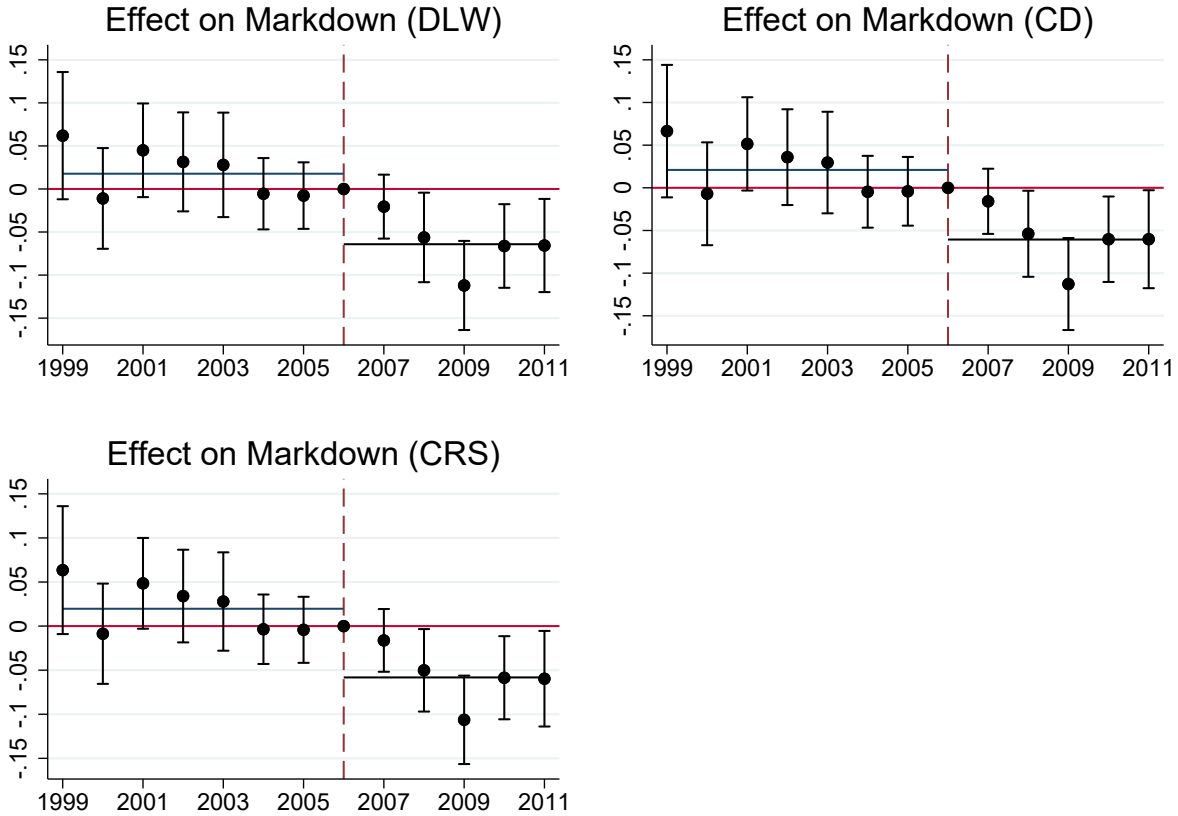
Notes: This graph plots the estimates of the effect of highway on markdown over time, along with their 95% confidence intervals. The vertical dashed line indicates the year of highway completion. The horizontal black lines indicate the simple time-average of the estimates before and after the highway completion.

areas. In any case, the result that connection reduces markups is consistent with the idea that the expressway lowered not only travel costs but trade costs as well, and induced more competition into these regions. Indeed, this is the central argument in Asturias et al. (2019).

Figure 5 is the analog to Figure 4 for markups. The trends are less stark, and, except in the case of the CRS measure, not statistically different from zero after the highway.

In sum, there is some evidence that markups fell as a result of the GQ, but this evidence is weaker than that for markdowns. In any case, a *drop* in markups cannot be a contributing factor to the observed decline in markdowns as the markup is in the denominator of the markdown.

Table 3: Impact of Highway on Markup

	(1) Markup (DLW)	(2) Markup (CD)	(3) Markup (CRS)
$\mathbb{1}\{\text{tercile=1}\} \times \mathbb{1}\{\text{t>2006}\}$	-0.009	-0.002	-0.025***
	(0.008)	(0.006)	(0.006)
$\mathbb{1}\{\text{tercile=2}\} \times \mathbb{1}\{\text{t>2006}\}$	0.004	0.003	-0.014**
	(0.009)	(0.007)	(0.006)
log(output)	-0.017***	-0.018***	0.083***
	(0.004)	(0.003)	(0.004)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	238,456	238,456	238,456
Adjusted R-squared	0.690	0.715	0.650

Notes: Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

4.3 Labor Compensation

We now explore payments to labor more closely. In particular, we continue to combine the first two terciles into a single group, and we examine other dependent variables impacted by the expressway connection: labor compensation, labor's share, and the shares of labor in the market at either the state or district level. The results are shown in Table 4.

Table 4: Impact of Highway on Labor Payment Measures

	(1) log(Labor compensation)	(2) labor's share	(3) Labor market share (state)	(4) Labor market share (district)
$\mathbb{1}\{\text{tercile=1,2}\}$ $\times \mathbb{1}\{\text{t>2006}\}$	0.111***	0.048***	0.005***	0.008***
	(0.016)	(0.009)	(0.002)	(0.003)
log(output)	0.539***	-0.206***	0.009***	0.015***
	(0.006)	(0.007)	(0.001)	(0.001)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	238,456	238,456	238,456	238,456
Adjusted R-squared	0.957	0.506	0.814	0.760

Notes: Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Figure 5: Impact of Highway Proximity on Markups over Time
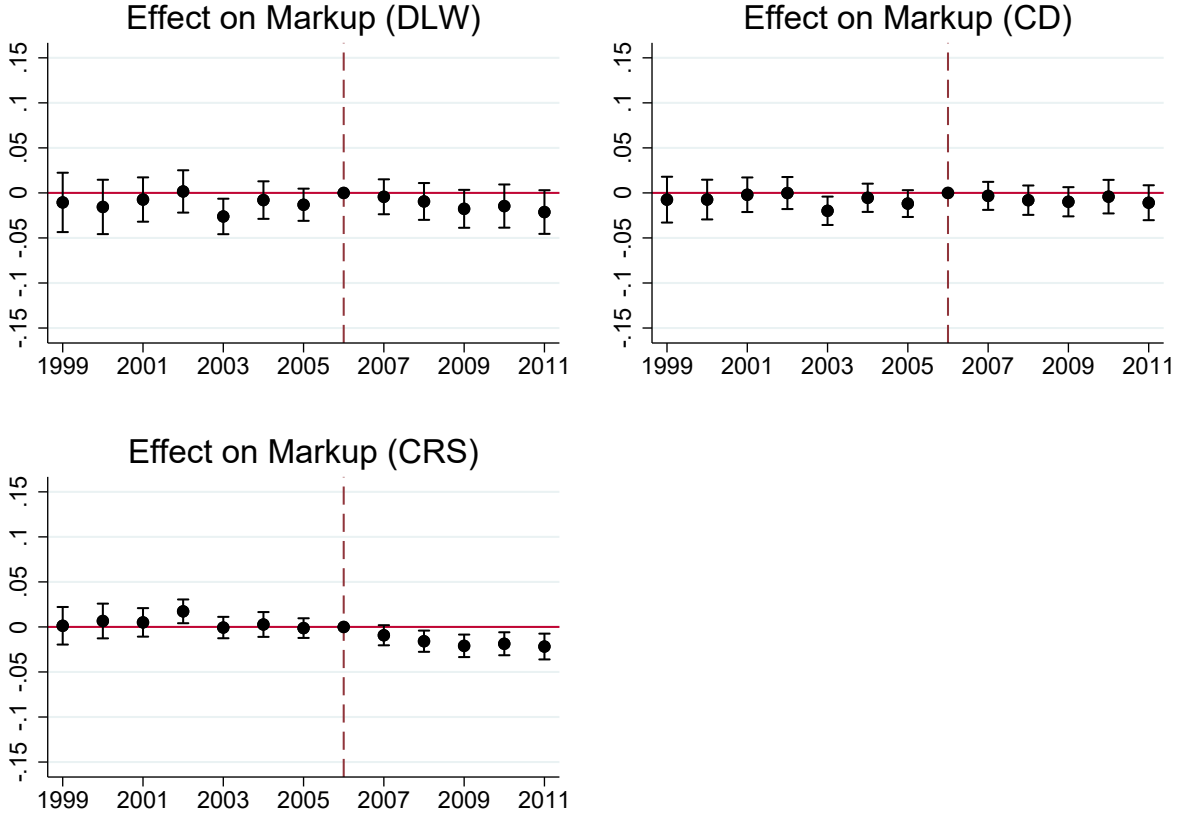
Notes: This graph plots the estimates of the effect of highway on markup over time, along with their 95% confidence intervals. The vertical dashed line indicates the year of highway completion.

Focusing on column (1), it is not surprising that (log) labor compensation is strongly correlated with size. Although the coefficient may seem low relative to advanced economies, it is comparable to labor's payment share as shown in Table 1. Again, more to the point, the coefficient on the interaction term shows that labor compensation increases by 11 (log) percentage points for those newly connected firms. Moreover, in column (2), we see that labor's share increases by 5 percentage point. Hence, we have direct evidence that labor compensation and labor's share is directly impacted. [19]

Columns (3) and (4) examine how connection to the expressway impacts the market share in the labor market. We find a significant increase in the market share of connected firms. The market share increases by 0.5 percentage points in the state-industry and 0.8

---

[19]The magnitude of the change in labor's share is lower than the magnitude of the drop in the markdown, however. Presumably this reflects a negative correlation between increases in labor compensation and increases in material expenditures across firms in connected districts.

percentage points in the district-industry. Recall equation (16): in theory markdowns are increasing in labor market share and decreasing in the elasticity of labor supply. Therefore, the drop in markdowns of connected firms cannot be explained by decreases in market share in the labor market.

Finally, in Figure 6 we check for the possibility of pre-existing trends in labor's share and labor compensation. We find no evidence at all of a pre-trend in labor's share as shown in the left panel. The right panel, however, exhibits three years that are significantly below zero. Nevertheless, applying a linear trend and using a Chow test for structural break in trend, we cannot reject a straight line, i.e., the lack of a structural break in trend in 2006. Again, the upward pretrend would not be inconsistent with the gradual construction of the expressway, of course, and so our results are best interpreted as a before and after comparison indicated by the horizontal lines which are the average pre- and post-2006 coefficients.

Figure 6: Impact of Highway Proximity on Labor Payment Measures over Time

Notes: This graph plots the estimates of the effect of highway over time on labor's value-added share and log of labor compensation, along with their 95% confidence intervals. The vertical dashed line indicates the year of highway completion.

23

4.4 Effect on Elasticity of Labor Supply

Next, we examine whether the effective elasticity of labor supply changed in response to the expressway connection. We have a fixed geographic sense of a labor market, but the expressway may have changed the effective pool from which workers can be drawn. In that sense, the labor supply elasticity is not necessarily picking up the elasticity of an individual worker or even a fixed set of workers, but the elasticity of supply that comes from neighboring areas responding to wages, or, in the case of a markdown, the ability of workers to find employment in neighboring areas if wages were suppressed.

To operationalize this, we start with equation (16) as motivation for the following regression equation:

$$\frac{\mu_{nkit}^L}{\mu_{nkit}^M} = \chi_n + \chi_t + \alpha ln(py_{nkit}) + \gamma_1 * s_{nkit}^L + \gamma_2 * \chi_{tercile=1,2} * s_{nkit}^L +$$

$$\gamma_3 * \chi_{tercile=1,2} * \chi_{t>2006} + \gamma_4 * \chi_{tercile=1,2} * \chi_{t>2006} * s_{nkit}^L + u_{nkit} \quad (26)$$

where s_{nkit}^L is again the firm's share in the labor market. Here the triple interaction coefficient γ_4 is of particular interest as it reflects the post-access change in the inverse labor supply elasticity. If the elasticity of labor supply increased, we would expect this coefficient to be negative.

In estimating a different version of (16), however, Brooks et al. (2021) note that the firm's labor compensation shows up in the denominator of the markdown and the numerator of the labor market share. Hence, any measurement error will cause a spurious correlation. To avoid this, they use a two-stage regression, which we also adopt, instrumenting for labor market share by using the firm's share in the local product market. Intuitively, the more output a firm produces, the more labor it should hire, but this should not directly affect the ratio of materials and labor payments (i.e., the markdown). (In the case that factor ratio increases with the scale of operations, we can add log output as a separate control in the regression.)

Table 5 shows the second stage results of these equations. (The first stage is quite strong and is included as Table A.5 in the appendix.) Column (1) measures the labor market share at the state level, while column (2) measures it at the district level. The estimated coefficient on firm's labor market share of 0.319 in column (1) implies a labor supply elasticity of about 3.1 at the state level before 2006, whereas the coefficient of 0.092 at the district level imply an elasticity of 10.3 at the district level. Both estimates are strongly statistically significant and the larger elasticity at the district level is consistent with more integration in the labor

market across districts than across states. [20]

Nevertheless, our primary focus is the interaction terms with post-2006, and those firms in proximity to the expressway. These coefficients are puzzling, however. The triple interaction coefficient, γ_4 above, is insignificant at the state level and the district level. Instead, we see that firms' markdowns fell in connected areas after the expressway is completed, but that decline happens uniformly across all firms rather than disproportionately for firms who have more labor market power. Thus, the observed decline in markups does not have the direct interpretation of an increase in labor supply elasticity that comes from increased labor market integration.

Table 5: Relationship between Markdown and Labor Market Share

	(1)	(2)
	Dependent Variable = Markdown (CD)	
	State-level Labor Market	District-level Labor Market
Firm's Share	0.319***	0.092***
	(0.050)	(0.034)
Firm's Share × $\mathbb{1}\{t>2006\}$	-0.033	-0.016
	(0.039)	(0.028)
Firm's Share × $\mathbb{1}\{tercile=1,2\}$	0.003	0.005
	(0.063)	(0.042)
Firm's Share × $\mathbb{1}\{tercile=1,2\}$× $\mathbb{1}\{t>2006\}$	0.076	0.072**
	(0.062)	(0.036)
$\mathbb{1}\{tercile=1,2\}$× $\mathbb{1}\{t>2006\}$	-0.057**	-0.069***
	(0.022)	(0.026)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	104,813	104,813
First Stage F-statistics	1423	4608
$\epsilon_L^{before\ highway}$	3.10***	10.31***
	(0.37)	(2.54)
$\epsilon_L^{after\ highway}$	2.74***	6.53***
	(0.42)	(0.94)

Notes: We define labor markets as industry and location specific. We choose 4-digit industry level. The location level is chosen as state (in specification 1) or district (in specification 2). In all regressions, we use firm's share in the local product market as an instrument for its labor market share. Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

We have also examined whether the decline in markdown is instead a spurious artifact

---

[20]The coefficients drop precipitously once we control for log output and the implied elasticities become quite large, 12 and 44, respectively. We view these elasticities as implausibly large, and we therefore interpret the log output control as a second proxy for labor market share.

of a decline in input prices that comes from connection. A decline in the prices of materials, together with a low elasticity of substitution with respect to materials, could lower the ratio of material expenditures to labor compensation, which we might falsely interpret as a decline in markdown. We indeed find that connection leads to significantly lower input prices, and the measured markdowns can be significantly decreasing in materials prices, but this does not undermine our measured impact of the highway connection on markdowns as shown in Table A.6 of the appendix.

In sum, the decrease in markdowns is not easily explained with the existing theory as a decline in monopsony power coming from either lower labor market shares or increased labor supply elasticity. It remains an open question.

4.5 Aggregate Labor Share

We last examine the impact of monopsony power and the change in monopsony power that comes from the expressway on aggregate labor's share in the manufacturing sector. Following equation (4), we calculate the actual and counterfactual labor's share in the data in every year. We construct two counterfactuals: labor's share in the absence of all markdowns (which is higher than observed) driven by market-power and labor's share in the absence of highway (which is lower than observed). The latter counterfactual is constructed by subtracting the coefficients in the CD case of Table 2 from the firm-specific markdowns of those firms located closer to the expressway.

The results are below in Figure 7. The solid line shows the actual, observed movements in labor's share in manufacturing. The dashed line above it shows the counterfactual labor's share in the absence of market-power driven markdowns. The overall impact of monopsony is sizable, especially early in the sample, confirming the results of Brooks et al. (2021). It ranges from a high of 7 percentage points (in 1999) to a low of 3 percentage points (in 2009).[21]

Monopsony can therefore explain a substantial amount of why labor's share is low in Indian manufacturing. Starting in 2006, we add the dash-dot line, which shows the impact of the expressway by plotting what the counterfactual labor's share would have been in the absence of the highway. Labor's share in this counterfactual world is smaller, indicating that the expressway increased observed labor's share by between 1.8 (2007) and 2.3 (2011) percentage points. The impact of the expressway is smaller for three reasons. First, the expressway does not completely eliminate monopsony power for firms closer to it. Second,

---

[21]It is surprising that weighted markdowns are rising over time in Figure 1, but their impact on labor's share is falling over time. We conjecture that this is driven by a changing correlation between markdowns and labor's share.

not all firms are impacted by the expressway. Third, the impact occurs in a time where monopsony power is less important for labor's share.
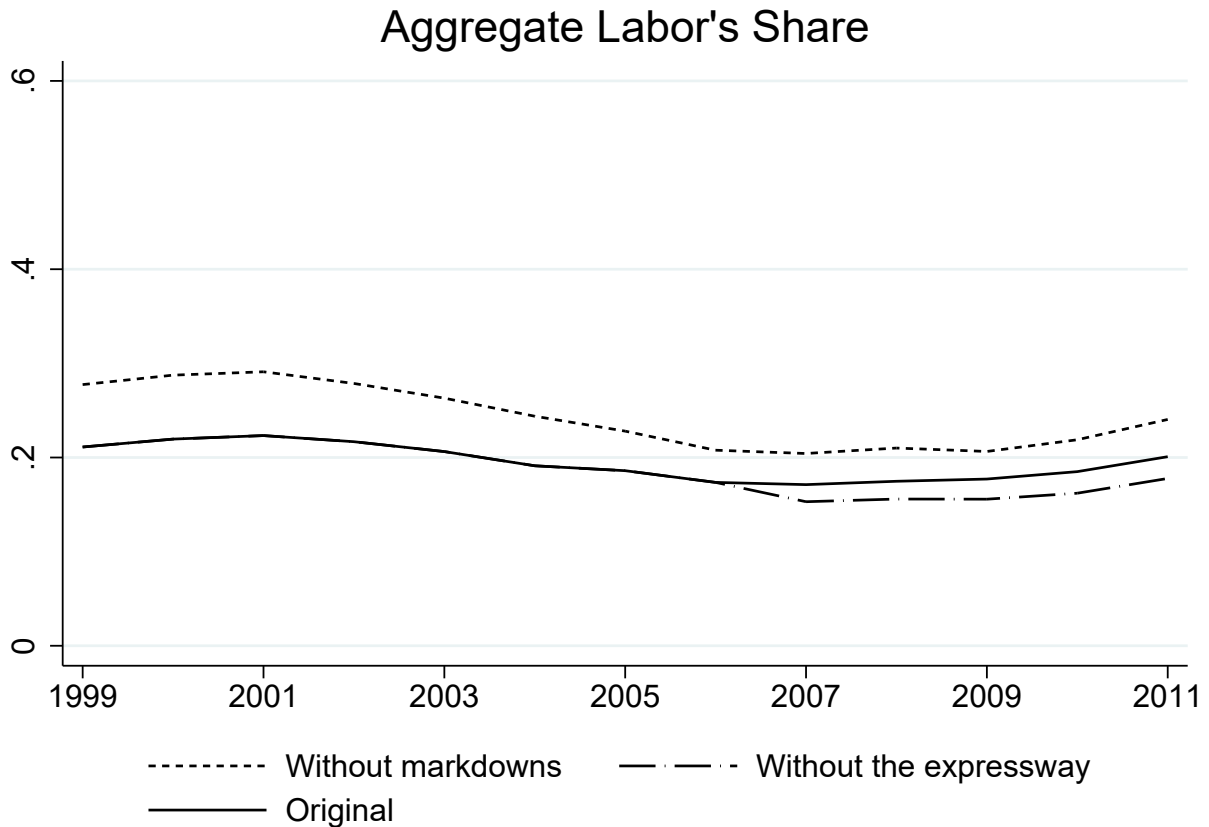

Figure 7: Aggregate Labor's Share in Manufacturing: Observed and Counterfactual

Notes: This graph plots the time path of the observed aggregate labor's share (solid line), the implied labor's share in a counterfactual economy without estimated markdowns (dashed line), and the implied labor's share in a counterfactual economy without the expressway (dash-dotted line).

In sum, our empirical results show a sizable impact for the expressway both in decreasing markdowns and increasing labor's share.

5 Conclusion

This paper presents some of the first evidence available on the role of infrastructure in affecting firm markdowns by evaluating the impact of India's Golden Quadrilateral expressway expansion on the monopsony power in the labor market. Firms in districts closer to the highway exhibit significantly lower markdowns than firms in more remote areas. The impact of highway on markdowns is substantial: the average markdowns are effectively eliminated

27

for firms within close proximity to the highway. These lower markdowns have raised labor's share by about 2 percentage points.

The causes of the lower markdowns for connected firms remain a puzzle. We show that the impact of highway on markdowns cannot be explained by a decrease in markups. We also find no evidence that labor supply elasticities increase when districts are connected to the highway. This brings into question whether the decrease in markdown is driven by increased integration across labor markets.

We must leave to future research the goal of identifying clear mechanisms. Nevertheless, the results in this paper are important in showing highway infrastructure investment reduces monopsony power and promotes competition in the marketplace.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer, "Structural identification of production functions," *Econometrica*, 2015, *83* (6), 2411–2415.

Alder, Simon, "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development," *Working Paper*, 2016.

_ **and Illenin Kondo**, "Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis," *Working Paper*, 2020.

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

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

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

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor Market Power," *Society for Economic Dynamics*, 2018. Meeting Papers 170.

Binswanger, Hans P. and eds. Mark R. Rosenzweig, *Contractual Arrangements, Employment, and Wages in Rural Labor Markets in Asia.*, New Haven: Yale University Press, 1984.

Braverman, Avishay and Joseph E. Stiglitz, "Sharecropping and the Interlinking of Agrarian Markets.," *American Economic Review*, 1982, *72* (4), 695–715.

_ **and** _ , "Landlords, Tenants, and Technological Innovations.," *Journal of Development Economics*, 1986, *23*, 313–332.

Brooks, Wyatt and Kevin Donovan, "Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua," *Econometrica*, 2020, *88* (5), 1965–1997.

Brooks, Wyatt J, Joseph P Kaboski, Yao Amber Li, and Wei Qian, "Exploitation of Labor? Classical Monopsony Power and Labor's Share," *Journal of Development Economics*, 2021. Forthcoming.

Brooks, Wyatt, Joseph P. Kaboski, and Amber Yao Li, "Agglomeration, Misallocation, and (the Lack of) Competition," *American Economic Journal: Macroeconomics*, 2020. Forthcoming.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

de Loecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, October 2012, *102* (6), 2437–71.

__ , **Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik**, "Prices, Markups and Trade Reform," *Econometrica*, March 2016, *84* (10), 445–510.

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

Edmond, Chris, Virgiliu Midrigan, and Daniel Xu, "Competition, Markups and the Gains from International Trade," *American Economic Review*, 2015, *105* (10), 3183–3221.

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

Gandhi, Amit, Salvador Navarro, and David Rivers, "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Gouin-Bonenfant, Emilien, "Productivity Dispersion, Between-firm Competition and the Labor Share," *Society for Economic Dynamics.*, 2018. Meeting Papers 1171.

Hershbein, Brad, Claudia Macaluso, and Chen Yeh, "Monopsony in the U.S. Labor Market," *Working Paper*, 2020.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market," Technical Report w25954, National Bureau of Economic Research 2019.

Lu, Will Jianyu, "Transport, Infrastructure and Growth: Evidence from Chinese Firms," *Working Paper*, 2020.

Morlacco, Monica, "Market Power in Input Markets: Theory and Evidence from French Manufacturing," *Working paper*, 2019.

World Bank, "India's Transport Sector: The Challenges Ahead," *World Bank Reports Series*, 2002, *1*, 1–65.

Appendix

Table A.1: Impact of Highway on Markdown Controlling for State-specific Secular Trends

	(1) Markdown (DLW)	(2) Markdown (CD)	(3) Markdown (CRS)
$\mathbb{1}\{\text{tercile}=1\} \times \mathbb{1}\{t>2006\}$	-0.095***	-0.093**	-0.086**
	(0.033)	(0.037)	(0.035)
$\mathbb{1}\{\text{tercile}=2\} \times \mathbb{1}\{t>2006\}$	-0.110***	-0.111***	-0.104***
	(0.032)	(0.036)	(0.034)
log(output)	0.369***	0.365***	0.320***
	(0.014)	(0.015)	(0.013)
Firm FE	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes
Observations	238,456	238,456	238,456
Adjusted R-squared	0.846	0.851	0.848

Notes: Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.2: Impact of Highway on Markdown Controlling for Secular Trends in Coastal Areas

	(1) Markdown (DLW)	(2) Markdown (CD)	(3) Markdown (CRS)
$\mathbb{1}\{\text{tercile=1}\} \times \mathbb{1}\{t>2006\}$	-0.076***	-0.074***	-0.071***
	(0.020)	(0.021)	(0.019)
$\mathbb{1}\{\text{tercile=2}\} \times \mathbb{1}\{t>2006\}$	-0.087***	-0.086***	-0.082***
	(0.024)	(0.024)	(0.023)
log(output)	0.369***	0.365***	0.320***
	(0.014)	(0.015)	(0.014)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2000\}$	-0.048	-0.047	-0.046
	(0.035)	(0.037)	(0.035)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2001\}$	-0.022	-0.018	-0.021
	(0.030)	(0.032)	(0.030)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2002\}$	-0.037	-0.035	-0.038
	(0.031)	(0.034)	(0.032)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2003\}$	-0.025	-0.025	-0.031
	(0.033)	(0.035)	(0.033)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2004\}$	-0.058*	-0.064*	-0.065*
	(0.034)	(0.037)	(0.034)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2005\}$	-0.054	-0.058	-0.062
	(0.039)	(0.041)	(0.038)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2006\}$	-0.019	-0.027	-0.032
	(0.037)	(0.040)	(0.037)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2007\}$	-0.023	-0.029	-0.036
	(0.038)	(0.041)	(0.038)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2008\}$	-0.005	-0.011	-0.017
	(0.045)	(0.047)	(0.044)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2009\}$	-0.024	-0.037	-0.043
	(0.046)	(0.049)	(0.046)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2010\}$	-0.028	-0.036	-0.042
	(0.042)	(0.045)	(0.043)
$\mathbb{1}\{\text{coastal district}\} \times \mathbb{1}\{t=2011\}$	-0.033	-0.038	-0.042
	(0.047)	(0.050)	(0.047)
Firm FE	Yes	Yes	Yes
$\mathbb{1}\{\text{coastal district}\} \times$ Year FE	Yes	Yes	Yes
Observations	238,456	238,456	238,456
Adjusted R-squared	0.845	0.850	0.848

Notes: The name of the coastal districts in India comes from the Centre for Coastal Zone Management and Coastal Shelter Belt, hosted by Institute for Ocean Management, Anna University Chennai (http://iomenvis.nic.in/index2.aspx?slid=3680&sublinkid=259&langid=1&mid=1). Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.3: Impact of Highway on Markdown Controlling for NREGA Impact

	(1) Markdown (DLW)	(2) Markdown (CD)	(3) Markdown (CRS)
$\mathbb{1}\{\text{tercile}=1\} \times \mathbb{1}\{t>2006\}$	-0.081***	-0.080***	-0.077***
	(0.021)	(0.021)	(0.019)
$\mathbb{1}\{\text{tercile}=2\} \times \mathbb{1}\{t>2006\}$	-0.095***	-0.096***	-0.091***
	(0.025)	(0.026)	(0.024)
log(output)	0.370***	0.365***	0.321***
	(0.015)	(0.016)	(0.014)
nrega-job card	0.324*	0.384**	0.365**
	(0.187)	(0.194)	(0.184)
nrega-labor expenditure	-6.414*	-6.454	-6.659*
	(3.876)	(3.990)	(3.826)
Firm FE	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes
Observations	225,185	225,185	225,185
Adjusted R-squared	0.846	0.851	0.849

Notes: Information about district-level job cards and labor expenditure distributed by NREGA comes from the Ministry of Rural Development (https://nregarep2.nic.in/netnrega/dynamic2/dynamicreport_MPR.aspx). We scale the number of job cards and labor expenditure by the total population in each district. We take the inverse hyperbolic sine transformation of the series and assign zero to years before the program starts. Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.4: Impact of Highway on Markdown Using Alternative Distance Measure

	(1) Markdown (DLW)	(2) Markdown (CD)	(3) Markdown (CRS)
$\mathbb{1}\{\text{tercile}=1\} \times \mathbb{1}\{t>2006\}$	-0.062***	-0.063***	-0.060***
	(0.022)	(0.023)	(0.022)
$\mathbb{1}\{\text{tercile}=2\} \times \mathbb{1}\{t>2006\}$	-0.052**	-0.059**	-0.058**
	(0.024)	(0.024)	(0.023)
log(output)	0.369***	0.365***	0.320***
	(0.015)	(0.015)	(0.014)
Firm FE	Yes	Yes	Yes
State×Year FE	Yes	Yes	Yes
Observations	238,456	238,456	238,456
Adjusted R-squared	0.845	0.850	0.848

Notes: The grouping of highway proximity in this table is based on the distance to the straight line connecting the four and five vertices of the GQ provided by Asturias et al. (2019). Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.5: Relationship between Markdown and Labor Market Share - First Stage

	(1)	(2)
	Dependent Variable = Firm's Labor Market Share	
	State-level Labor Market	District-level Labor Market
Output Share	0.718***	0.834***
	(0.013)	(0.012)
Output Share × $\mathbb{1}\{t>2006\}$	0.102***	0.042***
	(0.013)	(0.009)
Output Share × $\mathbb{1}\{tercile=1,2\}$	0.036**	0.012
	(0.016)	(0.013)
Output Share × $\mathbb{1}\{tercile=1,2\}$ × $\mathbb{1}\{t>2006\}$	-0.056***	-0.009
	(0.016)	(0.011)
Lagged Labor Market Share	0.136***	0.058***
	(0.011)	(0.009)
Lagged Labor Market Share × $\mathbb{1}\{t>2006\}$	-0.098***	-0.039***
	(0.014)	(0.009)
Lagged Labor Market Share × $\mathbb{1}\{tercile=1,2\}$	-0.054***	-0.029***
	(0.015)	(0.010)
Lagged Labor Market Share × $\mathbb{1}\{tercile=1,2\}$ × $\mathbb{1}\{t>2006\}$	0.058***	0.009
	(0.017)	(0.011)
$\mathbb{1}\{tercile=1,2\}$ × $\mathbb{1}\{t>2006\}$	0.001	-0.001
	(0.001)	(0.001)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	104,813	104,813
Adjusted R-squared	0.962	0.971

Notes: We define labor markets as industry and location specific. We choose 4-digit industry level. The location level is chosen as state (in specification 1) or district (in specification 2). In all regressions, we use firm's share in the local product market as an instrument for its labor market share. Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.6: Relationship between Markdown and Labor Market Share Controlling for Intermediate Input Price

	(1)	(2)	(1)	(2)
		Dependent Variable = Markdown (CD)		
	State-level Labor Market		District-level Labor Market	
Firm's Share	0.311***	2.303***	0.093***	1.180***
	(0.048)	(0.153)	(0.035)	(0.092)
Firm's Share $\times$ $\mathbb{1}\{t>2006\}$	-0.019	-0.037	-0.004	-0.002
	(0.037)	(0.037)	(0.025)	(0.025)
Firm's Share $\times$ $\mathbb{1}\{tercile=1,2\}$	-0.013	-0.072	0.000	0.009
	(0.066)	(0.073)	(0.043)	(0.045)
Firm's Share $\times$ $\mathbb{1}\{tercile=1,2\}\times$ $\mathbb{1}\{t>2006\}$	0.057	0.059	0.058*	0.055
	(0.061)	(0.065)	(0.033)	(0.034)
$\mathbb{1}\{tercile=1,2\}\times$ $\mathbb{1}\{t>2006\}$	-0.047**	-0.046**	-0.058**	-0.055**
	(0.021)	(0.021)	(0.024)	(0.024)
log(input price)	0.007***	0.040***	0.007***	0.050***
	(0.002)	(0.003)	(0.002)	(0.004)
log(input price) * Firm's share		-0.228***		-0.128***
		(0.015)		(0.008)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	99,897	99,897	99,897	99,897
First Stage F-statistics	1353	153.5	5021	445

Notes: We define labor markets as industry and location specific. We choose 4-digit industry level. The location level is chosen as state (in specification 1 and 2) or district (in specification 3 and 4). In all regressions, we use firm's share in the local product market as an instrument for its labor market share. A firm's intermediate input price is measured as the average of all input prices weighted by the total value of each input. Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.

Table A.7: Relationship between Markdown and Labor Market Share Dropping Small Firms

	(1)	(2)
	Dependent Variable = Markdown (CD)	
	State-level Labor Market	District-level Labor Market
Firm's Share	0.307***	0.094***
	(0.049)	(0.036)
Firm's Share × 𝟙{t>2006}	-0.016	-0.013
	(0.036)	(0.028)
Firm's Share × 𝟙{tercile=1,2}	0.012	0.004
	(0.062)	(0.043)
Firm's Share × 𝟙{tercile=1,2}× 𝟙{t>2006}	0.062	0.068*
	(0.060)	(0.035)
𝟙{tercile=1,2}× 𝟙{t>2006}	-0.049**	-0.066***
	(0.020)	(0.026)
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	103,482	102,527
First Stage F-statistics	1423	4782
$\epsilon_L^{\text{before highway}}$	3.14***	10.28***
	(0.38)	(2.54)
$\epsilon_L^{\text{after highway}}$	2.74***	6.56***
	(0.42)	(0.94)

Notes: This table excludes small firms in the bottom 10 percent of the distribution of labor market shares. We define labor markets as industry and location specific. We choose 4-digit industry level. The location level is chosen as state (in specification 1) or district (in specification 2). In all regressions, we use firm's share in the local product market as an instrument for its labor market share. Observations in these regressions are weighted by survey-provided sampling weights. Standard errors clustered at the district level are in parentheses. *** significant at 1%. ** significant at 5%. * significant at 10%.