Road to Productivity: The Effect of Roads on Manufacturing TFP in India

Subham Kailthya* Uma Kambhampati†

This version: August 23, 2018
Recent version: click here

Abstract

In this article, we examine the impact of roads on manufacturing total factor productivity (TFP). High road density lowers transport cost, which boosts manufacturing TFP. We exploit exogenous variation in the partisan alignment of states with the centre to identify impact. Using establishment-level panel data on Indian manufacturing during 1998-2012, we find that road density raises TFP by 10%-34% at the industry-level. This translates to an average elasticity of 0.12. The impact of roads on TFP differs by firm's size, age and location. Smaller firms, incumbents and those in urban states see most pronounced gains. Furthermore, long-differenced results show that an increase in road density raises the growth rate of TFP. Placebo tests on our instrument and several robustness checks reinforce our findings. Thus, addressing the shortfall in transport infrastructure can unlock large economic dividend.

Keywords: Manufacturing; Total Factor Productivity; Transport Infrastructure; Roads; Political Economics; India

JEL codes: D24; L60; R11; R40; D72

*Corresponding author, Department of Economics, University of Reading, Whiteknights. Email: subham01@gmail.com

†Department of Economics, University of Reading

1 Introduction

Transportation plays a vital role in stimulating economic growth. But, in many emerging economies including India and China, inadequate transport infrastructure impedes manufacturing growth (Jedwab and Moradi 2016, Ghani et al. 2016*a*). For instance, the recent World Bank Enterprise Survey (WBES) reports that inadequate transport infrastructure is a major obstacle for one in every ten manufacturing firms in India (WB 2014).[1] To address the shortfall in transport infrastructure and unlock its growth potential, many developing countries have embarked on ambitious multi-billion dollar projects.[2] Yet, in spite of such big-budget spending, our understanding of the role of transportation in manufacturing growth especially in developing countries, is far from complete. In this article, we analyse the impact of India's *entire* road infrastructure on its manufacturing Total Factor Productivity (TFP).

To date, a majority of studies that focus on the economic impact of transportation in developing countries have adopted two broad approaches. The first approach studies grand transport projects such as the Golden Quadrilateral (GQ) project in India (see Ghani et al. 2016*a*,*b*, Datta 2012, etc.) or the National Trunk Highway System (NTHS) in China (see Faber 2014) that offer project-specific insight. The second approach considers the effect of a synthetic index of infrastructure that includes a range of infrastructure variables (see Mitra et al. 2002, 2012, etc.). Because the contribution of individual components in the latter approach is not separable, it is not well suited to learning how transportation *per se* affects manufacturing productivity.[3] Using establishment-level panel data on Indian manufacturing firms during 1998-2012, we examine the contribution of India's entire road

[1] Inadequate transport infrastructure is both a regional and a global constraint on manufacturing: about 21% of the firms surveyed in South Asia and around 19% of firms surveyed across all countries in the WBES during the years 2010-2017 indicate lack of transport infrastructure to be a significant obstacle. See `http://www.enterprisesurveys.org/`. In addition, it constrains infrastructure-dependent plants in benefitting the most from liberal industrial policies such as the de-licensing of the product space in Indian manufacturing (Gupta et al. 2008).

[2] For instance, India's 2017-18 budget announced a record spending of Rs.3.96 trillion ($59 billion) to build and modernize its transport infrastructure and provide 'renewed impetus' to manufacturing. This includes an allocation of Rs.649 billion for road building, a 12% rise in comparison to the previous budget (Bloomberg 2017). In contrast, expenditure allocation towards healthcare and education grew by 2.76% and 3.84% respectively, during the same period.

[3] While focusing on showcase transportation projects offer detailed project-specific insight, they may not be generalizable. If we consider that showcase projects are often carefully selected to maximize economic returns at low social/ political cost (e.g. projects that have low land acquisition costs or are located in politically favoured states), the findings from these studies may be closer to the upper bound instead of resembling the average effect. On the other hand, studies that use an aggregate index of infrastructure that includes roads, railways, electricity and a host of other infrastructure variables cannot easily disentangle the impact of its individual components and hence offer limited policy insight. Hulten et al. (2006) is an exception that studies the impact of national and state highways on Indian manufacturing productivity but does not control for endogeneity.

infrastructure on its manufacturing TFP. This allows us to draw generalized insights on the economic impact of roads within a developing country context. We focus on roads because it is the predominant mode of transportation accounting for 65% of freight movement and 80% of India's passenger traffic (NHAI cited in Ghani et al. 2016*a*).

Transport infrastructure can affect economic activity in several ways: First, a robust transport network lowers freight rates and improves travel times that integrate distant markets. The arterial network of roads and railways have played a fundamental role in increasing trade through improved market access. This relationship has been extensively documented (see, for example, Fernald 1999, Chandra and Thompson 2000, Michaels 2008, Duranton et al. 2014, Holl 2016, Redding and Turner 2015, among others).

Secondly, a reliable transport network affects a firm's logistics throughout the different stages of production. It lowers uncertainty at both the procurement and supply stages thereby reducing the need to hold on to large inventories (Shirley and Winston 2004, Li and Li 2013, Datta 2012). The cost-saving aspect is particularly relevant for developing countries where inventory levels are generally two to five times higher than in the United States, the developed country benchmark (Guasch and Joseph 2001). Gulyani (2001), for instance, examines the Indian automotive sector and finds that the 'total logistics cost' of inadequate transport infrastructure is very high. In a similar vein, Datta (2012) estimates that being on the GQ highway reduces inventory holdings by at least six days' worth of production. Thus. lower transport cost results in larger and spatially diversified markets that are serviced by manufacturers operating at low levels of inventory. Unsurprisingly, these gains get reflected in manufacturing productivity.

The constraints posed by inadequate transportation infrastructure are analysed by Storeygard (2016) who use night-time luminosity as a proxy for the economic size of cities in Africa. Using global oil prices to provide exogenous variation in transport costs, they find that the economic activity of African cities that are further away from roads suffer more than cities that are nearer during high oil price years. In India, Ghani et al. (2016*a*) examine the economic impact of the Golden Quadrilateral (GQ) project and find that industries located nearer to highways produce more output and have higher allocative efficiency. In a similar study, they find that the GQ highway affects rural and urban areas differently. While the entry rates of firms are similar across both rural and urban areas, rural areas experience a larger increase in employment and output (Ghani et al. 2016*b*). Thus, the literature suggests that improving transport infrastructure in general, and roads in particular, is likely to increase manufacturing productivity. But, the gains might differ both within and across industries.

An important challenge in estimating the impact of roads on economic activity is that their placement is not random (Banerjee et al. 2012, Shatz et al. 2011, etc.). For example, roads might be more prevalent in states that trade more i.e. when roads are built to support trade, or if they trade less, i.e. when roads are built to encourage trade (Duranton et al. 2014). Hence, state policies geared to attract specific industries might affect how much it invests on roads. Moreover, states with higher GDP are also likely to have larger resources to build roads. We overcome this inference problem by exploiting exogenous variation in the implementation of road projects due to political alignment of states with the central government. The premise is that, in a federal democracy, federal transfers favour states that are politically aligned with the centre over states that are not so aligned (see, for example, Sengupta (2011) for a theoretical framework and Arulampalam et al. (2009) for empirical evidence of partisan alignment boosting federal transfers in India). Politically motivated favouritism then increases the ability of aligned states to invest in visible transport infrastructure such as roads to improve their chances of victory in the next election. This exogenous variation in a state's road infrastructure, after controlling for railways and other state-specific fundamentals, allows us to maintain the exclusion restriction in identifying the effect of roads on TFP. Our study therefore provides a political-economic explanation for differences in road density across states, which have first order impact on manufacturing productivity.

Another concern is that omitted variables such as differences in state-specific characteristics or access to alternative modes of transportation might confound the relationship between roads and manufacturing productivity. For example, firms might face different macroeconomic conditions that vary with its location such as the available labour pool, the skill set of workers, access to railways etc. all of which might affect productivity. To address this second concern, we control for year-interacted state-specific characteristics – population, literacy, total main and marginal workers, total workers in agriculture and industry – and a state's railway density to absorb the effect of potential confounders and allow for different time trends according to these characteristics. We also include state-industry fixed effects to control for state-specific policies toward different industries that might systematically affect industrial productivity within states. In addition, general economic conditions during a year such as the level of inflation in the country or the exchange rate situation might affect production orders or even the state's capacity to build roads. We include year fixed effects to absorb such year-specific shocks. Once we include the full set of controls, we are reasonably confident in pinning down the impact of roads on productivity.

To preview our results, we find that a marginal increase in road density leads to a 10%-34% rise in manufacturing TFP at the industry level. The impact of roads on TFP differs by a firm's age, size and location. Smaller firms, incumbents and those located in urban states see most pronounced gains in productivity. To get a longer-term picture, we analyse the impact of changes in road density on changes in industry-level productivity between 1998 and 2012. We find that the results from these long-differenced regressions are qualitatively similar to those from state-industry panel regressions. Both show a strong positive relationship between an increase in road density and a rise in manufacturing TFP. Thus, in addition to scale-effects, roads affect the growth rate of productivity. We conduct elasticity estimates to get a more intuitive understanding of the results and find that the average elasticity is about 0.12. Placebo tests on our instrument and a battery of robustness checks ensure the reliability of our findings.

We organize the article as follows: In section 2, we introduce the dataset. In section 3, we specify a simple theoretical model that outlines the relationship between roads and TFP, discuss how we calculate TFP and present our identification strategy. In section 4, we report regression results on the impact of roads on TFP that accounts for endogeneity. We present results from robustness checks in section 5. Section 6 concludes.

2 Data

We combine detailed data on organised manufacturing in India with state-level data on transportation infrastructure and socio-demographic characteristics to create a state-industry panel running from 1998 until 2012. We discuss this below.

2.1 Manufacturing Data

We use establishment-level panel data on manufacturing activity during 1998-2012 to estimate TFP. We obtained this data from the Annual Survey of Industries (ASI), a pan-India survey of organized manufacturing establishments administered by the Central Statistical Organization (CSO), Government of India. While ASI covers the entire country, we exclude Union Territories that do not hold legislative assembly elections from our sample. We will discuss this further in section 3.3. ASI provides exhaustive data on the book values of a firm's assets and liabilities, employment and labour, receipts and expenses along with several other economic variables for a financial year (e.g. the 1998 survey reports data for 1998-99), but we will only refer to the initial year for simplicity.

ASI covers organized manufacturing establishments registered under the Factories Act, 1948 that employ more than 10 workers if they use electricity in their manufacturing process or 20 workers if they do not. It follows a Circular Systematic sampling design that divides the sampling frame into two sectors: a census sector and a sampling sector. The census sector consists of establishments in the states of Manipur, Meghalaya, Nagaland, Tripura and Union Territory of Andaman and Nicobar Islands where all establishments are surveyed. In other states, the census sector includes establishments with more than 100 workers or if they file joint returns i.e. returns for multiple units within a state. The sampling sector, on the other hand, consists of establishments that are not included in the census sector. The sampling design considers the state, sector and 4-digit NIC codes in stratifying the sample. Because of the survey design, we use the sampling multipliers provided by ASI throughout the article and cluster the standard errors of the regression coefficients at the state-industry level (see, Abadie et al. (2017) on clustering as a design problem). Thus, our results are representative of the entire population of manufacturing firms in the organized sector in India.

With the data at hand we construct an unbalanced panel of establishment-level economic activity starting from 1998, when ASI first released firm-level panel data, until 2012. While the data contains firm identifiers that allow us to observe individual firms over time, ASI, in compliance with the Collection of Statistics Act, 2008, no longer disseminates district identifiers after 2007. Hence, in this article, we conduct our analysis at the state-industry level. Analysis at the state-industry level is appropriate because it helps us to factor in state-specific endowments such as raw materials, labour pool, etc. that affect the distribution of industries within a state. Another reason was the need to cover a longer time span (i.e. to have larger Ts). This not only increases the precision of our estimates but also allows us to observe if the impact persists over time. In contrast, if we reverse engineered district identifiers, as in Martin et al. (2017), we could only go up to the year 2007 (after which ASI discontinued releasing district codes) whereas, with our approach using annual state-industry panel, we are able to extend the analysis further up to the year 2012. Besides, another advantage of this approach is that it is forward compatible as new data is released in future.[4]

[4]While we are aware of the matching method used in Martin et al. (2017) to recover the district identifiers of firms by matching economic variables in the panel data with several rounds of cross-sectional data that include district identifiers, we decided to go ahead with a state-industry panel for our analysis because of the gains from using the approach, as already discussed.

2.2 Transport Infrastructure

As indicated earlier, our main covariate of interest is road density. We focus on roads since it is responsible for about 65% of freight and 80% of all passenger traffic in India (NHAI cited in Ghani et al. (2016*a*)). But, using the total length of roads in a state might be problematic because states vary widely in their geographical size. To overcome this, we focus on road density, the total length of roads in a state divided by a state's land area, to obtain a normalised metric of access to roads in each state. Moreover, focusing on a physical measure rather than the money value to capture market access is advantageous for two main reasons. First, it ensures that our measure is free from distortions due to price volatility. Secondly, a physical measure is less vulnerable to misreporting, which is not an innocuous problem in weak institutional settings that typically lack in transparency and accountability. We obtained data on the total length of roads by state-year from the Ministry of Road Transport and Highways from 1998 until 2012. Summary statistics of our sample indicate that the average road density across states was 4.86 km/sq.km in 2012. Delhi had the highest road density of 21.6 km/sq.km whereas, Jammu and Kashmir had the lowest road density of 0.20 km/sq.km.

Railways present the main alternative to road transportation in India (see Donaldson 2018, on the impact of railways on trade in colonial India). Hence, to clearly identify the effect of roads on manufacturing productivity we need to include a state's railway density as a control variable in our model. To operationalise this, we gathered data on the length of railways (in route km) for every state during the study period from the Ministry of Railways and divided it by a state's land area to obtain normalised values of railway density.

2.3 Socio-Demographic Characteristics

We also include several important socio-demographic variables as controls in our model to reduce omitted variable bias. These include population, literacy rate, main and marginal workers as well as data on total agricultural and industrial workers. Data for these variables are obtained from the 2001 census. Thus, we obtain estimates that pin down the impact of roads on TFP after conditioning on a range of control variables.

3 Empirical Strategy

3.1 Model Specification

To analyse the effect of roads on manufacturing productivity, we first estimate firm-level productivity as a function of local determinants (see Holl 2016, Martin et al. 2011) using a Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^{\beta_1} L_{it}^{\beta_2} \tag{1}$$

where Y_{it} is the value added output of firm i at time t. L_{it} denotes labour employed in production and, K_{it} is the capital stock. We note that firm i belongs to industry j and is located in state s but we suppress the subscripts for simplicity in the above equation. We then compute $A_{jst} = 1/n \sum_i A_{ijst}$ as the average TFP of industry j in state s at time t by taking the average across firms indexed $i = 1 \dots n$ in industry j within state s at time t. A_{jst} depends on a vector of state specific characteristics, Z_{st}, and access to roads, $ROAD_{st}$. We adopt this approach because ASI data only allows us to identify the state where a firm is located, as already mentioned. Since our goal is to examine the effect of the total road network on industry-level TFP, conducting the analysis at the state-industry level seems to be most appropriate. We model productivity as[5]:

$$A_{jst} = exp[\gamma(ROAD_{st} + Z_{st})] \tag{2}$$

We take log on both sides of eq.(1) to get:

$$y_{it} = \beta_1 k_{it} + \beta_2 l_{it} + \alpha_{jt} \tag{3}$$

and for industry-level log TFP values from eq.(2) to get,

$$\alpha_{jst} = \gamma ROAD_{st} + \eta_{jst} \tag{4}$$

where lower case letters now denote the log of the respective variables. The term η_{jst} might further include state characteristics, state-industry fixed effects or year effects that affect industry-level productivity along with a random noise term.

We estimate the impact of roads in two steps. First, we estimate firm-level TFP based on eq.(3) and obtain average industry-level TFP for different states across the years. In the

[5] We note that our results hold if we assume $A_{jst} = \gamma(ROAD_{st})Z_{st}$.

second step, we use this productivity measure as the dependent variable to estimate the effect of roads in eq.(4).

3.2 Measuring TFP

In this study, we estimate TFP using the value-added definition of production. The results are also robust to the gross-output definition. We focus on multifactor productivity because differences in TFP reflect shifts in the production isoquants with higher-TFP producers on a higher isoquant than lower-TFP producers (Syverson 2011). We estimate TFP as the residual of a Cobb-Douglas production function as:

$$\alpha_{it} = y_{it} - \hat{\beta}_1 k_{it} + \hat{\beta}_2 l_{it} \tag{5}$$

where $\hat{\beta}_1$ and $\hat{\beta}_2$ are the input coefficients of capital and labour respectively, estimated at the 3-digit NIC industry classification using the Ackerberg et al. (2015) method (henceforth, ACF). We use the ACF method because: first, it accounts for the endogeneity of input choice in estimating productivity. The endogeneity arises because firms can observe their productivity before choosing inputs which means that inputs are likely to be correlated with productivity. Secondly, the ACF method overcomes the functional dependence problem that affects identification of the labour coefficient present in Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP) methods. Thus, while OP and LP invert investment (OP) and intermediate input (LP) demand functions that are unconditional on labour input, ACF inverts investment or intermediate input that are conditional on labour input to overcome functional dependence and correctly identifies the labour coefficient in the first stage (see Ackerberg et al. (2015) for a detailed discussion on the methodology, and Matthias et al. (2016) for a recent application of this method). We note that our results using the ACF method are consistent with both the LP and the Wooldridge (2009) methods of estimating TFP.

To estimate TFP, we first deflate sales, labour, raw material, energy, and capital to arrive at their respective real values. Since the book value of capital is measured at historic costs, we use the perpetual inventory method (PIM) that accounts for differences in the vintage of capital stock. We then estimate firm-level TFP by 3-digit NIC industry classification (see Appendix A for the variables used in estimating TFP).

To do this, we estimate firm-level TFP using 264 thousand observations covering 54 thousand establishments at 3-digit NIC level. Next, we normalise log TFP by dividing

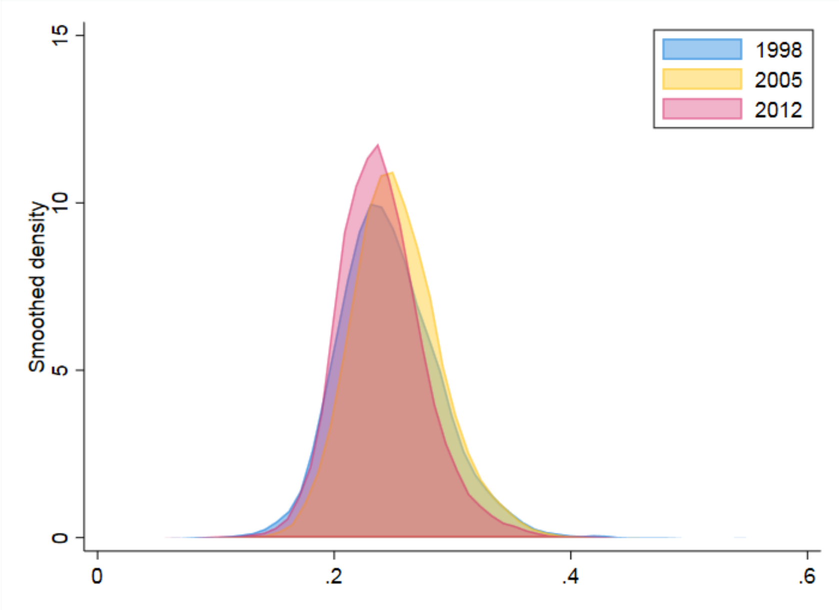


Figure 1: TFP in 1998, 2005 and 2012.
Figure plots the density of log value-added TFP in 1998, 2005 and 2012. TFP is normalised by dividing it by employment-weighted average productivity for an industry-year.

it by the employment-weighted average productivity (in logs) for an industry-year.[6] In Figure 1, we plot the log normalised value added TFP (TFP(VA)) for three different years – 1998, 2005 and 2012. We observe that productivity increased consistently between the years and showed signs of increasing concentration of high-TFP firms in the industry.

We then calculate industry-level averages of TFP(VA). Table 1 presents summary statistics for TFP(VA) (see Table 13 in Appendix B for summary statistics relating to the log normalised gross-output TFP values (TFP(GO)). We observe that TFP(VA) is about 0.16 at the median for the entire sample. It is higher for small firms (0.20) than large ones (0.13). The 90:10 percentile TFP(VA) ratio – a measure of dispersion of productivity – is about 3.7. This illustrates significant heterogeneity in TFP(VA). The firms at the 90th percentile are about four times more productive than firms at the 10th percentile.

We now turn to discuss our identification strategy.

3.3 Identification

Our identification strategy relies on three key aspects of India's political economy landscape: (a) fiscal dependence of states on federal funding for road building projects; (b) politically motivated tactical spending on 'visible' infrastructure projects by a state's ruling party;

[6]This approach is followed in Ghani et al. (2016*a*). We adopt this approach for comparability and note that our results hold even without normalising the TFP values.

Table 1: TFP in Indian Manufacturing (Value Added), 1998-2012

Groups	Obs.	Mean	p10	p25	p50	p75	p90
All	6714	0.162	0.070	0.097	0.153	0.212	0.262
By Age:							
Young	5071	0.169	0.068	0.087	0.147	0.220	0.299
Incumbent	6595	0.163	0.069	0.095	0.155	0.216	0.267
By Size:							
Small	6333	0.202	0.069	0.101	0.192	0.277	0.345
Large	6173	0.128	0.069	0.085	0.116	0.155	0.201
By Location:							
Rural	3203	0.138	0.065	0.080	0.117	0.177	0.239
Urban	3511	0.186	0.086	0.129	0.187	0.230	0.274
TFP_{2012}/ TFP_{1998}	298	1.334	0.759	0.953	1.204	1.589	2.057

Notes: Table shows the dispersion of log value-added Total Factor Productivity (TFP) in Indian manufacturing during 1998-2012 estimated using the ACF method (Ackerberg et al. 2015). The firm-level log TFP values are normalised by dividing it by employment-weighted average productivity (in logs) for an industry-year. Row 'All' in the table corresponds to all observations in our sample. 'Young' includes industry-level estimates of only firms aged $\leq$ 4 years whereas, 'Incumbents' include those that are aged >4 years. 'Small' includes firms in the industry with fixed-assets lower than the median fixed assets in the industry, while 'Large' includes those with fixed assets greater than the industry median value. 'Rural' shows TFP for industries located in rural states where a rural state is defined to have urbanisation rate lower than the median value across all states in our sample. 'Urban' shows TFP for industries located in states where the urbanisation rate exceeds the median value. TFP_{2012}/ TFP_{1998} is the ratio of log TFP values in 2012 and 1998.

and, (c) the potential for manipulation of the size and timing of federal transfers to favour politically aligned states. The centre-state revenue dependence together with the potential for electoral gains from investing in visible infrastructure projects create strong incentives for the central government to favour federal transfers to politically aligned states. Owing to these institutional features, we use the political alignment of states with the centre as an instrument for road density.

In India, roads are predominantly publicly provided. While the central government funds the majority of road building works, the respective state governments are mainly responsible for its implementation. The size and timing of federal transfers allows aligned states to strategically increase their expenditure on public goods such as roads with the intention of strengthening the electoral position of incumbent parties. This is especially important when seats are tightly contested (Arulampalam et al. 2009, Johansson 2003, Baskaran et al. 2015, Bracco et al. 2015).[7] In decentralized democracies, partisan alignment

[7] Moreover, in parliamentary systems with single member districts, the chances of winning for the party is closely aligned with the performance of co-partisans which increases public provision that wins votes while reducing the incentive for corruption. As Bohlken (2016) points out, the role of partisan alignment is important because it is able to reconcile incentives for private rents that ministers keep to themselves (as argued in Lehne et al. (2018)) and providing public goods and controlling wider corruption in infrastructure projects.

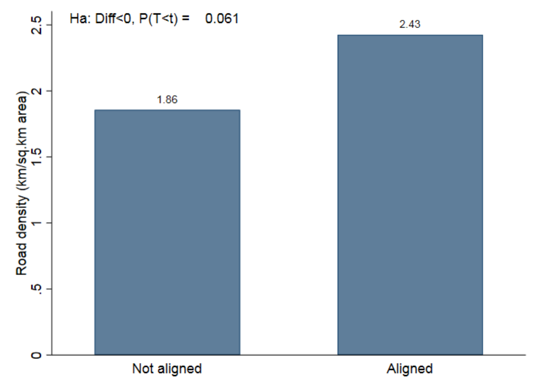
of lower-level jurisdictions with the centre leads to 'tactical', rather than 'programmatic' federal transfers that systematically affects the provision of local public goods (Solé-Ollé 2013, Solé-Ollé and Sorribas-Navarro 2008, Sengupta 2011, Khemani 2003).[8]

The hierarchical dependence relationship is likely to affect implementation of a range of development projects. But, 'visible' infrastructure projects are especially vulnerable to tactical/ targeted spending since the ruling party can credibly claim credit for the roads it builds as they are easily observed by the electorate (Mani and Mukand 2007). Wilkinson (2006), for example, documents that when Indian states hold competitive elections politicians announce several infrastructural projects with the aim of building political support and to raise money for campaign finance. Whether these promises are delivered and the projects completed depend on the ruling coalition. A different political party in power might shift the goal away from completing on-going projects onto new ones (Williams 2017). Hence, political economic factors olay a key role in determining when and where road projects are implemented.

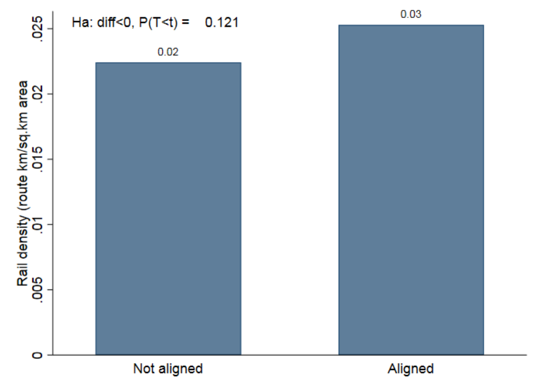
We exploit the exogenous variation in partisan alignment between the state's ruling party and the centre over election cycles that are staggered across states as an instrument for road density. To operationalize this, we define a ruling party as the political party winning the maximum number of seats in a state for each legislative assembly election held between 1998 and 2012. We then construct a binary variable, *Aligned*, that takes a value of one if the state's ruling party is aligned with the centre (or the central coalition) and a value of zero if it is not aligned. To do this, we obtain data on state assembly election results during the period of our study from the Election Commission of India. In the context of coalition governments at the centre, we carefully map the various state parties with the coalition to account for any re-configurations (including parties dropping out during the life of the coalition government[9]).During our study period, the National Democratic Alliance (NDA) headed by the Bharatiya Janata Party (BJP) was in power from 1998 until 2003, while the Congress Party-led United Progressive Alliance (UPA) was in power from 2004 to 2014. Finally, we drop Union Territories from our analysis because they do not hold legislative elections.

[8]In India, the Finance Commission is responsible for determining the extent of equalization transfers from the centre to the states to neutralize vertical imbalance through a formulaic approach. But as Khemani (2007) notes 'formula-based transfers ... have limited success in curbing political influence ... the formula itself is determined by a political process or is not binding and leaves room for political discretion ...'.

[9]For instance, the All India Anna Dravida Munnetra Kazhagam (AIADMK), a leading regional party in Tamil Nadu was part of the NDA alliance in parliamentary elections in 1998 but withdrew its support a year later leading to the BJP government's collapse and an early re-election in 1999 where AIADMK realigned with the Congress Party. We update the ruling coalition to account for any changes at the centre.



(a) Road density



(b) Railway density

Figure 2: Political Alignment and Transport Density: Road vs. Railway.
Panel (a) plots the average road density (in km per sq.km area) across Indian states by alignment status during 1998-2012 whereas, panel (b) plots the average railway density (in route km per sq.km area) in a similar way. We observe that road density is higher when a state's ruling party is aligned with the centre in comparison to when it is not aligned. The difference is statistically significant with a p-value of 0.06 based on a one-sided t-test. For railway density, the difference is not statistically significant.

To what extent does political alignment affect road density? In Figure 2 panel (a), we compare the average road density per sq. km land area of a state – our measure of access to roads – between aligned and non-aligned states. It shows that during 1998-2012, the average road density in aligned states was 2.43 km/sqkm area whereas, for non-aligned states the average road density was lower at 1.86 km/sqkm area. This difference is statistically significant with a p-value of 0.06 based on a one-sided t-test.

We then compare the density of roads with the density of railways between Aligned and Non-Aligned states. *Ex-ante* we would expect a state's road density to differ along partisan alignment status but less so for railways. This is because the extension of railways unlike roads, depends heavily on the historical networks that have been already laid out. In contrast, roads are easy to provide than railways, which make them vulnerable to manipulation. Figure 2 panel (b) plots the relationship between railway density and political alignment. We find that, as expected, while the difference between aligned and non-aligned states is statistically significant for road density, it is not so for railways. The finding that political alignment affects road density but not railway density increases our confidence in the choice of our instrument.

We now proceed to discuss the regression results.

4 Results

In this section, we first present regression results that relate to the impact of roads on industry-level TFP. We then examine the heterogeneity in the impact of roads by firm's size, age and location. Finally, we present elasticity estimates that are more intuitive to understand.

4.1 Industry-level Effects

To conduct our analysis, we construct a weakly balanced panel of industry-level TFP(VA) by state-year during 1998-2012 by averaging firm-level TFP. This yields 6,714 observations. In all regressions, we use the multipliers supplied by ASI and weight the observations by an industry's employment in the initial study period. Hence, our results are representative of the entire population of organized manufacturing firms in India.

We begin by estimating a semi-log panel fixed effects model where we regress TFP(VA) on road density and control for state-industry and year fixed effects. Thus, we estimate an equation of the form:

$$TFP_{jst} = \gamma(ROAD_{st}) + \theta_{js} + \phi_t + \mu_{jst} \tag{6}$$

where TFP_{jst} is the average industry-level TFP(VA) for industry j in state s at time t that correspond to α_{jst} in eq.(4). $ROAD_{st}$ is the road density for a state in a given year and γ its marginal effect – the main coefficient of interest. θ_{js} denotes state-industry fixed effects; ϕ_t denotes year effects and μ_{jst} is an error term clustered at the state-industry level. We cluster our errors at the state-industry level because we want to account for possible correlations within a state-industry pair over time that arise due to state-specific industrial policies and to account for ASI's sampling design.

Table 2 presents results from estimating the marginal effect of road density, γ. The panel fixed effects results (Table 2 columns 1 and 4) indicate that road density is positive but not statistically significant. However, as already discussed, the fixed effects estimates are likely to be biased due to the non-randomness of road placement. In fact, the coefficients are likely to be downward biased since the major thrust of the government has been to extend connectivity to remote habitations where productivity is much lower (see 1 for differences in rural-urban manufacturing TFP).

Columns 2 and 3 in Table 2 present results from IV regressions, where we instrument road density with the exogenous variation in the political alignment of a state's ruling party

Table 2: Effect of Road Density on TFP: Main Results

Variables	Valued added			Gross Output		
	FE (1)	IV-I (2)	IV-II (3)	FE (4)	IV-I (5)	IV-II (6)
Road density	0.003	0.066***	0.100***	0.003	0.067***	0.103***
	(0.00)	(0.02)	(0.03)	(0.00)	(0.02)	(0.03)
Rail density	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Year x State variables	No	No	Yes	No	No	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6714	6651	6651	6714	6651	6651
RMSE	0.043	0.050	0.054	0.042	0.049	0.054
F		41.99	35.36		41.99	35.36

Notes: In cols.(1)-(3) the dependent variable is log value-added TFP while in cols.(4)-(6) it is log gross-output TFP, both estimated using ACF method. FE shows results from a panel fixed effects model; IV-I from a panel 2SLS model that controls for state-industry and year fixed effects whereas, IV-II shows results from a panel 2SLS model that additionally controls for time-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – and a state's railway density. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

with the centre across several election cycles that are staggered across states. Further, we include additional controls to maintain the exclusion restriction and avoid omitted variables bias. We estimate the full model with additional controls as:

$$TFP_{jst} = \gamma(Road_{st}) + \delta_t X_t + \zeta(Rail_{st}) + \theta_{js} + \phi_t + \varepsilon_{jst} \tag{7}$$

where X_{st}, is a vector of year interacted state-specific characteristics in logs with δ_t its corresponding coefficient vector that allow for different time trends according to these characteristics. We also control for a state's railway density, $Rail_{st}$ with coefficient ζ to purge the effect of railways on TFP and include state-industry and year fixed effects like in eq.(6). ε_{jst} denotes an idiosyncratic shock term.

Columns 2 and 5 in Table 2 present results from a panel Two Stage Least Squares (2SLS) model where we regress TFP(VA) and TFP(GO) respectively, on road density instrumented by *Aligned*, conditional on state-industry and year fixed effects. Before we go any further, we need to consider whether our instrument is appropriate. Column 1 in Table 3 presents the first stage results where we regress road density on $Aligned_{st}$, conditioning on all controls, except rail density, to maintain the exclusion restriction. We include rain density in column 2. It shows that road density is 0.12-0.17 points higher when the state's ruling party is aligned with the centre than when it is not aligned. These

Table 3: First Stage: Effect of Alignment on Road Density

	IV-I (1)	IV-II (2)
Aligned	0.173***	0.115***
	(0.03)	(0.02)
Rail density	No	Yes
Year FE	Yes	Yes
Year x State variables	No	Yes
State-Industry FE	Yes	Yes
Observations	6651	6651
RMSE	0.357	0.324

Notes: Table shows first stage estimates from regressing Road density on *Aligned* conditional on controls. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

results are statistically significant at the 1% level of significance demonstrating that our instrument is suitable for the task at hand.

Returning to the IV estimates in Table 2, we find that our estimates for both the VA and GO measures are very similar. But, in what follows we will concentrate on the VA estimates. Column 3 provides estimates with the full set of controls where we find that road density has a highly significant impact on TFP. Our results indicate that an additional km of road per sq.km area leads to about a 10 percentage point increase in TFP. We cluster the standard errors at the state-industry level for reasons already mentioned.

One concern is whether political alignment indeed affects road density or if it is merely an artefact of the data. To dispel this concern we conduct a placebo test motivated by Chetty et al. (2009) and Martin et al. (2017), where we randomly assign alignment status to a state during a year and use the constructed variable as our placebo instrument for road density. *Ex-ante* we would expect placebos to yield insignificant estimates. We conduct 100 placebo runs in total where each draw is from a Bernoulli distribution that preserves a state's historical alignment probability between 1998 and 2012. Table 4 presents results from one such draw where we find that the placebo instrument is not informative i.e. it has no significant effect on road density and neither does it affect TFP even after conditioning on the full set of controls.

Table 5 summarises the results from running the placebo test 100 times for TFP(VA) and TFP(GO), respectively. We get an insignificant result about 93 times for TFP(VA) and 92 times for TFP(GO) respectively, which largely reject the placebo instruments – the randomly assigned alignment status – in favour of true alignment status. Only 5

Table 4: Placebo Instrument Test

	TFP(VA) (1)	TFP(GO) (2)	First Stage (3)
Road density	0.274	0.305	
	(0.46)	(0.50)	
Aligned (Placebo)			0.005
			(0.01)
Rail density	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes
Observations	6698	6698	6698
RMSE	0.098	0.107	0.326
F	0.456	0.456	

Notes: Results from a single placebo test where we instrument a state's road density with a randomly assigned alignment status based on a Bernoulli distribution that preserves a state's alignment probability between 1998-2012. Placebos were assigned 100 times. This table shows one example of a placebo run. We include the full set of controls. See Table 2 for the full list of conditioning variables. RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

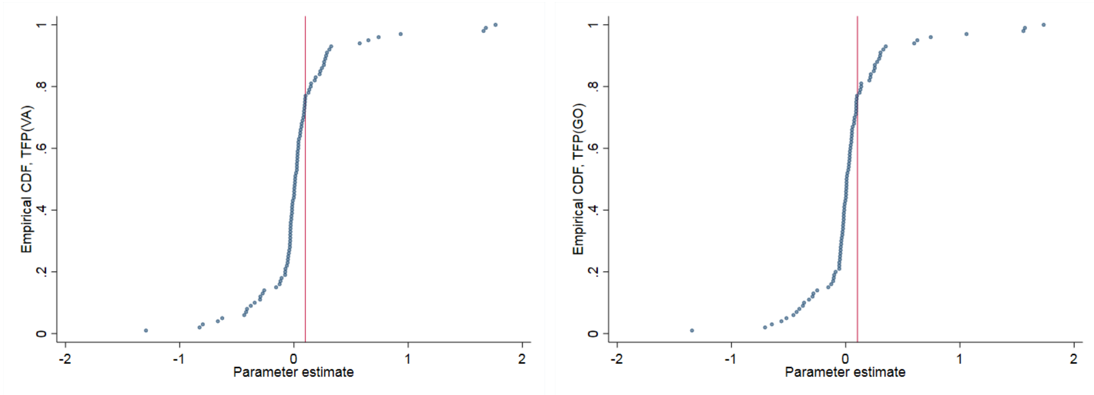
times out of 100 runs do we get a significant positive coefficient for the effect of road density on TFP(VA) at the 5% level of significance (6 times for TFP(GO)) whereas, the coefficient has incorrect sign and is negative in the remaining 2 cases for both TFP(VA) and TFP(GO).

Table 5: Summary Statistics of Placebo Tests

Variable	Above 0 (1)	Below 0 (2)	Insignificant (3)
TFP(VA)	5	2	93
TFP(GO)	6	2	92

Notes: In cols.(1) and (2), we show the number of runs when each outcome of interest was above or below 0 and significant at the 5% level, respectively, whereas in col.(3), we show the number of runs when the result did not turn out to be significant at the 5% level. Placebos were assigned 100 times.

In Figure 3, we plot the empirical cumulative distribution function (CDF) for the 100 placebo tests for TFP(VA) and TFP(GO) in panels (a) and (b), respectively. The true coefficients as shown by the vertical lines are towards the right of the CDFs in each case implying that the placebo test was successful.



(a) Empirical CDF, TFP(VA) (b) Empirical CDF, TFP(GO)

Figure 3: Figure shows empirical CDF distribution of 100 placebo runs and their true coefficients for value-added and gross output TFP (in logs), respectively. The vertical lines show the coefficients on true alignment status.

In the next section, we discuss the extent to which the effect of roads on TFP differ by age, size and location of firms within an industry.

4.2 Impact Heterogeneity

Our results so far clearly indicate that higher road density translates into positive gains in manufacturing TFP. What is not clear, however, is whether the impact differ by firm's size, age or location, a question that might be of particular relevance to policymakers. In this section, we present results that reveal the heterogeneity in the impact of roads along these axes.

4.2.1 Young vs Incumbent Firms

We begin by examining whether roads differentially affect young firms relative to incumbents. We classify young firms as those established in the last four years whereas, firms above this threshold are the incumbents. We then aggregate TFP(VA) by a firm's age at the state-industry level for every year in the data following this classification. The number of observations is smaller than the full sample as some industries do not have firms in a given year that belongs to the young category or the incumbent category, as the case maybe.

Table 6 presents results for young and incumbent firms in the industry in Panels A and B respectively. We focus on columns 3 and 6 that correspond to the full-specification model for TFP(VA) and TFP(GO), respectively. The results indicate that a marginal increase in

Table 6: Effect of Roads on TFP: Young vs. Incumbent

Variables	Valued added			Gross Output		
	FE	IV-I	IV-II	FE	IV-I	IV-II
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Young Firms in Industry			
Road density	0.007	0.049*	0.078*	0.008	0.051*	0.079*
	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)	(0.04)
Observations	5071	5023	5023	5071	5023	5023
RMSE	0.076	0.081	0.081	0.076	0.081	0.081
F		32.68	25.39		32.68	25.39
			Panel B: Incumbent Firms in Industry			
Road density	0.003	0.066***	0.104***	0.003	0.067***	0.107***
	(0.00)	(0.02)	(0.03)	(0.00)	(0.02)	(0.03)
Observations	6595	6532	6532	6595	6532	6532
RMSE	0.045	0.052	0.056	0.044	0.051	0.056
F		39.67	32.40		39.67	32.40
Rail density	No	No	Yes	No	No	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Year x State variables	No	No	Yes	No	No	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: In cols.(1)-(3) the dependent variable is log value-added TFP while in cols.(4)-(6) it is log gross-output TFP both estimated using ACF method. FE shows results from a panel fixed effects model; IV-I from a panel 2SLS model that controls for state-industry and year fixed effects whereas, IV-II shows results from a panel 2SLS model that additionally controls for time-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – and a state's railway density. Panel A shows results that correspond to young firms aged$\leq$ 4 years in the industry whereas, Panel B shows results for incumbent firms in the industry aged $>$ 4 years. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

road density has a significant positive impact on TFP for both young and incumbent firms. However, the magnitude of impact is larger for incumbents. A marginal increase in road density results in a 10 percentage point increase in TFP(VA) for incumbents (Column 3, Panel B in Table 6) compared to an 8 percentage point increase for younger entrants (Column 3, Panel A in Table 6). This is hardly surprising because younger entrants have already factored in location decisions at the entry-stage. Thus, additional productivity gains from higher road density for younger firms, once they have established, are relatively small. In comparison, incumbents can now restructure their operation in response to higher road density, which gets reflected in their productivity.

Table 7: Effect of Roads on TFP: Small vs. Large

Variables	Valued added			Gross Output		
	FE	IV-I	IV-II	FE	IV-I	IV-II
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Small Firms in Industry					
Road density	0.005	0.080***	0.122***	0.006	0.082***	0.127***
	(0.00)	(0.02)	(0.03)	(0.00)	(0.02)	(0.03)
Observations	6333	6268	6268	6333	6268	6268
RMSE	0.063	0.071	0.075	0.061	0.069	0.074
F		41.92	33.53		41.92	33.53
	Panel B: Large Firms in Industry					
Road density	0.001	0.051***	0.081***	0.001	0.050***	0.082***
	(0.00)	(0.02)	(0.03)	(0.00)	(0.02)	(0.03)
Observations	6173	6111	6111	6173	6111	6111
RMSE	0.041	0.046	0.049	0.041	0.046	0.049
F		37.06	30.05		37.06	30.05
Rail density	No	No	Yes	No	No	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Year x State variables	No	No	Yes	No	No	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: In cols.(1)-(3) the dependent variable is log value-added TFP while in cols.(4)-(6) it is log gross-output TFP both estimated using ACF method. FE shows results from a panel fixed effects model; IV-I from a panel 2SLS model that controls for state-industry and year fixed effects whereas, IV-II shows results from a panel 2SLS model that additionally controls for time-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – and a state's railway density. Panel A shows results that correspond to small firms in the industry with fixed assets lower than the industry median value of fixed assets whereas, Panel B shows results for large firms in the industry with fixed assets that exceed the industry median value of fixed assets. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2.2 Small vs Large Firms

To consider whether the impact of roads on small firms is different in comparison to large ones, we classify firms as small if the real value of their fixed assets (land, buildings, machinery etc.) is lower than the industry's median value and we consider them as large if their fixed assets exceed the median value. We then aggregate the TFP(VA) by firm size at the state-industry level for every year in the data resulting in about 6,300 observations that relate to small firms and 6,100 observations for large firms.

Table 7 presents coefficient estimates of road density on TFP for small and large firms in Panels A and B, respectively. We find that, at the margin, higher road density benefits small firms more than they do large ones. Increasing road density raises TFP(VA) by about 12 percentage points (Column 3, Panel A in Table 7) for small firms whereas, it increases TFP(VA) by 8 percentage points (Column 3, Panel B in Table 7) for large firms. That productivity gains are higher for small firms might be due to the relative flexibility of small-sized firms with less fixed assets. In contrast, large firms face rigidities in restructuring production that constrain productivity gains from higher road density.

That smaller firms see larger productivity gains might be due to the flexibility that small size brings. Smaller firms can restructure operations quickly in response to better connectivity whereas, larger firms face rigidities. This translates into higher productivity gains for smaller firms relative to larger firms.

4.2.3 Rural vs Urban States

In this section, we examine if a denser road network affects firms in rural states differently than firms in urban states. We define a state as rural if its urban population is lower than the median value of urban population for all states in our sample.

In Table 8, we examine how roads impact TFP(VA) of firms in rural states (Panel A) and compare it with firms in urban states (Panel B). Further, we disaggreagte the impact of firm's location by its size and age (see Table 14 in Appendix B for the corresponding impact on TFP(GO)). We condition the estimates on the full set of controls in each column. We find that higher road density increases TFP(VA) in rural states by about 13 percentage points (Panel A, Column 1 in Table 8) whereas, the increase is about 22 percentage points for urban states (Panel B, Column 1 in Table 8). This difference might be due to the presence of complementarities in production in urban states that allow industries to capitalise on the full benefits arising from higher road density.

Considering the impact of roads on TFP by a combination of size, age and location characteristics, we find that the largest impact of roads is on younger entrants in urban states, where TFP(VA) rises by more than 34 percentage points (Panel B, column 4) followed by small-sized firms in urban states where the TFP(VA) rises by 33 percentage points (Panel B, column 2). Thus, we find that the effect of roads on manufacturing TFP is considerably heterogeneous and they differ by age, size and location of firms.

In the next section, we turn to results from long-differenced regressions.

4.3 Long-Differenced Results

Thus far, we consider the impact of roads on manufacturing productivity using state-industry panel data during 1998-2012. However, building roads takes time and road

Table 8: Effect of Roads on TFP (Value added): Rural vs. Urban

Variables	All (1)	Small (2)	Large (3)	Young (4)	Incumbent (5)
	Panel A: Rural States				
Road density	0.129***	0.149***	0.111***	0.037	0.142***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Observations	3143	2886	3022	2683	3047
RMSE	0.053	0.069	0.048	0.072	0.057
F	26.04	25.04	18.67	21.82	22.55
	Panel B: Urban States				
Road density	0.219**	0.334***	0.118	0.344**	0.189**
	(0.09)	(0.12)	(0.07)	(0.17)	(0.08)
Observations	3508	3382	3089	2340	3485
RMSE	0.075	0.115	0.055	0.128	0.069
F	12.11	12.12	10.81	8.06	11.90
Rail density	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes

Notes: The columns indicate log value-added TFP values for all firms in the industry (in col.(1)), for small firms (in col.(2)), for large firms (in col.(3)), for young firms (in col.(4)) and for incumbent firms in the industry (in col.(5)). Panel A shows results for industries located in rural states where we classify a state to be rural if its urban population is less than the median urban population of all states in our sample. Panel B shows results for industries located in urban states i.e. states which exceed the median urban population. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). The coefficients on road density are from separate regressions of the respective dependent variables on road density, controlling for year-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – a state's railway density and state-industry and year fixed effects. RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

infrastructure might need to reach a tipping point before it begins to affect productivity. On the other hand, it is possible that the impact of roads is only transitory, spurring a short burst in productivity as roads are built/ upgraded, but decays over time as firms readjust their operations. To investigate this, we consider the long horizon impact of road density on industry-level TFP by long-differencing our dependent variable, TFP(VA) and TFP(GO), as well as our main covariate, road density, over a span of 14 years. We do this by simply subtracting the values of these variables in 1998 from their corresponding values in 2012. Differencing in this way takes out time invariant factors at the state-industry level. Because our dependent variables are in logs, we can interpret the long-differenced dependent variable as the growth rate in productivity. We then regress the long-differenced

Table 9: Effect of Roads on TFP: Long-Differenced Results

	ACF Method		LP Method		
	ΔTFP (VA) (1)	ΔTFP (GO) (2)	ΔTFP (VA) (3)	ΔTFP (GO) (4)	First Stage (5)
Δ Road density	0.078***	0.078***	0.077***	0.079***	
	(0.02)	(0.02)	(0.02)	(0.02)	
CAY$_{98\text{-}12}$ x Aligned$_{12}$					0.087***
					(0.01)
Industry FE	Yes	Yes	Yes	Yes	Yes
State variables	Yes	Yes	Yes	Yes	Yes
Observations	298	298	298	298	298
RMSE	0.086	0.084	0.084	0.085	0.525
F		62.5			

Notes: Cols.(1) and (2) relate to long-differenced log value-added and gross-output TFP values between 1998 and 2012 respectively, estimated using Ackerberg, Caves and Frazer (2015) method. Cols.(3) and (4) relate to long-differenced log value-added and gross-output TFP respectively, estimated using the Levinsohn and Petrin (2003) method for the same time period. Col.(5) shows first stage results. ΔRoad density is the change in road density between 1998 and 2012 where road density is defined as total state-wide road length (in km) divided by its land area (in sq. km). We instrument road density with the consecutive number of aligned year between 1998-2018 (CAY$_{98\text{-}12}$) interacted with $Aligned_{12}$. All regressions control for industry fixed effects and state characteristics – population, literacy, total main and marginal workers and total main workers in agriculture and industry – and difference in rail density between 1998 and 2012. RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

variables on the change in road density between 1998 and 2012, which we can write as:

$$\Delta TFP_{js} = \beta(\Delta ROAD_s) + \delta X_s + \theta_j + u_{js} \tag{8}$$

where the prefix Δ denotes that the variable is long-differenced and X_s denotes a vector of state-specific variables. In eq.(8), we examine the impact of a marginal increase in the change in road density between 1998 and 2012, $\Delta ROAD_s$, on the growth rate of productivity, ΔTFP_{js} . We instrument $\Delta ROAD_s$ by the consecutive number of years for which a state was aligned with the centre between 1998 and 2012 interacted with $Aligned_{12}$. We condition the estimates on the on the full set of controls X_s.

Table 9 shows results from these estimations. In columns 1 and 2, the dependent variables are respectively $\Delta TFP(VA)$ and $\Delta TFP(GO)$, calculated using the ACF method. After controlling for endogeneity, we find that a marginal increase in the change in road density leads to about an 8 percentage point increase in the change in productivity. Thus, we can conclude that, at the industry level, roads not only affect the scale of productivity (see Table 2), but also its growth (see Table 9).[10]

[10]Columns 3 and 4 in Table 9 show that the long-differneced effects of roads on TFP estimated using the LP method are similar to that obtained by employing the ACF method in columns 1 and 2.

Table 10: Elasticity of TFP (Value Added) with respect to Road density

Variables	All (1)	Rural (2)	Urban (3)	Young (4)	Incumbent (5)	Small (6)	Large (7)
log(Road density)	0.120***	0.118***	0.442***	0.102*	0.120***	0.151***	0.091***
	(0.03)	(0.04)	(0.16)	(0.05)	(0.03)	(0.04)	(0.02)
Rail density	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6651	3143	3508	5023	6532	6268	6111
RMSE	0.053	0.056	0.076	0.082	0.055	0.074	0.047
F	89.604	85.432	20.085	58.538	85.634	80.282	78.734

Notes: Dependent variable is log TFP(VA) for the sub-sample indicated in columns (1-7). The results are from a panel 2SLS model that regresses TFP(VA) on log(Road density) where the latter is instrumented by *Aligned*. All regressions control for time-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – a state's railway density and include year and state-industry fixed effects. RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.4 Elasticity of TFP

The above results show how road density affects manufacturing productivity either on an annual basis using panel data or in the longer term using long-differenced regressions. Another useful and perhaps more intuitive way to understand the results is to look at elasticity – the percentage change in TFP due to a one percent increase in road density. To obtain elasticity esimates we regress TFP(VA) (in logs) on log (Road density) where we instrument the latter with *Aligned* and include the full list of controls. Table 10 presents elasticity estimates. It also shows how elasticity differs by firm's size, age and location. We find that overall, a one percent increase in road density leads to a 12 percentage point increase in TFP(VA). Mirroring earlier results from our semi-log model, the elasticity estimates show that the productivity of small firms, incumbents and those in urban states are more elastic to road density in relation to their respective counterparts. The productivity of large firms are the least elastic to road density (9.1%) whereas, firms in urban states are the most elastic (44%).

In the next section, we present results from robustness checks.

Table 11: Effect of High Road Density on TFP (VA and GO): FE and IV

Variables	Valued added			Gross Output		
	FE (1)	IV-I (2)	IV-II (3)	FE (4)	IV-I (5)	IV-II (6)
High Road density	0.003	0.442*	0.360**	0.003	0.449*	0.370**
	(0.00)	(0.23)	(0.16)	(0.00)	(0.23)	(0.16)
Rail density	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Year x State variables	No	No	Yes	No	No	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6714	6651	6651	6714	6651	6651
RMSE	0.043	0.120	0.100	0.042	0.122	0.102
F		5.76	8.73		5.76	8.73

Notes: In cols.(1)-(3) the dependent variable is log value-added TFP while in cols.(4)-(6) it is log gross-output TFP both estimated using ACF method. FE shows results from a panel fixed effects model; IV-I from a panel 2SLS model that controls for state-industry and year fixed effects whereas, IV-II shows results from a panel 2SLS model that additionally controls for time-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – and a state's railway density. High Road density is a dummy variable that takes a value of 1 if a state's road density exceeds two-third of the road density distribution in a state-year across all states in the sample and is zero otherwise. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5 Robustness Checks

We conduct several robustness checks to validate our results. We briefly discuss them below.

Alternative measure of road density: Here, we capture road density by a a binary variable, *high road density*, that takes a value of 1 if the road density for a state-year exceeds two-thirds of the distribution of road density across all states in the sample, and is zero otherwise. Relaxing the functional form specification in this way allows us to capture discontinuities in the effect. The results are presented in Table 11 where we consider the effect of high road density on TFP(VA) and TFP(GO) in columns 3 and 6 respectively, and where high road density is instrumented by *Aligned*. The first stage results show that our instrument is appropriate (see Table 15 in the Appendix B). After controlling for endogeneity and including the full set of controls, we find that high road density raises log TFP by about 36%-37% points relative to the base group, i.e. states with low road density. In comparison to previous results that use a continuous measure of road density, we observe that the magnitude of impact when using a binary measure is relatively large. The discontinuous jump in productivity from being in high road density states might be

Table 12: Effect of Roads on TFP: Robustness Checks

Variables	LP(VA) (1)	Winsorised (2)	No Split States (3)	IV-A (4)	IV-B (5)
			Panel A: Value Added		
Road density	0.100***	0.100***	0.070***	0.063***	0.103***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
RMSE	0.053	0.054	0.046	0.048	0.055
			Panel B: Gross Output		
Road density	0.103***	0.101***	0.071***	0.064***	0.106***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
RMSE	0.054	0.053	0.046	0.047	0.054
Rail density	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	6651	6654	6463	6651	6651
F	35.36	35.36	34.89	72.60	36.38

Notes: The dependent variable in col.(1) is log value-added TFP estimated using LP method (Levinsohn and Petrin 2003). In col.(2), we winsorise the top 1% and bottom 1% of log value-added TFP estimated using ACF method. In col.(3), we exclude the states of Bihar, Jharkhand, Madhya Pradesh, Chhattisgarh, Uttar Pradesh and Uttarakhand that were re-organized in 2000 from our analysis. In cols.(4) and (5), we present results from regressions where we instrument road density with 'Aligned x Seat share' and 'Aligned x Election year dummy', respectively. Panel A relates to value-added TFP whereas, panel B relates to gross-output TFP. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). All regressions include the full set of controls – time-interacted state-specific variables, a state's railway density along with year and state-industry fixed effects. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

due to agglomeration effects as firms form spatial clusters or other complementarities in production that kick-in to amplify the impact of road density.

Alternative Measures of TFP: Throughout this article, we mainly focus on TFP estimated by the ACF method since it avoids the functional dependence problem that affects labour coefficients in Olley and Pakes (1996) type methods, as mentioned before. In Table 12, we present estimates using alternative measures of TFP. In column 1, we present results from regressions where we estimate TFP(VA) using the Levinsohn and Petrin (2003) method (see Table 16 in Appendix B for first stage results). We find that the effect of roads on TFP is robust to the choice of alternative methods in estimating productivity. Importantly, the results are quite similar to those obtained by the ACF method.[11]

[11] We note that our results hold even if we use the Wooldridge (2009) method but do not show the results to conserve space.

Treating Outliers: One important issue in working with establishment-level data is that outliers might unduly influence our estimates. To avoid this, we winsorise the top 1% and the bottom 1% of the distribution of log TFP and re-run our estimates. Column 2 in Table 12 presents results after 1%/99% winsorisation of the dependent variable. We find that, even after treating outliers in this way, our results remain almost identical, evidence that our results are robust to outliers.

Excluding Split States: In 2000, three new states – Jharkhand, Chhattisgarh and Uttarakhand – were carved out of Bihar, Madhya Pradesh and Uttar Pradesh, respectively. It is possible that changes to state boundaries that reorganized the institutional mechanisms governing states affected state policies and hence industrial productivity in split states. One example of this is Uttarakhand which implemented tax incentive schemes to attract industries that affected productivity (see Chaurey 2016). To ensure that our results are robust to these changes, we exclude these six states and re-estimate our results. We find that excluding these states lowers the effect of roads on productivity to 7 percentage points, as expected, but remains statistically significant at the 1% level (Column 3 in Table 12).

Alternative Instruments: Thus far, we instrument road density with *Aligned*, a binary indicator for whether a state's ruling party is aligned with the centre during a year. However, since *Aligned* is a binary variable, we do not get a sense of whether the share of seats of the ruling party in a state has a bearing on the provision of roads, conditional on alignment. To examine this, we interact *Aligned* with the seat share of the ruling party and use this as our instrument (see column 4 in Table 12). This effectively weights aligned parties in different states in proportion to the share of seats they won in legislative assembly elections.

Besides this, it is also important to examine if political factors are indeed driving our results. One way to examine this is to consider alignment during an election year. To do this we interact *Aligned* with a dummy for an election year (see column 5 in Table 12) and use it as an instrument for road density.

We observe that: first, both the alternative instruments are informative (see Table 16 in Appendix B); and, second, the results in columns 4 and 5 in Table 12 are qualitatively similar to the main results in column 3 in Table 2. In fact, as expected, the coefficient on road density with 'Aligned x Election year dummy' as instrument (IV-A) is marginally higher than when we used *Aligned* as instrument. In contrast, accounting for the seat share of aligned ruling-parties across states (IV-A) slightly lowers the estimate of roads.

6 Conclusion

In this article, we examine the impact of roads on manufacturing productivity. By reducing transport costs and improving firm logistics, higher road density encourages specialisation, reduces inventory costs, extends the size of markets and increases trade. Here, we examine how road density affects multifactor productivity in organized Indian manufacturing during 1998-2012.

An important challenge in estimating the effect of roads on manufacturing TFP is that the placement of roads is not random. For example, roads might be more prevalent in states that trade more i.e. when roads are built to support trade, or if they trade less i.e. when roads are built to encourage trade. Moreover, states producing more output are also likely to have greater capacity to invest in roads. We overcome this inference problem by exploiting exogenous variation in the timing and duration of partisan alignment of states with the centre across multiple election cycles that are staggered across states. The premise is that in a federal democracy, the centre is likely to favour states that are aligned with it over those that are not. This influences the transfer of resources to aligned states, allowing them to deliver more public goods such as roads. Thus, *ex-ante* we expect road density within states to be higher when states are aligned with the centre than when states are not so aligned. Our first stage regression estimates confirm that the average road density in India is higher during times when a state is aligned than when it is not aligned. Our study therefore provides a political-economic explanation for differences in road density across states, which have first order impact on manufacturing productivity.

To avoid omitted variable bias, we include state-industry fixed effects to control for unobserved time-invariant heterogeneities such as industry-specific policies within a state. We control for railway density and include year interacted state-specific variables such as population, literacy, total main and marginal workers etc. To absorb year specific volatility in cost or orders we include year fixed effects.

After controlling for endogeneity and including the full set of controls, we find that at the industry-level an additional km of road per sq.km of area leads to a 10%-34% log point rise in TFP. We find that this impact varies by age, size and location of firms within the industry. Smaller firms, incumbents and those in urban states see most pronounced gains. Analysing changes between 1998 and 2012, we find that the results confirm the findings from panel data analysis. Moreover, the results are consistent across both the value-added and the gross-output TFP measures and across different methods in estimating productivity. To gain an intuitive understanding of the results we conduct elasticity estimates where we find that, on average, a one percentage point increase in road density leads to a 12 percentage point increase in industry-level TFP. The heterogeneity in elasticity estimates mirror findings from the semi-log model.

We conduct several robustness checks to validate our results. We use productivity estimated using alternative methods as our dependent variable. We vary how we measure road density and use different instruments to test the sensitivity of the estimates. Furthermore, we winsorise the dependent variable to guard against outliers and conduct placebo instrument tests where we replace the true instrument, *Aligned*, with a random draw of alignment status. We find that the results are robust to a battery of checks.

Thus, we find that higher road density boosts manufacturing productivity. But, its impact differs by age, size and location of firms. Moreover, higher road density affects not only the scale but also the growth rate of productivity. These findings might be particularly relevant to policymakers in developing countries who intend to address the shortfall in transportation infrastructure and raise manufacturing productivity.

Appendix

A Variables Used to Estimate TFP

In estimating TFP, we use the following variables: gross value of output (or value added output), total man-days worked, raw materials, power and fuel. The values of gross and value-added output were converted into real terms by deflating the nominal values by industry-specific wholesale price indices (WPI) whereas, expenses on raw materials and power and fuels were deflated by overall WPI. The price indices were obtained from the Office of the Economic Adviser, GoI (see `http://eaindustry.nic.in/home.asp`).

We follow the methodology in Balakrishnan et al. (2000) to measure capital stock (see also Topalova and Khandelwal (2011) and Kathuria and Sen (2014) for application of this method). We apply the perpetual inventory method (PIM) and adjust the book value of capital to reflect replacement cost instead of historic cost in which they are measured. To arrive at a measure of capital stock at replacement costs for a base year, we first assume that our base year is 2006. This choice is driven by the fact that we have maximum observations for that particular year. We then compute a revaluation factor assuming that the life of a machine is twenty years, and both the price of capital and the growth of investment changes at a constant rate throughout the assumed twenty years lifetime of capital stock. We use the revaluation factor to convert base year capital to capital at replacement cost in current prices. We then deflate the current value by a deflator based on Gross Fixed Capital Formation (GFCF) series obtained from the Ministry of Statistics and Programme Implementation (MOSPI), GoI. Finally, we obtain the capital stock for every period by summing over investments in subsequent years.

B Tables

Table 13: TFP in Indian Manufacturing (Gross Output), 1998-2012

Groups	Obs.	Mean	p10	p25	p50	p75	p90
All	6714	0.161	0.070	0.096	0.153	0.211	0.260
By Age:							
Young	5071	0.168	0.069	0.087	0.147	0.219	0.298
Incumbent	6595	0.163	0.069	0.096	0.156	0.214	0.266
By Size:							
Small	6333	0.198	0.067	0.100	0.190	0.273	0.337
Large	6173	0.130	0.071	0.087	0.120	0.158	0.203
By Location:							
Rural	3203	0.137	0.065	0.080	0.118	0.176	0.239
Urban	3511	0.185	0.086	0.132	0.187	0.228	0.274
TFP_{2012}/ TFP_{1998}	298	1.336	0.776	0.966	1.233	1.602	1.977

Notes: This table shows the dispersion of log gross-output Total Factor Productivity (TFP) in Indian manufacturing during 1998-2012 estimated using the Ackerberg, Caves and Frazer (2015) method. The firm-level log TFP values are normalised by dividing it by employment-weighted average productivity (in logs) for an industry-year. Row 'All' in the table corresponds to all observations in our sample. 'Young' includes industry-level estimates of only firms aged $\leq$ 4 years, whereas 'Incumbents' include those that are aged >4 years. 'Small' includes firms in the industry with fixed-assets lower than the median fixed assets in the industry, while 'Large' includes those with fixed assets greater than the industry median value. 'Rural' shows TFP for industries located in rural states where a rural state is defined to have urbanisation rate lower than the median value across all states in our sample.'Urban' shows TFP for industries located in states where the urbanisation rate exceeds the median value. TFP_{2012}/ TFP_{1998} is the ratio of log TFP values in 2012 and 1998.

Table 14: Effect of Roads on TFP (Gross output): Rural vs. Urban

Variables	All (1)	Small (2)	Large (3)	Young (4)	Incumbent (5)
			Panel A: Rural States		
Road density	0.129***	0.149***	0.111***	0.041	0.141***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Observations	3143	2886	3022	2683	3047
RMSE	0.052	0.068	0.048	0.069	0.056
F	26.04	25.05	18.67	21.82	22.55
			Panel B: Urban States		
Road density	0.224**	0.337***	0.123	0.345**	0.194**
	(0.09)	(0.12)	(0.08)	(0.16)	(0.08)
Observations	3508	3382	3089	2340	3485
RMSE	0.076	0.114	0.056	0.130	0.069
F	12.11	12.12	10.81	8.06	11.90
Rail density	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes

Notes: The columns indicate log gross-output TFP values for all firms in the industry (in col.(1)), for small firms (in col.(2)), for large firms (in col.(3)), for young firms (in col.(4)) and for incumbent firms in the industry (in col.(5)). Further, Panel A shows results for industries located in rural states where we classify a state to be rural if its urban population is less than the median urban population of all states in our sample while Panel B shows results for industries located in urban states i.e. states which exceed the median urban population. Road density is the total state-wide road length (in km) divided by its land area (in sq. km). The coefficients are from separate regressions of the respective dependent variables on road density, controlling for year-interacted state characteristics (in logs) – population, literacy, total main and marginal workers, total main workers in agriculture and industry – a state's railway density and state-industry and year fixed effects. RMSE=Root Mean Squared Error; F-stats>10 $\sim$ valid instruments. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 15: First Stage: Effect of Alignment on Dummy for High Road Density

	IV-I (1)	IV-II (2)
Aligned	0.026**	0.032***
	(0.01)	(0.01)
Rail density	Yes	Yes
Year FE	Yes	Yes
State variables	No	Yes
State-Industry FE	Yes	Yes
Observations	6651	6651
RMSE	0.255	0.254

Notes: First stage estimates from regressing High Road density on *Aligned* conditional on controls. High Road density is a dummy variable that takes a value of 1 if road density in a state exceeds the 60^{th} percentile of the overall road density, and zero otherwise. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 16: Robustness Checks: First Stage, Effect of Alignment on Road density

Variables	LP (1)	Winsorised (2)	No Split States (3)	IV-A (4)	IV-B (5)
Aligned	0.102***	0.102***	0.100***		
	(0.02)	(0.02)	(0.02)		
Aligned x Seat Share				0.003***	
				(0.00)	
Aligned x Election Year					0.097***
					(0.02)
Rail density	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Year x State variables	Yes	Yes	Yes	Yes	Yes
State-Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	6111	6111	5945	6111	6111
RMSE	0.315	0.315	0.318	0.311	0.317

Notes: Results from first stage estimates. Standard errors clustered at state-industry level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References

Abadie, A., Athey, S., Imbens, G. W. and Wooldridge, J. (2017), When should you adjust standard errors for clustering?, Technical report, National Bureau of Economic Research.

Ackerberg, D. A., Caves, K. and Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**(6), 2411–2451.

Arulampalam, W., Dasgupta, S., Dhillon, A. and Dutta, B. (2009), 'Electoral goals and center-state transfers: A theoretical model and empirical evidence from India', *Journal of Development Economics* **88**(1), 103–119.

Balakrishnan, P., Pushpangadan, K. and Babu, M. S. (2000), 'Trade liberalisation and productivity growth in manufacturing: Evidence from firm-level panel data', *Economic and Political weekly* pp. 3679–3682.

Banerjee, A., Duflo, E. and Qian, N. (2012), On the road: Access to transportation infrastructure and economic growth in China, Technical report, National Bureau of Economic Research.

Baskaran, T., Min, B. and Uppal, Y. (2015), 'Election cycles and electricity provision: Evidence from a quasi-experiment with Indian special elections', *Journal of Public Economics* **126**, 64–73.

Bloomberg (2017), 'Modi to spend $59 billion to upgrade india's infrastructure'.
URL: *https://www.bloomberg.com/news/articles/2017-02-01/modi-plans-59-billion-rail-road-push-as-bombardier-ge-invest*

Bohlken, A. T. (2016), Development or rent-seeking: How political influence shapes infrastructure provision in India, Technical report, Mimeo.

Bracco, E., Lockwood, B., Porcelli, F. and Redoano, M. (2015), 'Intergovernmental grants as signals and the alignment effect: Theory and evidence', *Journal of Public Economics* **123**, 78–91.

Chandra, A. and Thompson, E. (2000), 'Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system', *Regional Science and Urban Economics* **30**(4), 457–490.

Chaurey, R. (2016), 'Location-based tax incentives: Evidence from India', *Journal of Public Economics* .

Chetty, R., Looney, A. and Kroft, K. (2009), 'Salience and taxation: Theory and evidence', *American economic review* **99**(4), 1145–77.

Datta, S. (2012), 'The impact of improved highways on Indian firms', *Journal of Development Economics* **99**(1), 46–57.

Donaldson, D. (2018), 'Railroads of the Raj: Estimating the impact of transportation infrastructure', *American Economic Review* **108**(4-5), 899–934.

Duranton, G., Morrow, P. M. and Turner, M. A. (2014), 'Roads and trade: Evidence from the US', *Review of Economic Studies* **81**(2), 681–724.

Faber, B. (2014), 'Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System', *Review of Economic Studies* **81**(3), 1046–1070.

Fernald, J. G. (1999), 'Roads to prosperity? assessing the link between public capital and productivity', *American economic review* **89**(3), 619–638.

Ghani, E., Goswami, A. G. and Kerr, W. R. (2016*a*), '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.

Ghani, E., Goswami, A. G. and Kerr, W. R. (2016*b*), 'Highways and spatial location within cities: Evidence from India', *The World Bank Economic Review* **30**(Supplement_1), S97–S108.

Guasch, L. J. and Joseph, K. (2001), 'Inventories in developing countries', *World Bank Policy Research Woeking Paper 2552* .

Gulyani, S. (2001), 'Effects of poor transportation on lean production and industrial clustering: Evidence from the Indian auto industry', *World development* **29**(7), 1157–1177.

Gupta, P., Hasan, R. and Kumar, U. (2008), 'What constrains Indian manufacturing?', *ERD Working Paper SeriesNo. 119, Asian Development Bank* .

Holl, A. (2016), 'Highways and productivity in manufacturing firms', *Journal of Urban Economics* **93**, 131–151.

Hulten, C. R., Bennathan, E. and Srinivasan, S. (2006), 'Infrastructure, externalities, and economic development: a study of the Indian manufacturing industry', *The World Bank Economic Review* **20**(2), 291–308.

Jedwab, R. and Moradi, A. (2016), 'The permanent effects of transportation revolutions in poor countries: Evidence from Africa', *Review of economics and statistics* **98**(2), 268–284.

Johansson, E. (2003), 'Intergovernmental grants as a tactical instrument: empirical evidence from Swedish municipalities', *Journal of Public Economics* **87**(5-6), 883–915.

Kathuria, V. and Sen, K. (2014), *Productivity in Indian Manufacturing: Measurement, Method and Analysis*, Routledge.

Khemani, S. (2003), *Partisan politics and intergovernmental transfers in India*, The World Bank.

Khemani, S. (2007), The political economy of equalization transfers, *in* 'Fiscal Equalization', Springer, pp. 463–484.

Lehne, J., Shapiro, J. N. and Eynde, O. V. (2018), 'Building connections: Political corruption and road construction in India', *Journal of Development Economics* **131**, 62–78.

Levinsohn, J. and Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The Review of Economic Studies* **70**(2), 317–341.

Li, H. and Li, Z. (2013), 'Road investments and inventory reduction: Firm level evidence from China', *Journal of Urban Economics* **76**, 43–52.

Mani, A. and Mukand, S. (2007), 'Democracy, visibility and public good provision', *Journal of Development Economics* **83**(2), 506–529.

Martin, L. A., Nataraj, S. and Harrison, A. E. (2017), 'In with the big, out with the small: Removing small-scale reservations in India', *American Economic Review* **107**(2), 354–86.

Martin, P., Mayer, T. and Mayneris, F. (2011), 'Spatial concentration and plant-level productivity in France', *Journal of Urban Economics* **69**(2), 182–195.

Matthias, A. J., Beata, J., Molly, L. and Aaditya, M. (2016), 'Services reform and manufacturing performance: Evidence from india', *The Economic Journal* **126**(590), 1–39.

Michaels, G. (2008), 'The effect of trade on the demand for skill: Evidence from the interstate highway system', *The Review of Economics and Statistics* **90**(4), 683–701.

Mitra, A., Sharma, C. and Véganzonès-Varoudakis, M.-A. (2012), 'Estimating impact of infrastructure on productivity and efficiency of Indian manufacturing', *Applied Economics Letters* **19**(8), 779–783.

Mitra, A., Varoudakis, A. and Veganzones-Varoudakis, M.-A. (2002), 'Productivity and technical efficiency in Indian states' manufacturing: the role of infrastructure', *Economic development and cultural change* **50**(2), 395–426.

Olley, G. S. and Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

Redding, S. J. and Turner, M. A. (2015), Transportation costs and the spatial organization of economic activity, *in* 'Handbook of regional and urban economics', Vol. 5, Elsevier, pp. 1339–1398.

Sengupta, B. (2011), 'Provision of public goods in a federal economy: The role of party politics', *European Journal of Political Economy* **27**(1), 104–119.

Shatz, H. J., Kitchens, K. E. and Rosenbloom, S. (2011), *Highway infrastructure and the economy: implications for federal policy*, Rand Corporation.

Shirley, C. and Winston, C. (2004), 'Firm inventory behavior and the returns from highway infrastructure investments', *Journal of Urban Economics* **55**(2), 398–415.

Solé-Ollé, A. (2013), 'Inter-regional redistribution through infrastructure investment: Tactical or programmatic?', *Public Choice* **156**(1-2), 229–252.

Solé-Ollé, A. and Sorribas-Navarro, P. (2008), 'The effects of partisan alignment on the allocation of intergovernmental transfers. differences-in-differences estimates for Spain', *Journal of Public Economics* **92**(12), 2302–2319.

Storeygard, A. (2016), 'Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa', *The Review of Economic Studies* **83**(3), 1263–1295.

Syverson, C. (2011), 'What determines productivity?', *Journal of Economic literature* **49**(2), 326–65.

Topalova, P. and Khandelwal, A. (2011), 'Trade liberalization and firm productivity: The case of India', *Review of economics and statistics* **93**(3), 995–1009.

WB (2014), World bank enterprise survey, Technical report, World Bank.

Wilkinson, S. I. (2006), 'The politics of infrastructural spending in India', *Department of Political Science, University of Chicago, mimeo* **31**.

Williams, M. J. (2017), 'The political economy of unfinished development projects: Corruption, clientelism, or collective choice?', *American Political Science Review* **111**(4), 705–723.

Wooldridge, J. M. (2009), 'On estimating firm-level production functions using proxy variables to control for unobservables', *Economics Letters* **104**(3), 112–114.