INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 01542

July 2016

Roads to Innovation

Firm-Level Evidence from China

Xu Wang

Xiaobo Zhang

Zhuan Xie

Yiping Huang

Development Strategy and Governance Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

AUTHORS

Xu Wang is a PhD candidate at the National School of Development, Peking University, Beijing.

Xiaobo Zhang (x.zhang@cgiar.org) is a senior research fellow in the Development Strategy and Governance Division of the International Food Policy Research Institute, Washington, DC and a distinguished professor of economics in the National School of Development, Peking University, Beijing.

Zhuan Xie is a research fellow at State Administration of Foreign Exchange, China, Beijing.

Yiping Huang is a distinguished professor of economics in the National School of Development, Peking University, Beijing and a member of the Central Bank's Monetary Policy Committee, China.

Notices

[1.] IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by the International Food Policy Research Institute.

[2.] The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

Contents

Tables

Figures

ABSTRACT

Although both infrastructure and innovation play an important role in fostering a country's economic growth, discussion in the literature about how the two are connected is limited. This paper examines the impact of road density on firm innovation in China using a matched patent database at the firm level and road information at the city level. Regional variation in the difficulty of constructing roads is used as an instrumental variable to address the potential endogeneity problem of the road variable. The empirical results show that a 10 percent improvement in road density increases the average number of approved patents per firm by 0.71 percent. Road development spurs innovation by enlarging market size and facilitating knowledge spillover.

Keywords: infrastructure, innovations, transportation cost, knowledge diffusion

JEL: O31, O33, R11, R40

ACKNOWLEDGMENTS

This work was undertaken as a part of, and funded by, the CGIAR Research Program on Policies, Institutions, and Markets, which is led by IFPRI and funded by the CGIAR Fund Donors. We are grateful to the comments received from participants at the 2016 Chinese Economists Society North America Conference and the Conference of Urbanization and Infrastructure in Asia organized by Asian Development Bank Institute. This paper has not gone through IFPRI's standard peer-review process. The opinions expressed here belong to the author, and do not necessarily reflect those of PIM, IFPRI, or CGIAR.

1. INTRODUCTION

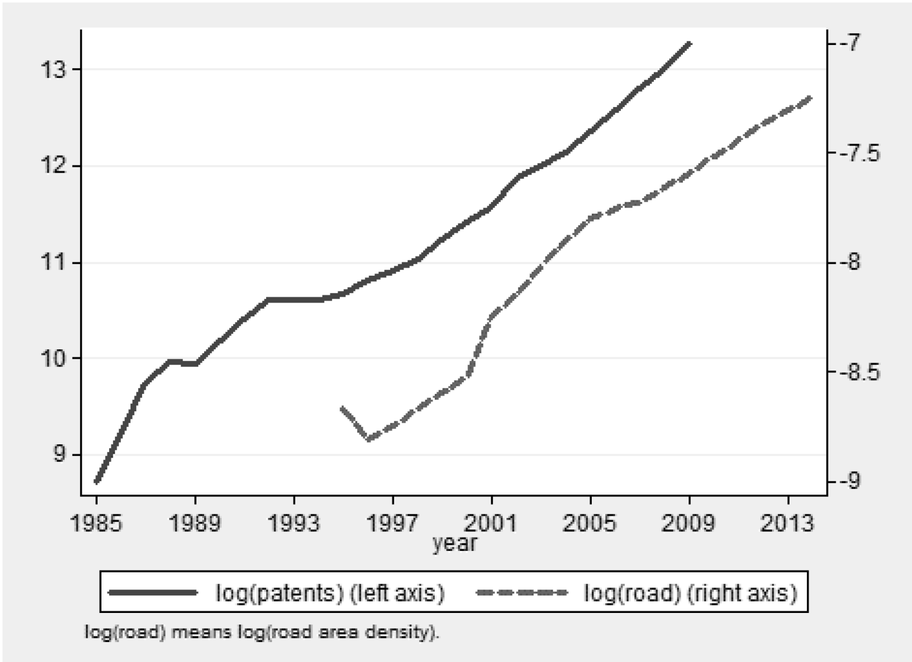
Infrastructure and innovation both play important roles in fostering a country's economic growth. As the "wheels" of economic activity (World Bank 1994), infrastructure provides access to basic services, lowers production and transaction costs, promotes trade, and helps accumulate physical and human capital. Following the pioneering work by Aschauer (1989), a large body of literature has investigated the economic impact of public infrastructure.[1] In parallel, the literature on innovation and growth has exploded in the past several decades (Romer 1990; Aghion and Howitt 1992). However, the literature on the relationship between infrastructure and innovation is comparably much more scant with a few exceptions. For example, Acemoglu, Moscona, and Robinson (2016) show that the spread of post offices in the United States in the 19th century, largely as a result of railway expansion, stimulated patenting activities. Agrawal, McHale, and Oettl (2014) find that better road access expands the distance among patents cited to each other in the United States, suggesting a strong knowledge diffusion effect of road development.

Following the spirit of those two papers, we examine the impact of road development on firm innovation in China. Our paper contributes to the literature in several dimensions. First, it is among the first to examine the relationship between infrastructure and firm innovation in a developing country.

The rapid infrastructure development and surge in patent activity in the past several decades in China offers a good setting to study the impact of road development on innovation. Since the early 1990s, China has made tremendous progress in building and upgrading transportation infrastructure. Road density, an important indicator of transportation infrastructure, has increased by 7.3 percent per year since 1990. In comparison, it has barely changed in India in the same period. Along with the rapid improvement in transportation infrastructure, the introduction of innovations has surged in the same period. The number of approved patents, a common measure of innovation, has grown by more than 20 percent per year. China has become the leading country filing patents since 2012. As Figure 1.1 shows, both road density and number of approved patents display an upward trend and strong co-movement over time. Across Chinese cities, the two variables also exhibit a high correlation, as Figure 1.2 reveals. Of course, the positive relationship across space and over time is just suggestive and does not imply any causality of road development on innovation. In our study, we control for other important factors that may shape innovations.

[1] See Romp and De Haan (2007) for relevant reviews.

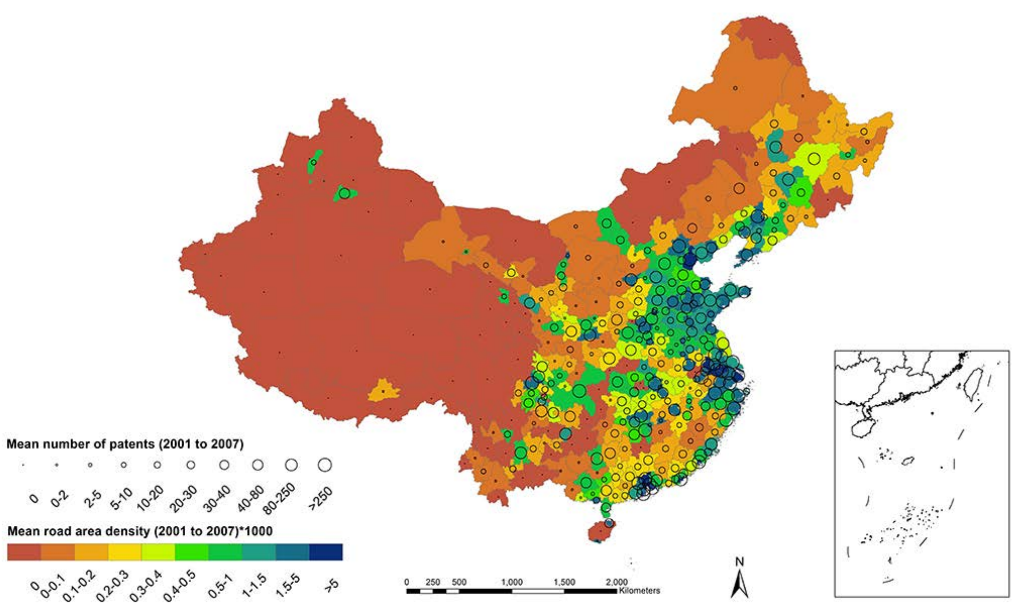
Figure 1.1 Temporal correlation between patents and road density



Source: The patents information is from the Chinese Patent Database (SIPO, various years). The road information is from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: The figure displays the temporal correlation between patents and road density. We aggregate the patents data and road data into year level.

Figure 1.2 Spatial correlation between number of patents and road density



Source: The patents information is from the Chinese Patent Database (SIPO, various years). The road information is from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: The figure depicts the spatial correlation between number of patents and road density. We aggregate the firm-level data into city-year level, and calculate mean number of patents for each city during the period 2001–2007.

Despite our effort to control for as many key variables as possible, the possibility exists that innovative activity and road infrastructure status are determined by some unobservable factors that are missing in the regressions. To address that problem, we exploit an instrumental variable approach. Following Saiz (2010) and Duflo and Pande (2007), we use mean slope in a city (that is, the incline of the road) to measure the relative difficulty (cost) of road construction at the city level. Yet the slope variable is time-invariant. To remedy that problem, we use the interaction of slope and yearly export price of road construction machinery as an instrumental variable and find that a 10 percent improvement in road density can increase the average of number of patents per firm by 0.71 percent. This effect is economically significantly large, equivalent to a 1.42 percent increase in an average firm's research and development (R&D) investment.

Second, in addition to quantifying the effect of road development on innovation, we further examine the main channels. It has been well regarded that knowledge is a fundamental factor behind innovation. However, knowledge (information) is often spatially dispersed (Krugman 1991; Jaffe, Trajtenberg, and Henderson 1993). With better road connections, people can travel long distances and communicate more widely, making it more likely for people to cross-fertilize ideas (Moretti 2004; Glaeser and Gottlieb 2009; Agrawal, Galasso, and Oettl 2014). Our study confirms that knowledge diffusion is an important channel through which road development sparks firm innovation. Apart from the knowledge-diffusion channel, we also examine the channel of market expansion. Consistent with the literature (Romer 1990; Aghion and Howitt 1992; Acemoglu and Linn 2004), we find that improved roads expand market size, which in turn leads to more innovation.

Third, our paper contributes to the emerging body of literature explaining the surge in patent activity in China. In the literature, foreign direct investment and government subsidies on patent applications are considered two key contributors to the observed rapid patent growth (Hu and Jefferson 2009; Hu 2010; Thoma 2013; Li 2012; Dang and Motohashi 2015). Our paper contributes to this branch of literature by showing that the expansion of transportation infrastructure is another explanatory factor in China's patent growth.

Finally, our paper is associated with the literature evaluating the impact of transportation infrastructure on various outcome variables, such as productivity (Fernald 1999), decentralization (Baum-Snow et al. 2012), the distribution of workers (Duranton and Turner 2012), trade-related activity (Donaldson 2010; Duranton, Morrow, and Turner 2014), and long-term gross domestic product (GDP) growth (Banerjee, Duflo, and Qian 2012). Our paper highlights the significant impact of transportation infrastructure on innovation, a largely neglected topic in the literature, in particular in the context of developing countries.

We organize the paper as follows. Section 2 describes data and methods. Section 3 reports estimation results and various robustness checks. Section 4 discusses the underlying mechanisms. Section 5 concludes.

2. DATA AND EMPIRICAL STRATEGY

Data

The main dataset we use is the Annual Survey of Industrial Enterprises in China (ASIEC) conducted by the National Bureau of Statistics of China. It covers all state-owned manufacturing firms and those non-state-owned manufacturing enterprises above a certain size for the 1998–2007 period.[2] More than 100,000 firms appeared at least once in the database, and among them, 27,575 firms were surveyed throughout the whole sample period.

We use approved patents as a measure of innovation. Patent data come from the Chinese Patent Database, which includes all patents approved before 2014 and patent applications before 2009.[3] The patent dataset and the ASIEC have been matched in Xie and Zhang (2015) for the period 1998–2008. We use their matched dataset in our analysis. In the matched database, of all the 2,081,656 firm-year observations, 2.8 percent have one or more patents. For those with at least one patent, more than 70 percent possess fewer than four patents (Table 2.1). Because a firm may not receive patents in all the years, its chance of receiving at least one patent over the whole sample period is higher than in a particular year.

Table 2.1 Patents summary

Number of patents per firm-year	Number of patents	Percentage
0	2,023,149	97.19
1 or more	58,507	2.81
Total	2,081,656	100
For those who have one or more patents		
1	23,428	40.04
2	11,728	20.05
3	5,953	10.17
4	3,900	6.67
5	2,535	4.33
6 or more	10,963	18.74
Total	58,507	100

Source: Calculated by authors based on the merged firm patent database (1998–2007) between the Chinese Patent Database (SIPO, various years) and the Annual Survey of Industrial Enterprises in China Database (NBSC, various years [a]).

Road data are obtained from the China City Statistical Yearbook. There are two road measures at the city level—road length and road area.[4] However, road length is not appropriate for our analysis for two reasons. First, road length masks the quality difference among different kinds of road. For example, although highways are of higher quality than rural roads, just looking at road length would not reveal the difference. Second, a change in statistical definition for the road length variable occurred in 2005, likely contributing to a jump in road length in 2005 (Figure A.1 in the appendix). In comparison, the area of

[2] It covers all firms with annual sales in excess of 5 million yuan, about US$0.6 million according to the exchange rate in 2005. They account for more than 85 percent of China's industrial output value.

[3] We use approved patents to measure a firm's innovation ability. Other indicators, including the firm's total factor productivity, R&D expenditure, technology secrets, and so on, have also been used in the literature to measure innovation (Bloom, Draca, and Van Reenen 2011). Patent data offer the most detailed and systematically compiled and managed information about innovation in China (Choi, Lee, and Williams 2011). More important, these data are systematically available unlike other commonly used indicators. This database contains 4,060,392 observations, including 1,097,000 invention patents, 1,620,069 utility model patents, and 1,343,323 design patents.

[4] Road length refers to the length of roads of all grades that have been put in use. Road area stands for the area of road surface.

paved road is less subject to the measurement problem of road quality and changes in statistical definition of road length. Since a spatially large city tends to have more paved roads, we use the share of road area in total city administrative area ("road area density") as a measure in our main analysis.

Estimation Specification

Our basic specification takes the following form:

$$Innovation_{i,c,t} = \beta_0 + \beta_1 * log(road_{c,t}) + \beta_2 * X_{it} + \mu_i + \alpha_t + \varepsilon_{i,t}, \tag{1}$$

where $\text{Innovation}_{i,c,t}$ represents the innovation level of firm *i* at city *c* in year *t*, $\log(road_{c,t})$ refers to the log form of road area density of city *c* in year *t*, μ_i stands for firm-level fixed effects, α_t refers to year fixed effects, and X_{it} is a vector of additional controls, including an industry specialization index, a product market competition measure, and a measure of local economic development.

Although the literature often assumes that the number of patents follows a Poisson or negative binomial distribution (Cameron and Trivedi 2013), we use ordinary linear regression for two reasons. First, as remarked by Angrist and Krueger (2001), it is more convenient to handle potential endogenous problems in a linear empirical framework than nonlinear models. Second, it is impossible to control for firm fixed effects in the Poisson or negative binomial regression because many firms do not have any patents during the period 2001–2007. The conditional maximum likelihood estimation of the Poisson or negative binominal model would drop the firms without patents for the entire time period, resulting in a loss of sample size by 93 percent.[5]

We use two measures for $\text{Innovation}_{i,c,t}$. The first one is the log form of total patents for a firm, "log (patents+1)," following Bloom, Draca, and Van Reenen (2011), Agrawal et al. (2014), and Acemoglu, Moscona, and Robinson (2016). The second measure is a dummy variable, "having at least one patent," being one if a firm has at least a patent in a specific year and zero otherwise.

Considering the potential lagging effect of roads, we employ the average road density of city *c* in year *t* during the last four years (*t*-3, *t*-2, *t*-1, and *t*) as the main measure of our road variable. If a firm was established last year, then we use the average road density from last year and this year; if a firm was just established this year, then we just use current road density. As a robustness check, we also try three-year lag and five-year lag to construct our key explanatory variable, and the results are similar.

One main empirical challenge in estimating equation 1 is possible correlations between unobserved factors $\varepsilon_{i,t}$ and road area density $\log(road_{c,t})$. For example, in areas with high growth potential, local governments may invest more in infrastructure. Meanwhile, local firms also likely respond to the same growth opportunities by investing in R&D. As a result, the observed spike in firm innovative activity is likely driven by unobserved growth potential rather than road development. Alternatively, when facing a tight budget, local governments' investment in roads may crowd out subsidies extended to firms in support of their innovations. When the error term and road variable are correlated, the estimates will be biased. To address this potential problem, we exploit an instrumental variable approach.

Instrumental Variable

Following Saiz (2010) and Duflo and Pande (2007), we use slope to measure the relative difficulty (cost) of constructing roads. The greater the mean slope for a city, the higher the cost to build roads. However, the slope variable at the city level is time-invariant, and it is problematic to use it as an instrument for the time-varying road variable. To overcome this problem, we also consider a time-varying variable, the

[5] In fact, we also run conditional maximum likelihood estimations of a negative binominal model as a robustness check. The results are robust.

weighted export price of road-building machinery, and interact it with the proportion of area with a slope of greater than 15 degrees to proxy for road construction costs.[6]

The slope variable itself may not meet the exclusion restriction because geographical conditions likely affect industrial structure, which in turn matters to firm behavior. The interaction of slope with the cost of road-building machinery can partly ameliorate the problem as it varies over time and is less likely to be correlated with unobserved factors, if any. Another way to reduce the omitted-variable problem is to control for as many related variables as possible in the regressions; such variables include an industry specialization index, a product market competition measure, and a measure of local economic development. As Duranton and Turner (2012) point out, the validity of instrumental variable estimation hinges upon the orthogonality of the dependent variable and the instrument conditional on control variables, not on unconditional orthogonality. Table 2.2 reports summary statistics for the variables.

Table 2.2 Variables summary

Variable	Observation	Mean	Standard deviation	Minimum	Maximum
Left-hand variables					
log(total patents+1)	1,630,400	0.041	0.271	0	8.406
log(invention+1)	1,630,400	0.009	0.112	0	8.321
log(design+1)	1,630,400	0.018	0.183	0	7.005
log(utility+1)	1,630,400	0.022	0.183	0	6.506
log(product+1)	1,630,400	0.040	0.267	0	8.161
log(process+1)	1,630,400	0.005	0.084	0	8.297
Right-hand variables					
log(road)	1,610,392	-6.457	1.495	-14.842	-2.741
log neighbor road (1st)	1,603,010	-6.907	1.064	-11.677	-4.532
log neighbor road (1st+2nd)	1,603,171	-6.993	0.888	-11.771	-5.213
log neighbor road (1st+2nd+3rd)	1,603,281	-7.132	0.752	-9.722	-5.589
industry specialization index	1,628,565	0.009	0.008	-0.027	0.513
competition index	1,514,100	0.030	0.012	0.000	0.061
log GDP per capita	1,624,964	9.973	0.728	7.601	11.502
Instrumental variables					
ratio15 (local)*eprice	1,554,680	0.403	0.390	0.000	2.323
ratio15 (1st neighbor)*eprice	1,546,477	0.453	0.353	0.001	2.007
ratio15 (1st+2nd neighbor)*eprice	1,550,971	0.484	0.324	0.002	2.010
ratio15 (1st+2nd+3rd neighbor)*eprice	1,552,045	0.496	0.288	0.002	1.882

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: "ratio15" stands for proportion of area with slope greater than 15 degrees; "eprice" indicates weighted average price of the exported equipment. Industry specialization index is a concentration index for each two-digit industry proposed by Ellison and Glaeser (1997). Competition index is log(1/HHI), where HHI is the Herfindahl-Hirschman index. Log GDP per capita is the log form of GDP per capita.

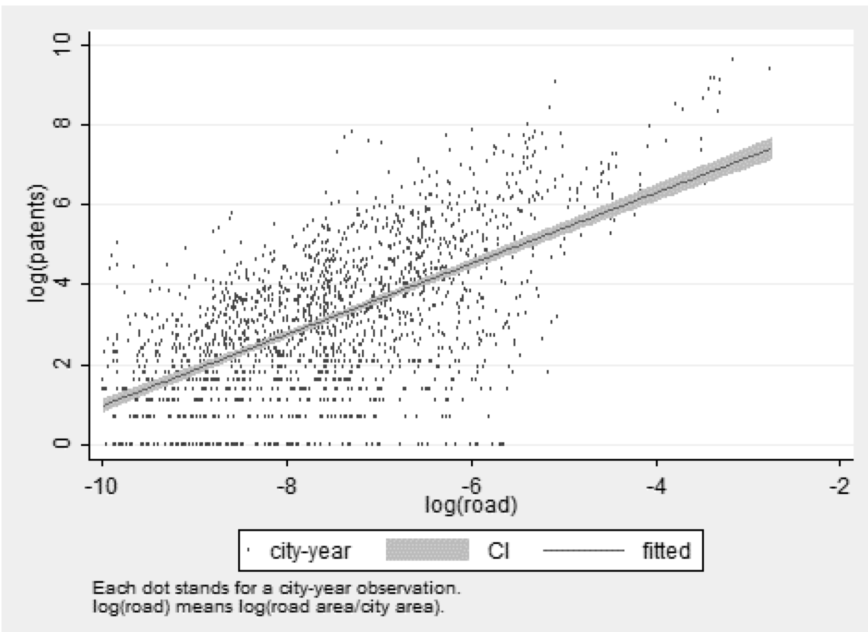
[6] The export price data for road-building machinery are from the Chinese Longitudinal Firm Trade Transaction Database. Road-building machines (with customs number 8,429) include motorized bulldozers, side-shovel bulldozers, road graders, graders, scrapers, mechanical shovels, excavators, shovel loaders, tamping machines, and road rollers. Because the road measure is based on the average for up to the past four years depending on a firm's survival status, we use the corresponding period of export price information to construct the instrumental variable.

3. ESTIMATION RESULTS

Benchmark Results

Table 3.1 presents the ordinary least squares (OLS) estimation results based on the basic specification in equation 1.[7] In the first regression on the log(patents +1), only the road density variable is included as an independent variable. It is significantly positive. After adding more control variables—industry specialization index, product market competition, and local economic development—one by one in regressions 2 through 4, the results hold. Columns 5 through 8 repeat the regressions in the first four columns by using "having at least a patent" as the dependent variable. Although the road variable is positive, it is no longer significant.

Figure 3.1 Raw correlation between patents and road density



Source: Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: The figure depicts the raw correlation between patents and road density. We aggregate the data into city-year level, and each dot in the figure represents a city-year observation.

To address potential endogeneity problems of the road variable in Table 3.1, Table 3.2 presents the first stage and second stage of instrument variable estimations. As shown in the first two columns on the first-stage regressions, the coefficient for the interaction term between slope and weighted price of imported construction machinery in two regressions on road density, one without any controls and one with a set of control variables, is significantly negative, indicating that the instrument variable has predictive power on road density.

[7] See Figure 3.1 for the correlation between number of patents and road density at the city-year level.

Table 3.1 Benchmark ordinary least squares results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	**Log(patents+1)**				**Having at least one patent**			
Log(road)	0.004**	0.004**	0.004*	0.004*	0.000	0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Industry specialization index		-0.006	-0.075	-0.079		0.006	-0.040	-0.041
		(0.078)	(0.116)	(0.116)		(0.053)	(0.077)	(0.077)
Competition index/100			-0.169	-0.211			-0.198	-0.234
			(0.221)	(0.220)			(0.144)	(0.144)
Competition index^2/10000			0.169	0.725			1.691	2.126
			(3.559)	(3.546)			(2.281)	(2.278)
Log GDP per capita				-0.001				0.002
				(0.003)				(0.002)
Firm fixed effects (FE)	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Age group FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,610,392	1,610,392	1,498,754	1,496,815	1,610,392	1,610,392	1,498,754	1,496,815
R-squared	0.601	0.601	0.606	0.607	0.514	0.514	0.520	0.521

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Industry specialization index is a concentration index for each two-digit industry proposed by Ellison and Glaeser (1997). Competition index is log(1/HHI), where HHI is the Herfindahl-Hirschman index. Log gross domestic product (GDP) per capita is the log form of GDP per capita. Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table3.2 Benchmark: First-stage and two-stage least squares results

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Log(road)		Log(patents+1)		Having at least one patent	
	First stage		2SLS			
Ratio15*eprice	-0.101***	-0.107***				
	(0.002)	(0.002)				
Log(road)			0.065***	0.071***	0.034**	0.038***
			(0.021)	(0.021)	(0.014)	(0.014)
Industry specialization index		0.708***		-0.127		-0.068
		(0.069)		(0.098)		(0.065)
Competition index/100		0.276*		-0.220		-0.238**
		(0.152)		(0.184)		(0.121)
Competition index^2/10000		16.319***		-0.551		1.370
		(2.442)		(2.997)		(1.924)
Log GDP per capita		-0.018***		-0.000		0.003
		(0.004)		(0.003)		(0.002)
Observations	1,495,375	1,384,771	1,495,375	1,384,771	1,495,375	1,384,771
R-squared	0.565	0.578				
Number of firms	361,408	338,801	361,408	338,801	361,408	338,801
First-stage F-stat			1992	2449	1992	2449

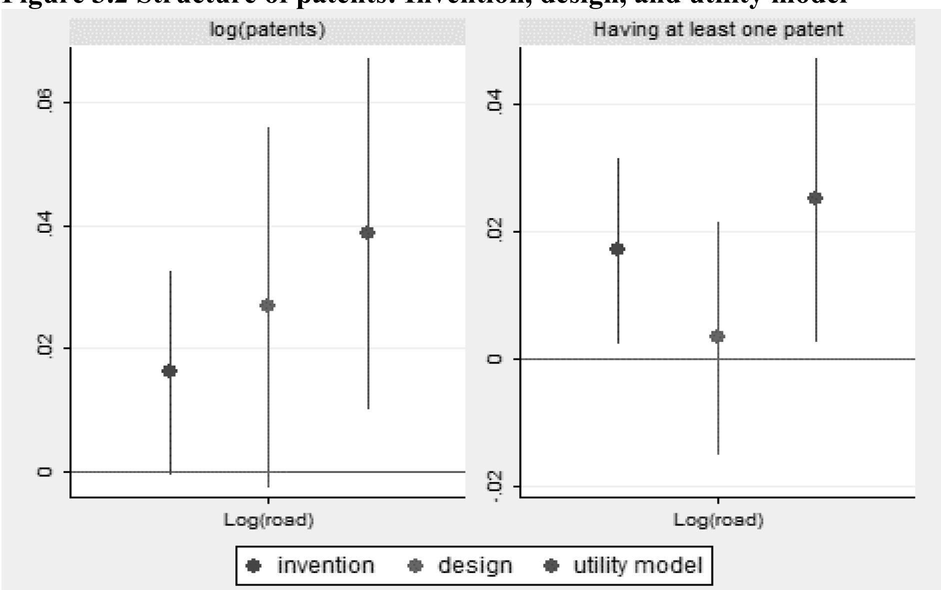
Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: For all specifications, we include the firm, year, and age group fixed effects. "Ratio15" stands for proportion of area with slope greater than 15 degrees; "eprice" indicates weighted average price of the exported equipment. Industry specialization index is a concentration index for each two-digit industry proposed by Ellison and Glaeser (1997). Competition index is log(1/HHI), where HHI is the Herfindahl-Hirschman index. Log gross domestic product (GDP) per capita is the log form of GDP per capita. Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Columns 3 through 6 display the two-stage least squares (2SLS) results on log(patents+1) and the dummy "having at least one patent." The first-stage F-statistics in all the four second-stage regressions pass the weak instrumental variable test. The road variable is statistically positive in all the four regressions. In column 4 with control variables, the coefficient for road density is 0.071, implying that 10 percent more roads would increase a firm's number of patents by 0.71 percent. Considering that the elasticity of corporate patenting to R&D expenditure in the innovation literature is estimated to be close to 0.5 (Aghion, Van Reenen, and Zingales 2013; Bloom, Schankerman, and Van Reenen 2013; Agrawal, Galasso, and Oettl 2014), a 10 percent increase in road density is roughly equivalent to a 1.42 percent increase in corporate R&D investment. The effect is economically sizable. According to column 6, a firm's probability of having patents would increase by 0.38 percent if road density increases by 10 percent. Per capita GDP is not significant in both the OLS and 2SLS regressions, probably due to inherent collinearity between firm fixed effects and per capita GDP at the city level. If we drop per capita GDP in the regressions, the main results still hold. Overall, the 2SLS estimates in Table 3.2 are larger than the corresponding OLS coefficients in Table 3.1, indicating a downward bias of OLS estimates.

Having investigated the impact of road density on total patents, we further consider the impact on three types of patent—invention, design, and utility model—and re-estimate equation 1.[8] The coefficient for the road variable and its confidence interval are reported in Figure 3.2. The coefficient for the road variable is positive and significant for the innovation and utility model patents, but not for design patents. Next, we divide patents into process and product innovations, repeat the estimation of equation 1, and present the results in Figure 3.3.[9] As shown in the figure, road density plays a significant role in shaping production innovation but not process innovation. It is very likely that road improvement enlarges market size, thereby making it possible to recover large fixed costs normally associated with product innovations.

Figure 3.2 Structure of patents: Invention, design, and utility model



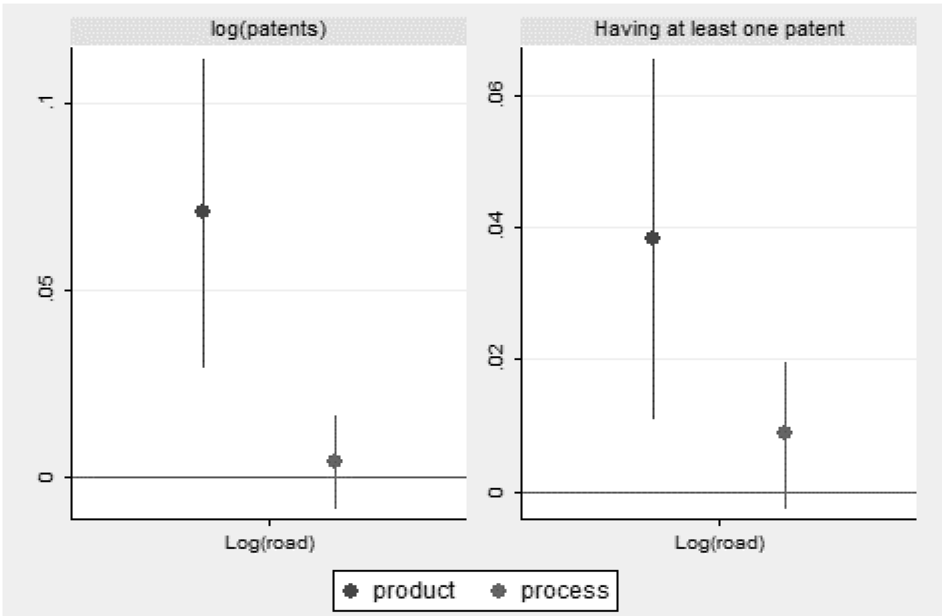
Source: Calculated by authors based on the regressions mentioned above. The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: The figure reports the coefficient of log(roads) and its 95 percent confidence interval in regressions for different types of patent. We use the same specification as columns 4 and 6 in Table 3.2 but change the dependent variable into three different kinds of patent: invention, design, and utility model.

[8] See Figure A.1 in the appendix for the patterns of the three types of patent.

[9] We use the following procedure to classify product and process innovations: if the name of an invention patent contains the word "method," we define it as a process innovation; otherwise, we define it as a product innovation. If it contains the words "method" and "and/or," it will be defined as belonging to both types. We acknowledge that the method is rather crude. The further refinement of the classification of product and process innovations remains a future research topic.

Figure 3.3 Structure of patents: Process and product



Source: Calculated by authors based on the regressions mentioned above. The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: The figure reports the coefficient of log(roads) and its 95 percent confidence interval in regressions for different types of patent. We use the same specification as columns 4 and 6 in Table 3.2 but change the dependent variable into two different kinds of patent: process patents and product patents.

Robustness Checks

In this subsection, we perform a variety of robustness checks for the benchmark specification of equation 1. Table 3.3 presents a falsification test on a randomly assigned road variable. To alleviate the concern about some omitted variables driving both road density and firm innovation that may lead to a spurious relationship, we randomly assign a firm to a city and then use the newly generated road density variable in the randomly matched cities to estimate equation 1. After repeating the process 50 times, bootstrapped coefficients and standard errors are obtained. The coefficients in Table 3.3 are statistically indifferent from zero, largely dismissing the concern about common trends.

Table 3.4 performs another falsification test on a subsample of state-owned enterprises (SOEs).[10] In principle, private firms are more responsive to market signals than SOEs, which tend to receive lucrative government supports and face less market competitive pressures. Therefore, we expect to see a much larger impact of road improvement on private firms than on SOEs. As Table 3.4 shows, this is indeed the case. Road improvement enhances the innovative activity only of private firms and not of SOEs.

We also change the lag length to three years or five years. The results are similar whether we use a three-year or five-year lag. Due to space limitations, the results are not reported here. Sine export prices of road construction machines from the Chinese Longitudinal Firm Trade Transaction Database are available for only 2000 to 2006, in the previous estimates, we use the price information of 2004–2006 as a proxy for the missing information in 2007. This may create some measurement errors. To address this concern, we drop observations in 2007 from the sample and re-estimate equation 1. The results still hold. To save space, we do not show the results here.

[10] In Table 3.4, we define a firm as an SOE if the state owns at least half of its registered capital.

Table 3.3 Falsification test

	(1)	(2)	(3)	(4)
Variable	**log(patents)**		**Having at least one patent**	
log(road)	OLS	2SLS	OLS	2SLS
Bootstrap coefficient	0.001	-0.070	0.001	-0.028
Bootstrap standard error	0.001	0.062	0.001	0.041
Mean t-stat	0.725	1.217	0.655	0.900
Other controls	YES	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations (mean)	1,467,134	1,355,364	1,467,134	1,355,364
R-squared (mean)	0.61		0.52	
Number of firms (mean)		333,452		333,452
First-stage F-stat (mean)		139		139

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: For each road variable in each city, we randomly assign a road value from another city and then rerun the regression. We repeat the process 50 times and obtain bootstrap coefficients and standard errors. Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. GDP = gross domestic product; OLS = ordinary least squares; 2SLS = two-stage least squares.

Table 3.4 Differential impact on state-owned and private enterprises

	(1)	(2)	(3)	(4)
Variable	**log(patents)**		**Having at least one patent**	
	Non-state-owned	**State-owned**	**Non-state-owned**	**State-owned**
log(road)	0.079***	0.013	0.047***	-0.034
	-0.023	-0.088	-0.015	-0.062
Other controls	YES	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	1,284,619	89,774	1,284,619	89,774
Number of firms	318,978	24,066	318,978	24,066
First-stage F-stat	2023	108.6	2023	108.6

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. We measure whether a firm is state owned by two standards. Here, we define a firm as state owned if the state owns at least half of its capital. In the appendix, we compare the share of capital between foreign investors and the state. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4. MECHANISMS AT WORK

Having shown that transportation infrastructure matters to firms' innovative behavior, we now turn to the underlying mechanisms. The literature has highlighted two major channels—market size expansion and knowledge diffusion. Road improvements lower transportation costs, which in turn expands firms' market size (Bougheas, Demetriades, and Morgenroth 1999; Jacoby and Minten 2009). A better-connected road system may also facilitate knowledge diffusion (Agrawal, Galasso, and Oettl 2014; Donaldson and Hornbeck 2013; Zheng and Kahn 2013).

Reduction in Transportation Costs and Market Expansion

Fernald (1999) constructed a vehicle intensity index (average vehicle share) to measure an industry's reliance on transportation infrastructure. We use that index to construct a "vehicle intensity" dummy, which equals one if the average vehicle share in a particular industry is larger than the median level among all the industries. We also use an alternative measure to capture an industry's reliance on transportation infrastructure. Theoretically, an industry with a larger weight per unit of value for its products implies high dependence on transportation (Duranton, Morrow, and Turner 2014). Based on the average weight per unit of value of products at the industrial level, we create a "weight intensity" dummy, being one if it is above the median level and zero otherwise. See Tables A.1 and A.2 in the appendix for more detailed information about these two classifications. As shown in columns 1 through 4 of Table 4.1, firms in industries with intensive vehicle use or heavier weight per unit of product value witness more rapid growth in innovation as roads improve. The results indicate that road improvement spurs firms to innovate probably through the channel of lower transportation costs.

As shown in the literature, lower transportation costs are often associated with greater market size (Jacoby and Minten 2009). Market size (access) has been regarded as one of the more important determinants of innovation (Aghion and Howitt 1992; Acemoglu and Linn 2004). Extensions of the Melitz (2003) model made by Bustos (2011) and Lileeva and Treer (2010) show that firms that enjoy a larger market are more likely to invest in technologies. Larger markets increase firm sales, enabling them to recoup the large fixed investment associated with R&D. We use two ways to test the market-size channel.

First, we examine the differential impact of road improvement on two different types of firm. Firms producing *homogeneous* goods, which tend to have lower markups and thus have to rely more on market expansion to generate profit, are often more responsive to changes in transport costs than those producing *differentiated* goods (Martincus and Blyde 2013). In other words, market size matters more to those producing homogeneous goods. Because market size expands as the cost of transportation drops, firms producing homogeneous goods may benefit more from road improvements. To text this, we divide the sample into two subgroups according to the average value of the GM index, a popular measure of industry-level product heterogeneity.[11] The higher the GM index, the more heterogeneous are the products in an industry (see Table A.3 in the appendix for more information). In columns 5 and 6 of Table 4.1, the interaction term between road density and a dummy indicating heterogeneous firms is included. The interaction term is negative and significant, suggesting that road improvement plays a greater role in facilitating innovations of firms producing homogeneous goods than those producing heterogeneous goods, consistent with the theoretical prediction. The result provides some indirect evidence that roads shape firms' innovation behavior via the channel of market-size expansion.

[11] The GM index, named after Gollop and Monahan (1991), measures the dissimilarity of input portfolios across firms in an industry. Kugler and Verhoogen (2012) aggregated the index to ISIC rev. 2 four-digit industries using a concordance from the Organization for Economic Co-operation and Development. We converted this index to the national industries classification (GBT4754-2002) based on the United Nations matching code. The UN matching code is available from http://unstats.un.org/unsd/cr/registry/regdnld.asp?Lg=1.

Table 4.1 Transportation cost mechanisms

Variable	(1) log(patents)	(2) Having at least one patent	(3) log(patents)	(4) Having at least one patent	(5) log(patents)	(6) Having at least one patent
log(road)	0.074***	0.045***	0.061***	0.037**	0.064**	0.010
	(0.023)	(0.016)	(0.024)	(0.016)	(0.032)	(0.022)
log(road)*"vehicle intensity" dummy	0.010***	0.005***				
	(0.002)	(0.001)				
log(road)*"weight intensity" dummy			0.012***	0.006***		
			(0.001)	(0.001)		
log(road)*"heterogeneous" dummy					-0.004**	-0.003***
					(0.002)	(0.001)
Other controls	YES	YES	YES	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	1,275,995	1,275,995	1,305,167	1,305,167	485,296	485,296
Number of firms	313,055	313,055	321,230	321,230	122,347	122,347
First-stage F-stat	971.5	971.5	985.6	985.6	338.1	338.1

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. The vehicle intensity index (vehicle share index), which is adopted from Fernald (1999), captures an industry's reliance on highways. Vehicle intensity dummy equals one if the vehicle intensity index of that industry is larger than the median level of all industry. The weight intensity index comes from Duranton, Morrow, and Turner (2014) and is defined as the mean weight per output value at the industry level. The "high-weight intensity" dummy is defined as one if the industry's weight intensity index is larger than the median level of all industries. In columns 5 and 6, we divide the sample into two groups by the value of the GM index (Gollop and Monahan 1991), which measures product heterogeneity at the industry level. The higher the GM index, the more heterogeneous the products in an industry are. The dummy equals one if the GM index is larger than its median level, and zero otherwise. See the appendix for more details. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses;*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Second, we examine the impact of road density on export status. A firm that exports goods to the world market has a larger market size than one focusing purely on the domestic market. In column 1 of Table 4.2, we focus on a subsample of firms that did not export at first and use a dummy variable indicating whether a firm exports or not later on as a dependent variable. It is apparent from the table that road access increases a firm's likelihood to export. In column 2, we change the dependent variable to log(sales). Road access increases the sales of firms that did not initially export. However, there is concern that as road conditions change, firms may move elsewhere. For example, firms may relocate to coastal areas to take advantage of export opportunities there. In Table A.4, we restrict our sample to those firms that have never changed their locations throughout the sample period. The results are similar. In a word, market size has expanded associated with road improvement, which in turn may lead to more instances of corporate innovation.

Table 4.2 Market expansion mechanisms

Variable	(1) Export dummy	(2) log(sales)	(3)
	Non-exporting at first	All firms	Non-exporting at first
log(road)	0.064***	0.832***	0.425***
	(0.022)	(0.063)	(0.066)
Other controls	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES
Year FE	YES	YES	YES
Observations	764,663	1,385,972	764,663
Number of firms	187,136	339,086	187,136
First-stage F-stat	2132	2364	2132

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. In column 1, we set the dependent variable as export dummy, and constrain the sample to those export firms at first. In columns 2 and 3, the dependent variables are log(sales). In column 3, the sample is restricted to firms that did not export at first. We also perform a robustness check by using firms that remained in the same place during those years in the appendix. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses;*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Knowledge Diffusion (Star Effect)

Lower transportation costs and better road connections also accelerate the mobility of people and diffusion of knowledge across space, allowing for ideas to cross-fertilize (Moretti 2004; Glaeser and Gottlieb 2009). In this subsection we investigate the channel of knowledge diffusion brought by road improvements.

As uncovered in Agrawal, Galasso, and Oettl (2014) and Agrawal, McHale, and Oettl (2014), innovation "stars" have disproportionately large knowledge spillover effects. We can test whether improved roads expedite the flow of knowledge from the innovation stars.

We first specify patent stars for each industry. The stars include firms above the 99th percentile in patent distribution in a given year, similar to the definition of Agrawal, Galasso, and Oettl (2014). If a city-industry has at least one star firm, we define the star dummy as one, and otherwise as zero. We drop those star firms from the regression sample. Following Agrawal, Galasso, and Oettl (2014), we divide the firms into two groups, one with a star nearby and one without a star nearby, and run separate regressions on the two subsamples. As shown in Table 4.3, firms in cities with innovation stars benefit more from road improvements than those without stars. This suggests that road improvement boosts knowledge spillover from star innovators within a city, leading to more patents.

Table 4.3 Knowledge diffusion (star effect)

Variable	(1)	(2)	(3)	(4)
	log(patents)		Having at least one patent	
	No star nearby	Star nearby	No star nearby	Star nearby
log(road)	-0.015	0.059**	-0.007	0.064***
	(0.024)	(0.024)	(0.023)	(0.022)
Other controls	YES	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	500,141	776,886	500,141	776,886
Number of firms	151,927	210,648	151,927	210,648
First-stage F-stat	564.1	1039	564.1	1039

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: A star firm is defined as an inventor above the 99th percentile in the patenting distribution of that industry in that year. If a city-industry has one or more star firms, we define the star dummy as one, and zero otherwise. We have already dropped those star firms from the sample before we run regressions. According to Agrawal, Galasso, and Oettl (2014), we divide the sample into two groups, one with a star nearby and one without a star nearby. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses;*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Spatial Spillover Effect

The star effect focuses only on knowledge diffusion within a city. Firms may also benefit from road improvements elsewhere (Donaldson and Hornbeck 2013). To capture the spillover effect from neighbors, we include road density in neighboring cities and their interactions with local road density in regressions following Shirley and Winston (2004) and Li and Li (2013).[12] As columns 1 and 2 in Table 4.4 show, road density among the first layer of neighbor cities is highly positive and significant. Moreover, its interaction with local road density is also significantly positive, indicating that road improvement in neighboring cities exerts a positive externality on the innovative activity of local firms. When expanding road density to multiple layers of neighboring cities, the results are robust. See columns 3 and 4 in Table 4.4 for results when considering the second layer of neighboring cities.

[12] Strictly speaking, our specification cannot fully distinguish the market size effect from the spatial spillover effect.

Table 4.4 Spatial spillover effect

Variable	(1) Log(patents+1)	(2) Having at least one patent	(3) Log(patents+1)	(4) Having at least one patent
log(road)	0.011	0.004	0.050***	0.027***
	(0.014)	(0.009)	(0.015)	(0.010)
log neighbor road (1st)	0.092***	0.057***		
	(0.014)	(0.009)		
log(road)*log neighbor road (1st)	0.005***	0.003**		
	(0.002)	(0.001)		
log neighbor road (1st+2nd)			0.262***	0.191***
			(0.063)	(0.043)
log(road)*log neighbor road (1st+2nd)			0.018***	0.013***
			(0.004)	(0.003)
Other controls	YES	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	1,379,979	1,379,979	1,380,041	1,380,041
Number of firms	337,633	337,633	337,640	337,640
First-stage F-stat	1503	1503	502.7	502.7

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. Ln neighbor road (1st) stands for the road density for the first layer of neighbors, and log neighbor road (1st+2nd) means the road density for the first two layers of neighbors. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses;*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5. CONCLUSION AND REMARKS

Using a matched patent database at the firm level and infrastructure information at the city level, we estimate the effect of transportation infrastructure on firm innovation. Our identification strategy exploits the variation in the difficulty of constructing roads—that is, the interaction between slope and export price of road construction machinery—as an instrument variable. We find that road improvement plays a strong role in sparking firm innovation. Road improvement lowers firms' transportation costs and expands their market size, which in turn contributes to greater instances of innovation. In addition, road network expansion creates a spatial externality in knowledge spillover. Firms benefit not only from road improvements in their own cities but also from those in neighboring cities.

Our findings have important policy implications. As remarked by Agrawal, Galasso, and Oettl (2014), when promoting firms' ability to innovate, policymakers need not limit themselves only to targeted R&D subsidies and tax credits, the two most popular industrial policies. We find that a 10 percent increase in local road density is roughly equivalent to a 1.42 percent increase in corporate R&D investment. When designing innovation policy, the role of infrastructure should be included in the toolkit.

Our study is subject to some limitations. The lack of patent citation data in China prevents us from studying the impact of road improvements on the quality of innovation and the web of links among innovations across regions. We leave these as future research topics.

APPENDIX: SUPPLEMENTARY TABLES

Table A.1 Selective industries grouped by average vehicle share

Low average vehicle share		High average vehicle share	
Industry	**Index**	**Industry**	**Index**
Leather products	0.200	Paper products	0.900
Rubber and plastic products	0.200	Oil and nuclear fuel processing	1.000
Textile industry	0.300	Food	1.300
General equipment	0.400	Beverage manufacturing	1.300
Vehicle manufacturing	0.400	Wood processing	1.700
Metal products	0.500	Nonmetal mineral products	2.800

Source: This index is from Fernald (1999) directly.

Notes The vehicle intensity index comes from Fernald (1999) and measures the reliance on highways. We divide the sample into two groups by median value (0.70). The "vehicle intensity" dummy equals one if an industry's vehicle intensity index is larger than the median level of all industries.

Table A.2 Selective industries grouped by weight per value

Low weight per value		High weight per value	
Industry	**Index**	**Industry**	**Index**
Communication equipment, computer	0.010	Wood, bamboo, rattan, etc.	2.170
Textile wearing apparel	0.060	Petroleum, nuclear fuel	2.330
Medicines	0.080	Plastic	8.510
Leather, fur, etc.	0.120	Coal	42.440
Electrical machinery equipment	0.150	Ferrous metal ores	42.440
Textile	0.250	Nonferrous metal ores	42.440
Rubber	0.320	Mining auxiliary	42.440

Source: This index is calculated by authors, based on the relevant index from Duranton, Morrow, and Turner (2014).

Notes: Following Duranton, Morrow, and Turner (2014), we use weight per value to measure the degree of dependence on transportation at the industry level. Industries with large weight per value are naturally more dependent on transportation. We divide the sample into two groups by the median value (0.34).

Table A.3 Homogeneous versus heterogeneous products

Industries with homogeneous products		Industries with heterogeneous products	
Industry	**GM**	**Industry**	**GM**
Oil	0.000	Nuclear fuel	0.540
Balls products	0.150	Forest chemical product	0.540
Edible ice	0.290	Raw material for organic chemistry	0.540
Sugar	0.360	Wood processing	0.560
Nonedible vegetable oil	0.400	Luggage and bags	0.570
MSG products	0.460	Cotton	0.590
Eggs	0.460	Fur tanning and garment processing	0.600
Beverages	0.460	Bamboo products	0.600

Source: This GM index is calculated by authors, based on the index from Kugler and Verhoogen (2012). We converted this index to the national industries classification (GBT4754-2002) based on the United Nations matching code. The UN matching code is available from http://unstats.un.org/unsd/cr/registry/regdnld.asp?Lg=1.

Notes: We divide the sample into two groups by the average value of the GM index (Gollop and Monahan 1991), which measures the dissimilarity of input portfolios across firms in an industry. A larger GM index value in an industry means it has more heterogeneous products. We divide the sample into two groups by the median value (0.53).

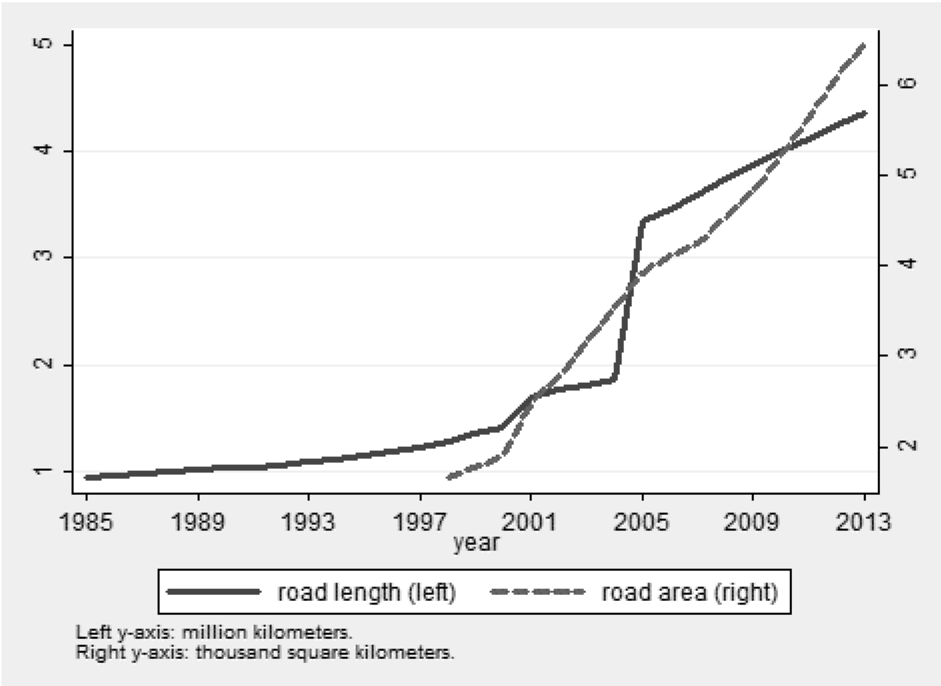
Table A.4 Market expansion mechanism (unchanged city code)

Variable	(1) Export dummy	(2) log(sales)	(3) log(sales)
	Non-exporting at first	All	Non-exporting at first
log(road)	0.065***	0.843***	0.432***
	(0.023)	(0.063)	(0.067)
Other controls	YES	YES	YES
Firm fixed effects (FE)	YES	YES	YES
Year FE	YES	YES	YES
Observations	763,770	1,384,771	763,770
Number of firms	186,928	338,801	186,928
First-stage F-stat	2218	2449	2218

Source: The firm-level dataset is the Annual Survey of Industrial Enterprises (NBSC, various years [a]). Innovation variables come from the Chinese Patent Database (SIPO, various years). The road and GDP information come from the China City Statistics Yearbook (NBSC, various years [b]).

Notes: Other controls include age group fixed effects, specialization index for each two-digit industry, log GDP per capita, and competition level and its square form. In this table, we constrain the sample to firms that did not change locations during the sample period. All specifications are estimated by two-stage least squares. Robust standard errors in parentheses;*** p < 0.01, ** p < 0.05, * p < 0.1.

Figure A.1 Road area and road length



Source: The road information comes from the China City Statistics Yearbook (NBSC, various years [b]).

REFERENCES

Acemoglu, D., and J. Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3): 1049–1090.

Acemoglu, D., J. Moscona, and J. A. Robinson. 2016. *State Capacity and American Technology: Evidence from the 19th Century.* NBER Working Paper 21932. Cambridge, MA, US: National Bureau of Economic Research.

Aghion, P., and P. Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60 (2): 323–351.

Aghion, P., J. Van Reenen, and L. Zingales. 2013. "Innovation and Institutional Ownership." *American Economic Review* 103 (1): 277–304.

Agrawal, A., I. Cockburn, A. Galasso, and A. Oettl. 2014. "Why Are Some Regions More Innovative Than Others? The Role of Small Firms in the Presence of Large Labs." *Journal of Urban Economics* 81: 149–165.

Agrawal, A., A. Galasso, and A. Oettl. 2014. *Roads and Innovation.* Rotman School of Management Working Paper 2478752. Toronto: University of Toronto.

Agrawal, A. K., J. McHale, and A. Oettl. 2014. *Why Stars Matter.* NBER Working Paper 20012. Cambridge, MA, US: National Bureau of Economic Research.

Angrist, J. D., and A. B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15 (4): 69–85.

Aschauer, D. A. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23 (2): 177–200.

Banerjee, A., E. Duflo, and N. Qian. 2012. *On the Road: Access to Transportation Infrastructure and Economic Growth in China.* NBER Working Paper 17897. Cambridge, MA, US: National Bureau of Economic Research.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang. 2012. *Roads, Railroads, and Decentralization of Chinese Cities.* Working paper. Accessed December 2015. http://www.lse.ac.uk/geographyAndEnvironment/whosWho/profiles/henderson/RoadsRailroadsDecentralization.pdf.

Bloom, N., M. Draca, and J. Van Reenen. 2011. *Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT, and Productivity.* NBER Working Paper 16717. Cambridge, MA, US: National Bureau of Economic Research.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81 (4): 1347–1393.

Bougheas, S., P. O. Demetriades, and E. L. Morgenroth. 1999. "Infrastructure, Transport Costs, and Trade." *Journal of International Economics* 47 (1): 169–189.

Bustos, P. 2011. "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of Mercosur on Argentinian Firms." *American Economic Review* 101 (1): 304–340.

Cameron, A. C., and P. K. Trivedi. 2013. *Regression Analysis of Count Data,* 2nd ed. New York: Cambridge University Press.

Choi, S. B., S. H. Lee, and C. Williams. 2011. "Ownership and Firm Innovation in a Transition Economy: Evidence from China." *Research Policy* 40 (3): 441–452.

Dang, J., and K. Motohashi. 2015. "Patent Statistics: A Good Indicator for Innovation in China? Patent Subsidy Program Impacts on Patent Quality." *China Economic Review* 35: 137–155.

Donaldson, D. 2010. *Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.* NBER Working Paper 16487. Cambridge, MA, US: National Bureau of Economic Research.

Donaldson, D., and R. Hornbeck. 2013. *Railroads and American Economic Growth: A "Market Access" Approach.* NBER Working Paper 19213. Cambridge, MA, US: National Bureau of Economic Research.

Duflo, E., and R. Pande. 2007. "Dams." *Quarterly Journal of Economics* 122 (2): 601–646.

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

Duranton, G., and M. Turner. 2012. "Urban Growth and Transport." *Review of Economic Studies* 79: 1407–1440.

Ellison, G., and E. L. Glaeser. 1997. "Geographic Concentration in US Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy* 105 (5): 889–927.

Fernald, J. G. 1999. "Roads to Prosperity? Assessing the Link between Public Capital and Productivity." *American Economic Review* 89 (3): 619–638.

Glaeser, E. L., and J. D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983–1028.

Gollop, F. M., and J. L. Monahan. 1991. "A Generalized Index of Diversification: Trends in US Manufacturing." *Review of Economics and Statistics* 73: 318–330.

Hu, A. G. 2010. "Propensity to Patent, Competition, and China's Foreign Patenting Surge." *Research Policy* 39 (7): 985–993.

Hu, A. G., and G. H. Jefferson. 2009. "A Great Wall of Patents: What Is behind China's Recent Patent Explosion?" *Journal of Development Economics* 90 (1): 57–68.

Jacoby, H. G., and B. Minten. 2009. "On Measuring the Benefits of Lower Transport Costs." *Journal of Development Economics* 89 (1): 28–38.

Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–598.

Krugman, P. R. 1991. *Geography and Trade.* Cambridge, MA, US: MIT Press.

Kugler, M., and E. Verhoogen. 2012. "Prices, Plant Size, and Product Quality." *Review of Economic Studies* 79 (1): 307–339.

Li, H., and Z. Li. 2013. "Road Investments and Inventory Reduction: Firm Level Evidence from China." *Journal of Urban Economics* 76: 43–52.

Li, X. 2012. "Behind the Recent Surge of Chinese Patenting: An Institutional View." *Research Policy* 41 (1): 236–249.

Lileeva, A., and D. Treer. 2010. "Improved Access to Foreign Markets Raises Plant-Level Productivity . . . for Some Plants." *Quarterly Journal of Economics* 125 (3): 1051–1099.

Martincus, C. V., and J. Blyde. 2013. "Shaky Roads and Trembling Exports: Assessing the Trade Effects of Domestic Infrastructure Using a Natural Experiment." *Journal of International Economics* 90 (1): 148–161.

Melitz, M. J. 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.

Moretti, E. 2004. "Human Capital Externalities in Cities." In *Handbook of Regional and Urban Economics,* vol. 4, edited by J. V. Henderson and J.-F. Thisse, 2243–2291. Amsterdam: Elsevier.

NBSC (National Bureau of Statistics of China). Various years. Annual Survey of Industrial Enterprises in China dataset. Beijing.

———. Various years. *China City Statistical Yearbook.* Road Data. Beijing: China Statistics Press.

Romer, P. M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–S102.

Romp, W., and J. De Haan. 2007. "Public Capital and Economic Growth: A Critical Survey." *Perspektiven der Wirtschaftspolitik* 8 (S1): 6–52.

Saiz, A. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125 (3): 1253–1296.

Shirley, C., and C. Winston. 2004. "Firm Inventory Behavior and the Returns from Highway Infrastructure Investments." *Journal of Urban Economics* 55 (2): 398–415.

SIPO, PRC (State Intellectual Property Office of the People's Republic of China). Various years. Chinese Patent database. Beijing.

Thoma, G. 2013. "Quality and Value of Chinese Patenting: An International Perspective." *Seoul Journal of Economics* 26: 33–72.

World Bank. 1994. *The World Bank Annual Report 1994.* Washington, DC.

Xie, Z., and X. Zhang. 2015. "The Patterns of Patents in China." *China Economic Journal* 8 (2): 122–142.

Zheng, S., and M. E. Kahn. 2013. "China's Bullet Trains Facilitate Market Integration and Mitigate the Cost of Megacity Growth." *Proceedings of the National Academy of Sciences* 110 (14): E1248–E1253.

RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to www.ifpri.org/pubs/pubs.htm#dp.
All discussion papers can be downloaded free of charge.

1541. *A systematic review of cross-country data initiatives on agricultural public expenditures in developing countries.* Richard Anson and Tewodaj Mogues, 2016.

1540. *Does female labor scarcity encourage innovation?: Evidence from China's gender imbalance.* Zhibo Tan and Xiaobo Zhang, 2016.

1539. *Linking smallholder farmers to commercial markets: Evidence from nongovernmental organization training in Nicaragua*. Ayako Ebata and Manuel A. Hernandez, 2016.

1538. *Can labor market imperfections explain changes in the inverse farm size–productivity relationship?: Longitudinal evidence from rural India.* Klaus Deininger, Songqing Jin, Yanyan Liu, and Sudhir K. Singh, 2016.

1537. *Labor adaptation to climate variability in eastern Africa.* Xiaoya Dou, Clark Gray, Valerie Mueller, and Glenn Sheriff, 2016.

1536. *A dynamic spatial model of agricultural price transmission: Evidence from the Niger millet market.* Anatole Goundan and Mahamadou Roufahi Tankari, 2016.

1535. *Qualitative methods for gender research in agricultural development*. Deborah Rubin, 2016.

1534. *Anchoring bias in recall data: Evidence from Central America*. Susan Godlonton, Manuel A. Hernandez, and Mike Murphy, 2016.

1533. *Contracting by small farmers in commodities with export potential: Assessing farm profits of lentil growers in Nepal.* Anjani Kumar, Devesh Roy, Gaurav Tripathi, P. K. Joshi, and Rajendra P. Adhikari, 2016.

1532. *Rent dispersion in the US agricultural insurance industry.* Vincent Smith, Joseph Glauber, and Robert Dismukes, 2016.

1531. *Long-term drivers of food and nutrition security.* David Laborde, Fahd Majeed, Simla Tokgoz, and Maximo Torero, 2016.

1530. *Understanding compliance in programs promoting conservation agriculture: Modeling a case study in Malawi.* Patrick S. Ward, Andrew R. Bell, Klaus Droppelmann, and Tim Benton, 2016.

1529. *A model of reporting and controlling outbreaks by public health agencies.* Alexander E. Saak and David A. Hennessy, 2016.

1528. *Boserupian pressure and agricultural mechanization in modern Ghana.* Frances Cossar, 2016.

1527. *Agricultural mechanization and agricultural transformation.* Xinshen Diao, Jed Silver, and Hiroyuki Takeshima, 2016.

1526. *Delegation of quality control in value chains*. Alexander E. Saak, 2016.

1525. *Structural transformation and intertemporal evolution of real wages, machine use, and farm size–productivity relationships in Vietnam.* Yanyan Liu, William Violette, and Christopher B. Barrett, 2016.

1524. *Can contract farming increase farmers' income and enhance adoption of food safety practices?: Evidence from remote areas of Nepal.* Anjani Kumar, Devesh Roy, Gaurav Tripathi, P. K. Joshi, and Rajendra P. Adhikari, 2016.

1523. *Effectiveness of food subsidies in raising healthy food consumption: Public distribution of pulses in India.* Suman Chakrabarti, Avinash Kishore, and Devesh Roy, 2016.

1522. *Findings across agricultural public expenditure reviews in African countries*. Stephen D. Mink, 2016.

1521. *Risk and sustainable crop intensification: The case of smallholder rice and potato farmers in Uganda.* Bjorn Van Campenhout, Emmanuel Bizimungu, and Dorothy Birungi, 2016.

1520. *Varietal integrity, damage abatement, and productivity: Evidence from the cultivation of Bt cotton in Pakistan.* Xingliang Ma, Melinda Smale, David J. Spielman, Patricia Zambrano, Hina Nazli, and Fatima Zaidi, 2016.

1519. *Institutional arrangements to make public spending responsive to the poor—(where) have they worked?: Review of the evidence on four major intervention types.* Tewodaj Mogues and Alvina Erman, 2016.

1518. *A poverty-sensitive scorecard to prioritize lending and grant allocation: Evidence from Central America.* Manuel A. Hernandez and Maximo Torero, 2016.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org