On the reverse causality between output and infrastructure: The case of China

Qu Feng *, Guiying Laura Wu

School of Social Sciences, Nanyang Technological University, Singapore

ARTICLE INFO

JEL Classification:
C23
E22
H54
O40
O50

Keywords:
Infrastructure
Productivity
Chinese economy

ABSTRACT

After the 2008 global financial crisis, promoting public infrastructure investment as a growth engine has been revived by economists. China has been considered as such a successful example of enhancing economic growth by massive infrastructure investments in the past decades. However, the literature has provided conflicting empirical results on the productivity effect of public infrastructure using aggregate data, mainly due to reverse causality. Thus, the estimated productivity effect could be either upward or downward biased. In this paper we rely on the institutional background of infrastructure investment in China, and explore several alternative ways to mitigate the reverse causality. Using China’s provincial-level data over 1996–2015 and within the framework of an aggregate production function estimation, we find that an upward bias dominates when estimating output elasticity of public infrastructure, and that weak evidence is found on the productivity effect of public infrastructure. This finding highlights the necessity of using alternative identification strategies or data types.

1. Introduction

After the 2008 global financial crisis, promoting public infrastructure investment as a growth engine has been revived by economists and policy makers. For example, a 4 trillion Chinese Yuan (equivalent to 600 billion US dollars) fiscal stimulus package was introduced by the Chinese government to invest mainly in the infrastructure in its western provinces in 2008 (Ouyang and Peng, 2015). Recently, as Chinese economy started to slow down in 2015, 1 trillion Chinese Yuan was further proposed to invest in infrastructure (Financial Times, August 5, 2015).

For a specific project on infrastructure investment, e.g., building an airport, it is straightforward to calculate its economic return if the benefits and costs of the project are well defined and recorded. However, its social return may not be fully captured in a financial evaluation framework. For a specific type of infrastructure, the literature has also developed various ways to identify its productivity effect, for example, Fernald (1999) for road in the US, Röller and Waverman (2001) for telecommunications infrastructure in OECD countries, and the recent works surveyed in Redding and Turner (2015) for transport infrastructure. In China, rates of return to railroad and road are found over 10% and 20%, respectively (Li and Li, 2013; Li and Chen, 2013).

To address whether public infrastructure investment as a whole enhances the growth of the whole economy, we take a macro view and focus on the productivity and return of the total public infrastructure investment. For this purpose, following the literature starting from Aschauer (1989), we estimate the output elasticity with respect to public infrastructure in an aggregate production function using China’s provincial panel data over 1996–2015.

The importance of studying China’s case is in two folds. First, it is well known that China is considered as an investment-driven economy with the investment-to-GDP ratio above 45% since 2009, far exceeding other developing countries and advanced economies. As a major component of the total investment, public infrastructure investment...
accounts for an average rate of 9.3% of China’s GDP during 1996-
2015. Thus, it is of policy significance to evaluate the productivity and
return of public infrastructure investment in China. Second, China’s
institutional framework may provide unique identification strategies for
the endogeneity problem due to the reverse causality between output and
public infrastructure when estimating its elasticity.

Using the framework of an aggregate production function estimation,
the literature has provided conflicting empirical results, mainly
due to reverse causality. As surveyed in Bom and Ligthart (2014),
the output elasticity of public capital varies from the highest estimate of
2.04 for Australia in one research to the lowest one of −1.7 for New
Zealand in another research. In between, many estimates are statisti-
cally not different from zero. The output elasticity of public infras-
tructure capital could be overestimated when a growth in output facilitates
an increase in public infrastructure investment. That is, public infras-
tructure investment could be induced by economic growth, instead of
driving economic growth. Alternatively, the output elasticity of pub-
lic infrastructure capital could be underestimated when public infras-
tructure investment is used as a countercyclical tool to boost economic
growth during economic recession.

In a recent study with a focus on the investment efficiency in China,
Shi and Huang (2014) argue that a downward bias is more likely in
China’s case. This is because the Chinese government tends to use
infrastructure investment as a choice for stimulating its economy when
a negative productivity shock is expected. Consistent with this logic,
they find that the output elasticity using a proxy approach developed by
Ackerberg et al. (2015) is even larger than that from the OLS approach.

Using China’s provincial panels over 1995–2011, they obtain a big and
positive output elasticity of public infrastructure, with a magnitude
around 0.22 to 0.29. This implies a rate of return more than 50%.3

In this paper we rely on the institutional background of infrastruc-
ture investment in China, and explore several alternative ways to miti-
gate the reverse causality between aggregate output and public infras-
tructure. Using different approaches we find that an upward bias dom-
inates when estimating output elasticity of public infrastructure using
China’s provincial-level data over 1996−2015. Within the framework of
an aggregate production function estimation, weak evidence is found
on the productivity effect of public infrastructure in China. This finding
suggests the necessity of using alternative identification strategies or
data types, e.g., a disaggregation approach using firm-level data, such
as Fisher-Vanden et al. (2015); Li et al. (2017); and Wu et al. (2017).

The rest of the paper is organized as follows. Section 2 introduces a
macroeconomic model using an aggregate production function, aug-
mented with public infrastructure capital. Various strategies of dealing
with the reverse causality are discussed in Section 3. Section 4 presents
the data and reports the empirical findings. Section 5 concludes.

2. Empirical model

To model the general idea that public infrastructure investment pro-
motes economic growth, following literature we introduce an aggregate
production function:

\[ Y = AK^{\gamma_L}L^{\gamma_L}, \]

where \( Y \) is the total output; \( L \) is the total labor force; and \( K \) is
the stock of non-infrastructure capital. The public infrastructure capital \( B \),
measuring the stock of public infrastructure investment, enters the pro-
duction function as a contributing component to the total productivity
factor (TFP) \( A \), i.e., \( A = A_0B^{\delta_B} \), where \( A_0 \) is the component of TFP
that is unrelated to public infrastructure. Thus, the aggregate production
function becomes

\[ Y = A_0B^{\delta_B}K^{\gamma_K}L^{\gamma_L}. \]

The stock variables, \( B \) and \( K \), accumulate according to the following
laws of motion:

\[ B_t = (1 - \delta_B)B_{t-1} + G_t \]

and

\[ K_t = (1 - \delta_K)K_{t-1} + I_t. \]

Here \( G \) measures the infrastructure investment in industries with exten-
nalities, such as electricity, gas, water, transport, information transmis-
sion, and \( I \) is the investment in non-infrastructure sectors. \( \delta_B \) and \( \delta_K \)
are depreciation rates of \( B \) and \( K \), respectively.

Under the assumption of constant returns to scale (CRS),
\( \gamma_K + \gamma_L + \gamma_I = 1 \), so that (1) becomes
\( Y/L = A_0(B/L)^{\delta_B}(K/L)^{\gamma_K}. \) Thus, the aggregate production function in the intensive form can be written as

\[ y = y_0 + \gamma_K b + \gamma_K k. \]

where \( y = \log(Y/L), b = \log(B/L), k = \log(K/L) \) and \( y_0 = \log(A_0). \) In
this equation, \( \gamma_K \) and \( \gamma_K \) are the output of elasticities of public infras-
tructure and non-infrastructure capital. The economic return of public
infrastructure, or the marginal output of public infrastructure, can be
measured as

\[ \partial Y/\partial B = \gamma_K Y/B. \]

To estimate the coefficients \( \gamma_K \), a panel data model based on the
aggregate production function above is used

\[ y_a = y_0 + \gamma_K b_a + \gamma_K k_a + \mu_t + T_i + \epsilon_a, \]

where \( y_a \) is the logarithm of GDP per labor in province \( a \) in year \( t \),
and \( b_a, k_a, \mu_t \) and \( T_i \) are the logarithm of public infrastructure stock per labor, and \( \epsilon_a \) is
the logarithm of non-infrastructure capital stock per labor. \( \mu_t \) denotes
province specific factors, such as different land area, location, weather,
endowments of raw materials and myriad other factors. Time effects \( T_i \)
can be used to control for national-level macro shocks, including busi-
ness cycles and counter-cyclical policies. \( \epsilon_a \) denotes idiosyncratic shocks
or measurement error in output. To deal with the non-stationarity in
macroeconomic variables, first-differencing Eq. (4) gives our estimat-
ing equation:

\[ \Delta y_a = \gamma_K \Delta b_a + \gamma_K \Delta k_a + \Delta T_i + \Delta \epsilon_a. \]
3. Dealing with reverse causality

When we write down Eq. (4) or (5), our aim is to identify the causal effect of public infrastructure on output. However, as pointed out, e.g., by Gramlich (1994), the causality could go from output to public infrastructure. Higher output may mean greater demand for the services from public infrastructure; higher output may also mean more income for expenditure on public infrastructure. Hence, a positive estimated elasticity could be mainly driven by this reverse causality. Thus, the OLS estimator of $\gamma_i$ in (5) (i.e., the first difference (FD) estimator of (4)) could be biased upward. Alternatively, in the literature as summarized by Borrion Lighthart (2014), due to the Keynesian multiplier effect, public infrastructure investment is often used to boost economic growth during the period of economic recession. In this case, output and public infrastructure investment could be negatively correlated. Thus, the OLS estimator of $\gamma_i$ in (5) (i.e., the first difference (FD) estimator of (4)) could be biased downward.

In the literature, there are several ways to deal with this endogeneity issue due to reverse causality. The first and general approach is the instrumental variable (IV) estimation, e.g., Holtz-Eakin (1994), Baltagi and Pinnoi (1995) and the more recent literature surveyed in Redding and Turner (2015). An alternative way to address the reverse causality is the simultaneous-equations approach, explicitly modeling the relationship between $y$ and $b$ in an additional equation, such as Röller and Waverman (2001) and Cadot et al. (2006). Another approach is to explore the heterogeneity of output effect from disaggregated data. A leading example is Fernald (1999). Recently, Calderon et al. (2015) use a panel cointegration approach to deal with the nonstationarity and establish only one cointegrating relation to address concerns with reverse causality in a panel data set with a long span of time periods.

In the Chinese context, Shi and Huang (2014) claim that the reverse causality could lead to a negative correlation between output and public infrastructure since “Chinese government tends to use infrastructure investment as a choice for reviving its economy when it expects a large negative TFP shock”, which will bias downward the estimated output elasticity of infrastructure. In their paper, the endogeneity due to reverse causality is interpreted as the negative correlation between $\Delta b_t$ and $\Delta y_t$, where this correlation is dealt with by the proxy approach developed by Ackerberg et al. (2015).

Different from Shi and Huang (2014), we argue that regarding the feedback effect of output on public infrastructure, a positive correlation is more likely to dominate in the case of China. Bai and Qian (2010) provide an interesting survey on the specific institutional background for infrastructure investment in China. Two stylized facts can be summarized from the survey. First, most infrastructure investment are made by state-owned enterprises with funds from both the central and the local governments. Second, among various jurisdiction levels, the provincial governments play a key role in infrastructure investment decision. Wu et al. (2017) survey several hypotheses on the investment incentives of the Chinese governments that have been discussed in the literature. In short, for the central government, first, infrastructure development is needed to fight against the worsening regional inequality by promoting the catch-up of lagging inland provinces with coastal provinces. This would imply a negative correlation between $b_t$ and $\mu_t$ in Eq. (4) and can be eliminated by first differenting as in Eq. (5). Second, infrastructure development is necessary to support the rapid economic growth of the country that fuels an ever-increasing demand for infrastructure services. This would imply a positive correlation between $\Delta b_t$ and $\Delta y_t$ in Eq. (5). Finally, for the provincial governments, under China’s regionally decentralized authoritarian system, infrastructure investment has been adopted as the most effective instrument by the local governments as their response the GDP yardstick competition. Hence a province with better growth prospects could expect to produce higher output and collect more fiscal revenue in the future, which in turn may allow the province to invest more in current infrastructure via various financing schemes. This would also imply a positive correlation between $\Delta b_t$ and $\Delta y_t$ in Eq. (5).

It is a well-known fact that the 30 provinces in China are at different levels of economic development, varying substantially in GDP per capita, public facilities and fiscal budget (Naughton, 2007). Hence, over a relative long span of time, such positive correlation generated by financing abilities cross provinces could overpower the negative correlation between output and public infrastructure due to the short-run countercyclical story or national policies to reduce regional disparity. Therefore, after including time effects in Eq. (5) to mitigate the effect of national-level countercyclical policies, we conject that the upward bias due to the reverse causality is more likely when estimating output elasticity of public infrastructure $\gamma_i$ in (5).

In this paper, we employ several ways to deal with or mitigate the endogeneity issue due to reverse causality. The first approach is to use an alternative measure of investment in fixed assets reported by the National Bureau of Statistics of China (NBS): Newly Increased Fixed Assets (NIFA hereafter) (xingzang guding zichan touzi in Chinese). Different from the usual measure of investment to construct public infrastructure capital and non-infrastructure capital in (4), Total Investment in Fixed Assets (TIFA hereafter) (quanshehui guding zichan touzi in Chinese), which measures total cost spent on constructing and purchasing fixed assets, NIFA measures investment in fixed assets that have been used for production after the process of construction and purchase is completed. Due to the time to build, NIFA is less likely to be affected by the current output. Thus, the reverse causality between output and public infrastructure (or non-infrastructure) capital is mitigated.

We also make use of a measure of $b_t$ in the level Eq. (4) (or $\Delta b_t$ in the differenced Eq. (5)) that is less likely to be affected by $y_t$ (or $\Delta y_t$). A natural candidate in the literature is the lagged value of $b_t$ (or $\Delta b_t$). Different from $b_t$ (or $\Delta b_t$), $b_{t-1}$ (or $\Delta b_{t-1}$) is less likely to be affected by $y_t$ (or $\Delta y_t$ under the assumption that the current output only affects the current and future, instead of the past, values of public infrastructure. As a stock variable accumulating all past public infrastructure investments, $b_{t-1}$ still provides service to future production.

As a general approach to deal with endogeneity, instrumental variable estimation is also used to consistently estimate $y_t$. In this paper, three different sets of instruments are explored. First, as in Holtz-Eakin (1994), twice-lagged variables $\Delta b_{t-2}$ and $\Delta b_{t-3}$ are employed as internal instruments for $\Delta b_t$ and $\Delta y_t$ in Eq. (5). Second, as widely documented in the literature one of distinctive institutional features of China’s economic miracle is that under the so-called “GDP tournament” scheme local governments have been playing an active role in promoting economic growth, including investing in infrastructure (Li and Zhou, 2005; Jin et al., 2005; Wang et al., 2017). Under this scheme, local governments compete with each other on GDP growth, and their investment behavior could affect each other. Thus, $\Delta b_t$ in neighboring

---

5. When infrastructure investment is used to reduce regional inequality at the growth of output, instead of the level of output, $\Delta b_t$ and $\Delta y_t$ could be negatively correlated, as in Shi and Huang (2014).


7. NIFA is not a formal measure of investment reported by NBS. It is reported to show the extent of how investment process in fixed assets has been completed in some years and some sectors. Since the data on NIFA are not available before 2002, TIFA is used as a formal measure of investment throughout the paper. We construct the data of NIFA before 2002 by using the components of basic construction and renovations of NIFA and their ratios in provinces and industries in China Statistics Yearbooks.

8. $\Delta b_{t-1}$ and $\Delta b_{t-2}$ could be correlated with $\Delta y_{t-1}$. It is worth noting that this IV approach is different from using $\Delta b_{t-1}$ and $\Delta b_{t-2}$ as regressors in the FD regression above.
provinces, denoted as $\Delta n_{it}$, can serve as an instrument for $\Delta h_{it}$. A recent study by Zheng et al. (2015) finds that infrastructure spending in a province is positively correlated with infrastructure spending in its neighboring provinces. In addition, since $\Delta y_{it}$ is only affected by $\Delta h_{it}$ and $\Delta k_{it}$ conditional on time dummies in Eq. (5), instruments of $\Delta h_{it-2}$ and $\Delta n_{it}$ have no direct effect on $\Delta y_{it}$. They affect $\Delta y_{it}$ only through $\Delta h_{it}$.

Third, we use the ages of provincial governors and party leaders as external instruments for public infrastructure in (5). In China’s current political system, provincial governors and party leaders retire at an age of 65 if they are not promoted to top-level officials in Chinese central government. Given that GDP growth is the most important key performance indicator and that investment is one of the major contributing factors to GDP growth, provincial governors and party leaders are less motivated to invest when their ages are closer to 65. In this case, the ages of provincial governors and party leaders could be negatively correlated with public infrastructure investment. In terms of exclusion restriction, like instruments of twice-lagged variables and neighboring public infrastructure, the ages of provincial governors and party leaders are considered to be irrelevant to output (or growth) in the aggregate production function (4) (or (5)).

The empirical results using the identification strategies above are reported in Section 4 below. Using a Chinese provincial panel data set during 1996–2015, we show that after dealing with the endogeneity issue due to reverse causality, the estimated output elasticities are notably smaller than the FD estimates, suggesting that an upward bias due to reverse causality is prevalent in China’s case.

### 4. Data and empirical results

Data on GDP ($y$) are obtained from the website of National Bureau of Statistics of China. We collect data for 30 provinces excluding Tibet over years 1996–2015. As in Shi and Huang (2014), the size of labor force ($L$) is calculated by the number of residents multiplied by the ratio of age cohort of 16–65. For the key variables public infrastructure investment ($G$) and non-infrastructure investment ($I$), we collect data on the total investment in fixed assets (TIFA) from Statistical Yearbooks of The Chinese Investment in Fixed Assets and China Statistical Yearbooks. These two series of statistics yearbooks report total investment in fixed assets by industry and by province. Infrastructure investment $G$ is measured by the sum of investments in the 3 industries: (1) production and supply of electricity, gas and water; (2) transport, storage and post; (3) information transmission, computer services and software. $I$ is defined as total investment minus $G$. Stock variables $B$ and $K$ are constructed as in (2) and (3) using depreciation rates $\delta_B = \delta_K = 10\%$. Table 1 reports the summary statistics for the variables used in the analysis. GDP, public infrastructure investment, non-infrastructure investment are deflated by the province-specific price indices of investment in fixed assets. The unit, mean and standard deviation for the real output per labor, real public infrastructure and non-infrastructure capital stocks per labor and other variables before taking logarithms are reported. These variables are used in the log form in regressions, so that the corresponding coefficients can be interpreted as elasticities.

We first report estimation results on elasticities $\gamma_B$ and $\gamma_K$ without dealing with reverse causality. Column (1) of Table 2 reports fixed effects (FE) estimates of $\gamma_B$ and $\gamma_K$ which are 0.057 and 0.303, respectively. To eliminate unit roots and common trends in the macro data, first-differencing is needed. Column (2) presents FD estimates, showing that the estimated elasticity of public infrastructure capital is 0.127 and significant at 1% level. Considering that the return of public infras-

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Summary statistics of variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>$y$</td>
<td>real output per labor</td>
</tr>
<tr>
<td>$b$</td>
<td>real infrastructure capital per labor</td>
</tr>
<tr>
<td>$k$</td>
<td>real non-infrastructure capital per labor</td>
</tr>
<tr>
<td>$n_{bit}$</td>
<td>real infrastructure capital per labor based on NIFA</td>
</tr>
<tr>
<td>$nb$</td>
<td>real infrastructure capital per labor in neighboring provinces</td>
</tr>
<tr>
<td>$G$</td>
<td>infrastructure investment flow</td>
</tr>
<tr>
<td>$L$</td>
<td>number of labor force</td>
</tr>
<tr>
<td>age 1</td>
<td>age of provincial governor</td>
</tr>
<tr>
<td>age 2</td>
<td>age of provincial party leader</td>
</tr>
</tbody>
</table>

For examples, the neighboring provinces of Shanghai are Jiangsu and Zhejiang, and Jiangxi’s neighbors are Zhejiang, Anhui, Hubei, Hunan, Fujian and Guangdong provinces. $nb_{it}$ is defined as the log (sum of infrastructure stock in neighboring provinces/sum of labor in neighboring provinces). The instrument used is its first difference.

9 We define a province as a neighboring province of $i$ if it shares common border of province $i$. For examples, the neighboring provinces of Shanghai are Jiangsu and Zhejiang, and Jiangxi's neighbors are Zhejiang, Anhui, Hubei, Hunan, Fujian and Guangdong provinces. $nb_{it}$ is defined as the log (sum of infrastructure stock in neighboring provinces/sum of labor in neighboring provinces). The instrument used is its first difference.

10 A similar argument can be seen in Li and Zhou (2005), Wang et al. (2017), in which age is an important factor for the career concerns of provincial leaders.

11 The definition of $G$ here is consistent with the description of physical infrastructure in Fig. 14.3 of Naughton (2007) for China. and the literature in general, e.g., Calderon et al. (2015), Shi and Huang (2014) also include investment in management of water conservancy, environment, and public facilities as part of public infrastructure investment. When we broaden the definition of infrastructure as in Shi and Huang (2014) in robustness checks, we obtain similar findings as in our benchmark results.

12 The choice of depreciation rate in the literature typically varies between 3% and 16%. Thus we set 10% as our benchmark depreciation rate and conduct robustness checks using other rates as alternatives. The main finding of our empirical exercise turns out to be not sensitive to the depreciation rate. To implement the perpetual inventory method, one has to start with an initial value for $B_0$ and $K_0$. In our application, we assume that $B_{1996} = G_{1996}/(\delta_g + g)$ and $K_{1996} = L_{1996}/(\delta_K + g)$, where $g = 10\%$, the average long-run growth rate during our sample period. This assumption is based on the property of a balanced-growth-path model, in which new investment is made to compensate depreciation and guarantee a constant growth in capital stock.

13 According to The Chinese Statistical Yearbook, the investment in fixed assets consists of three components, namely the investment in construction and installation, the investment in purchases of equipment and instrument, and the investment in other items. Price indices of investment in fixed assets are calculated as the weighted arithmetic mean of the price indices of the three components of investment in fixed assets. Under our definition, both infrastructure and non-infrastructure investment contain investment in all three components. Without knowing the exact proportion of each component, we apply the price indices of investment in fixed assets to both infrastructure and non-infrastructure investment.

14 Clustered standard errors are reported in parenthesis below estimates, adjusted for 30 clusters in province.
structure capital is \( \frac{\partial Y}{\partial \delta} = \gamma_y \frac{Y}{B} \) and \( Y/B = 2.254 \) averaging over 1996–2015 for depreciation rates \( \delta_k = \delta_y = 10\% \) in the sample, this elasticity indicates a return rate of 28.6\%. This means that investment in public sectors is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-

capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-
capital is very productive and profitable. To examine the change of return over time, FD estimates using subsamples are also reported in columns (3) and (4), 0.144 and 0.088 for periods of 1996–2007 and 2008–2015, respectively. This implies rates of return over time, FD estimates using subsamples are also reported to public infrastru-

### Table 2

**Output elasticities: Fixed-effects and first-differenced estimates.**

<table>
<thead>
<tr>
<th>Independent variables: Output per labor</th>
<th>FE (1)</th>
<th>FD (2)</th>
<th>FDnew (3)</th>
<th>FDlag (4)</th>
<th>FE (5)</th>
<th>FD (6)</th>
<th>FDnew (7)</th>
<th>FDlag (8)</th>
<th>FDnew (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure capital per labor</td>
<td>0.057</td>
<td>0.127***</td>
<td>0.144***</td>
<td>0.088**</td>
<td>0.037*</td>
<td>0.033</td>
<td>0.035</td>
<td>0.005</td>
<td>-0.030</td>
</tr>
<tr>
<td>Periods</td>
<td>(0.071)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Non-infrastructure capital per labor</td>
<td>0.303***</td>
<td>0.324***</td>
<td>0.340***</td>
<td>0.315***</td>
<td>0.228***</td>
<td>0.250***</td>
<td>0.210***</td>
<td>0.215***</td>
<td>0.402***</td>
</tr>
<tr>
<td>Periods</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.03</td>
</tr>
<tr>
<td>Overall R2</td>
<td>0.84</td>
<td>0.73</td>
<td>0.76</td>
<td>0.67</td>
<td>0.60</td>
<td>0.63</td>
<td>0.54</td>
<td>0.45</td>
<td>0.71</td>
</tr>
<tr>
<td>No. of observations</td>
<td>599</td>
<td>569</td>
<td>329</td>
<td>240</td>
<td>569</td>
<td>329</td>
<td>240</td>
<td>539</td>
<td>539</td>
</tr>
</tbody>
</table>

Notes.

1. FE and FD in columns (1)–(4) denote fixed effects regression and first difference regression, respectively.
2. FDnew in columns (5)–(7) refer to the first-difference estimates using newly increased fixed asset investment.
3. FDlag in columns (8) refer to the first difference estimates using the lags of both public infrastructure and non-infrastructure capital. In column (9) only the lagged value of public infrastructure capital is used.
4. Standard errors are reported in parentheses. The stars, *, ** and *** indicate the significance level at 10%, 5% and 1%, respectively.
5. Standard errors are adjusted for 30 clusters in province.
6. Depreciation rate 10% is used to calculate public infrastructure and non-infrastructure capital stocks.
7. For the definition, unit of variables and data sources, please refer to Table 1.

### Table 3

**Output elasticities: Instrumental variable estimates.**

<table>
<thead>
<tr>
<th>Independent variables: Output per labor</th>
<th>FD IV1 (1)</th>
<th>FD IV1 (2)</th>
<th>FD IV1 (3)</th>
<th>FD IV2 (4)</th>
<th>FD IV2 (5)</th>
<th>FD IV3 (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure capital per labor</td>
<td>-0.095</td>
<td>-0.050</td>
<td>-0.063</td>
<td>-0.098</td>
<td>0.059</td>
<td>0.140</td>
</tr>
<tr>
<td>Periods</td>
<td>(0.09)</td>
<td>(0.14)</td>
<td>(0.10)</td>
<td>(0.19)</td>
<td>(0.21)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Non-infrastructure capital per labor</td>
<td>0.332***</td>
<td>0.210***</td>
<td>0.379***</td>
<td>0.333***</td>
<td>0.258***</td>
<td>0.229***</td>
</tr>
<tr>
<td>Periods</td>
<td>(0.04)</td>
<td>(0.08)</td>
<td>(0.05)</td>
<td>(0.11)</td>
<td>(0.09)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Year effects</td>
<td>All</td>
<td>All</td>
<td>1996–2007</td>
<td>2008–2015</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Overall R2</td>
<td>0.62</td>
<td>0.60</td>
<td>0.63</td>
<td>0.62</td>
<td>0.70</td>
<td>0.71</td>
</tr>
<tr>
<td>No. of observations</td>
<td>509</td>
<td>269</td>
<td>240</td>
<td>509</td>
<td>509</td>
<td>509</td>
</tr>
<tr>
<td>Instruments</td>
<td>Δb₂₀, Δk₂₀</td>
<td>Δb₂₀, Δk₂₀</td>
<td>Δb₁₀, Δk₁₀</td>
<td>Δb₁₀, Δk₁₀</td>
<td>Δb₁₀, Δk₁₀</td>
<td>Δb₁₀, Δk₁₀</td>
</tr>
<tr>
<td>1st-stage regression coefficient</td>
<td>0.337</td>
<td>0.356</td>
<td>0.345</td>
<td>0.272</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td>1st-stage r-ratio</td>
<td>(0.94)</td>
<td>(4.93)</td>
<td>(5.36)</td>
<td>(3.07)</td>
<td>(-2.25)</td>
<td>(-2.11)</td>
</tr>
<tr>
<td>Sargan test (p-value)</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Notes.

1. FD IV denotes first difference instrumental variable regression.
2. Depreciation rate 10% is used to calculate the capital stocks.
3. Standard errors are reported in parentheses. The stars, *, ** and *** indicate the significance level at 10%, 5% and 1%, respectively.
4. Standard errors are adjusted for 30 clusters in province in columns (1)–(6).
Table 4
Output elasticities: Robustness checks.

<table>
<thead>
<tr>
<th>Independent variables:</th>
<th>A: Depreciation rates $\delta_b = 4%$, $\delta_k = 10%$</th>
<th>B: year-end employment</th>
<th>C: FE on Differenced data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FD</td>
<td>FDnew</td>
<td>FDlag</td>
</tr>
<tr>
<td>Infrastructure capital per labor</td>
<td>0.182***</td>
<td>0.075**</td>
<td>-0.003</td>
</tr>
<tr>
<td>Non-infrastructure capital per labor</td>
<td>0.301***</td>
<td>0.211***</td>
<td>0.219***</td>
</tr>
<tr>
<td>Regions</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Periods</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Overall $R^2$</td>
<td>0.74</td>
<td>0.61</td>
<td>0.45</td>
</tr>
<tr>
<td>No. of observations</td>
<td>569</td>
<td>569</td>
<td>539</td>
</tr>
<tr>
<td>Instruments</td>
<td>$\Delta b_{t-2}, \Delta k_{t-2}$</td>
<td>$\Delta b_{t-2}, \Delta k_{t-2}$</td>
<td>$\Delta b_{t-2}, \Delta k_{t-2}$</td>
</tr>
<tr>
<td>1st-stage regression coefficient</td>
<td>0.360</td>
<td>0.274</td>
<td>0.397</td>
</tr>
<tr>
<td>1st-stage t-ratio</td>
<td>(7.14)</td>
<td>(3.81)</td>
<td>(8.47)</td>
</tr>
</tbody>
</table>

Notes.
1. Panel A: depreciation rates of $4\%$ and $10\%$ are used to construct the capital stocks. Definitions of FD, FDnew, FDlag, FDIV1 and FDIV2 remain as in Tables 2 and 3.
2. Panel B: year-end employment is used to measure the labor force. Depreciate rates of $10\%$ remain as in Tables 2 and 3.
3. Panel C: FE, FEnew, FElag, FEIV1 and FEIV2 refer to fixed effects estimates using differenced data and those using newly increased fixed asset investment, lags of public infrastructure and private capital stocks, instruments of lagged values and neighboring public infrastructure, respectively.
4. Standard errors are reported in parentheses. The stars, *, ** and *** indicate the significance level at $10\%$, $5\%$ and $1\%$, respectively.
5. Standard errors are adjusted for 30 clusters in province in columns (1)–(15).
of non-infrastructure capital elasticity $\gamma_k$ is still of a big magnitude of 0.215 and significant, though decreasing from 0.324 in column (2). To further confirm the effect of reverse causality on estimating $\gamma_{F}$, Column (9) gives FD estimates using the lagged value of $\Delta b_k$ and current value of $\Delta k_r$. Same pattern remains as in column (8).

Table 3 reports IV estimates of elasticities using instruments of twice-lagged variables (FDIV1), neighboring public infrastructure (FDIV2) and the ages of provincial governors and party leaders (FDIV3), respectively. Non-infrastructure capital $\Delta k_r$ is also considered as endogenous and instrumented by $\Delta k_r_{t-2}$. The estimates of public capital elasticity using the full sample are -0.095 and -0.098 in columns (1) and (4), respectively. Similar to FDnew and FFlag estimates Table 2, after dealing with the reverse causality between $y$ and $b$ and $k$, the FD IV estimates of output elasticity public infrastructure drop to small negative numbers, and are no longer statistically significant from 0. The FD IV estimates of $\gamma_{F}$ using external instruments of the ages of provincial governors alone and both ages of provincial governors and party leaders are 0.059 and 0.140 in columns (5) and (6), respectively. Both are positive and of a big magnitude, but statistically insignificant.15 Columns (2)–(3) also give FDIV1 estimates using subsamples in the periods of 1996–2007 and 2008–2015.16 The estimates of $\gamma_{F}$ are small, negative, and insignificant.

Unlike $\gamma_b$, the corresponding estimates of $\gamma_{F}$ in columns (1), (4) and (5) in Table 3, 0.332, 0.333 and 0.258, are still positive and significant, and are comparable with the FD estimates in Table 2.17 Thus, the difference between $\gamma_{F}$ and $\gamma_k$ indicates the different roles that the public infrastructure and non-infrastructure capital play in the aggregate production function. Public infrastructure is more likely positively affected by the output than non-infrastructure capital.

Three robustness checks are reported in Table 4: using depreciation rates $\delta_t = 4\%$, $\delta_t = 10\%$ in panel A,18 replacing calculated labor force with year-end employment reported by NBS in panel B, and running fixed effects estimation on differenced data instead of pooled OLS in panel C. In panels A and B, we report FD and 4 estimates using NIFA (FDnew), lagged variables (FFlag) and two internal instruments $\Delta b_t$ and $\Delta k_{r-t}$ (FDIV1) and $\Delta b_t$ in neighboring provinces (FDIV2), corresponding to columns (5), (8) of Table 2 and columns (1), (4) of Table 3, respectively. In panel C, FE and FE estimates using NIFA (FDnew), lagged variables (FFlag) and 2 sets of instruments (FEIV1, FEIV2) are presented in columns (11)–(15), respectively. Consistent with the message delivered by Tables 2 and 3, estimates of $\gamma_{F}$ in columns (2)–(5), (7)–(10), (12)–(14) decrease substantially after reverse causality is taken into consideration. In column (15), using differenced data the fixed effects IV estimate of $\gamma_{F}$ is 0.228 but insignificant. No robust pattern of a big positive and significant estimates of $\gamma_{F}$ are found in various cases, sharply contrasted with the estimates of $\gamma_k$.

Table 5 shows the results of three additional robustness checks. First,

---

15 The first-stage regression results of regressing instruments for $\Delta b$ on exogenous variables in Eq. (5) are reported in the last three rows in Tables 3 and 4. For instruments of instruments of twice-lagged variables and neighboring public infrastructure, both are very informative. The magnitude of instrument of age of provincial governors (age1) is small but still statistically significant. Sargan test for overidentification is conducted in column (6) of Table 3. No evidence shows that instruments of age1 and age2 are invalid.

16 The year of 2008 as the cutoff point is used because a 4 trillion Chinese Yuan fiscal stimulus package was introduced by the Chinese government to invest mainly in the infrastructure in its western provinces in 2008. This event could lead to different output elasticities of infrastructure before and after 2008.

17 The mean value for the ratio $Y/K$ is 0.624 during our sample period. Thus the output elasticities of non-infrastructure capital from Tables 2 and 3 indicate a rate of return around 20%. This number is close to the results reported by Bai and Zhang (2014).

18 We also conduct robustness checks using other different depreciate rates, including combinations of i) $\delta_t = 5\%$, $\delta_t = 10\%$; ii) $\delta_t = 15\%$, $\delta_t = 15\%$; iii) $\delta_t = 10\%$, $\delta_t = 15\%$; iv) $\delta_t = 15\%$, $\delta_t = 10\%$. Main results remain.

---

<table>
<thead>
<tr>
<th>Table 5: Output elasticities: Additional robustness checks.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variables: $\Delta \gamma_b, \Delta k_{r-t}$</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>FD IV2</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>Infrastructure capital per labor</td>
</tr>
<tr>
<td>Non-infrastructure</td>
</tr>
<tr>
<td>capital per labor</td>
</tr>
<tr>
<td>Regions</td>
</tr>
<tr>
<td>Periods</td>
</tr>
<tr>
<td>Year effects</td>
</tr>
<tr>
<td>Overall $R^2$</td>
</tr>
<tr>
<td>No. of observations</td>
</tr>
<tr>
<td>Instruments</td>
</tr>
<tr>
<td>1st-stage regression coefficient</td>
</tr>
<tr>
<td>1st-stage t-ratio</td>
</tr>
</tbody>
</table>

Notes.
1. Panel D: an instrument based on a new measure of infrastructure investments in neighboring provinces, defined as their GDP competitors instead of their geographic neighbors.
2. Panel E: an alternative definition of infrastructure by including investments in industries related to management of water conservancy, environment, and public facilities.
3. Panel F: subsample of eastern region is used.
5. Standard errors are reported in parentheses. The stars, *, ** and *** indicate the significance level at 10%, 5% and 1%, respectively.
6. Standard errors are adjusted for 30 clusters in provinces in columns (1)–(6). Robust standard errors are used in columns (7)–(11).
panel D employs an alternative measure of infrastructure investments in neighboring provinces, defined as their GDP competitors instead of their geographic neighbors used in column (4) of Table 3. The estimated output elasticity of infrastructure becomes 0.077 and statistically insignificant. Second, in Panel E we consider an alternative definition of infrastructure by including investments in industries related to management of water conservancy, environment, and public facilities, i.e., the fourth category of infrastructure investment considered in Shi and Huang (2014). As in Table 4, FD, FDnew, FDlag, FDIV1 and FDIV2 estimates are reported in columns (2)–(6). Third, considering China’s geographic heterogeneity and different economic development across regions, we split the sample into 3 groups: eastern, central and western regions. Panel F presents 5 estimates as in panel E. As in Table 4, the same pattern emerges. Once the reverse causality between the output and infrastructure is mitigated, the estimates of $\gamma_3$ decrease remarkably and become statistically insignificantly in most cases. This evidence suggests that the reverse causality may lead to an upward bias.

5. Conclusion

This paper is motivated by the question whether infrastructure investment contributes to productivity gains and economic growth in China. We address this issue in the framework of an aggregate production function, in which public infrastructure capital is modelled as a contributing factor of TFP, and a panel data set of 30 Chinese provinces during 1996–2015 is used to estimate the output elasticities of public infrastructure and non-infrastructure capital stocks. In such a framework, the main identification problem is the reverse causality between the output and public infrastructure investment, which could lead to an upward or downward bias.

In this empirical study, we proposed several different ways to mitigate the reverse causality. Unlike Shi and Huang (2014), we find that an upward bias dominates when estimating output elasticity of public infrastructure in China’s context. After controlling for the reverse causality between the GDP growth and public investment, we find weak evidence of a big positive productivity effect of public infrastructure within the framework of an aggregate production function.

This, of course, does not deny the possibility that public infrastructure investment may play an important role in economic growth and development. As surveyed by Gramlich (1994), Shi and Huang (2014) and Calderon et al. (2015), there are other econometric issues that are not discussed in the short note. Instead, what we want to highlight here is the challenge of identifying the productivity effect of public infrastructure investment in the aggregate production function estimation.

19 For example, the neighbors of Jiangsu, the ranked 2nd in 2016, are Guangdong and Shandong. The information on Chinese provinces GDP ranking 2016 is from Wikipedia: https://en.wikipedia.org/wiki/List_of_Chinese_administrative_divisions_by_GDP framework. Dealing with reverse causality is of the first order importance, and it is difficult to find good external instruments due the nature of aggregate data. This difficulty suggests the unique value of using alternative identification strategies or data types, e.g., a disaggregation approach using firm-level data such as Fisher-Vanden et al. (2015); Li et al. (2017); and Wu et al. (2017).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.econmod.2018.05.006.

References


