Productivty Growth and Infrastructure Related Sectors

Hildegart Ahumada and Fernando Navajas

UTDT FIEL

22nd Dynamic Econometrics Conference,
Nuffield College, Oxford
September 2019

Outline

- Motivation: productivity growth and infrastructure
- A Global data base
- An overview of sectors
- Econometric approach:panel modeling using *Autometrics*
- Main results
- Final remarks

Motivation: productivity growth and infrastructure

- Infrastructure has been previously identified as one of the critical aspects for per capita income transitions (Izquierdo et al, 2016).
- Applied econometric papers have studied the effects of infrastructure capital on growth (eg Calderon and Serven, 2014, 2016; Egert et al, 2009; and Estache and Garsous, 2012) or on aggregate output (Calderon et al, 2015).
- However, there is no full agreement on which infrastructure sector matters most for growth, with results that may depend on the methodology employed.
- No available results on sectoral growth effects of infrastructure (on which sector).

Motivation: productivity growth and infrastructure

- Growth accounting and KLEMS modeling seem promising for studying *which on which* effects. However, data limitations precludes advances. No studies (except Mas, 2009) deal with infrastructure explicitly.
- We follow Izquierdo et al (2018) insights and explore sectoral productivity growth and the performance of infrastructure related-sectors in terms of their contribution to productivity growth.
- To analyse these issues we focus on a world panel.

A Global data base

- We employ a dataset based on the Groningen Growth and Development Centre (GGDC) data set (Timmer and de Vries, 2007; Trimmer et al, 2015)

- It reports annual labor productivity across 10 sectors following a conventional national accounts sectoral classification (Agriculture, Mining, Manufacturing, Utilities, Construction, Transport, Domestic Trade, Financial services, Government and Social services).

- We select a sample of 25 economies (8 from LAC, 9 from OECD and 8 from Asia) between 1971 and 2014, due to the availability of control variables and the possibility of extending the sample beyond the GGDC dataset subject to national accounts data.

A Global data base

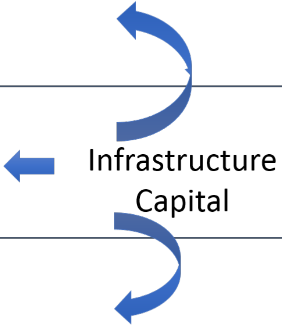
- Controls include different forms of capital that we use as proxies for capital in the infrastructure related sectors, measures of international trade openness, human capital and a political index.
- Infrastructure capital-Labor controls are taken from Penn World Tables (PWT) in combination with de GGDC database on sectoral employment, using "Structures" to proxy the construction sector, "Transport Equipment" to proxy transport. In the case of Utilities we use as a proxy for capital Total Installed Power Capacity from EIA database and employment in the utility sector from GGDC.

A Global data base

Variable	Description	Unit
y_agr	Labour Productivity - Agriculture Sector	Thousands of constant 2005 PPP US dollars per worker
y_min	Labour Productivity - Mining Sector	
y_man	Labour Productivity - Manufacturing Sector	
y_utl	Labour Productivity - Utilities Sector	
y_con	Labour Productivity - Construction Sector	
y_trh	Labour Productivity - Trade, Restaurants & Hotels Sector	
y_tsc	Labour Productivity - Transport, Storage & Communication Sector	
y_fire	Labour Productivity - Finance, Insurance & Real Estate Sector	
y_gvs	Labour Productivity - Government Services Sector	
y_csp	Labour Productivity - Community, Social & Personal Services Sector	
y_Eco	Labour Productivity - Economywide	
hc	Human capital index, based on years of schooling and returns to education	Index
pl_gdpo	Price level of Output-side real GDP (PPP/XR), price level of USA GDPo in 2011=1	Index
trade_gdp	Exports + Imports as fraction of GDP	% of gdp
trade_share	X+M as share of the sample's total X+M	% of sample total trade (X+M)
power_loss	Electric power transmission and distribution losses	% of output
power_installed	Total Installed Power Capacity per capita	Million kW per capita
savings_rate	Savings calculated as 1-(C(%)+G(%))	% of gdp
polity_index	Revised Combined Polity Score = Democ - Autoc, -10 being strongly autocratic, 10 strongly democratic	Index
polity_democ	0 = no democracy ; 10 = full democracy, adjusted	Index
polity_autoc	1 = no autocracy ; 10 = full autocracy, adjusted	Index
polity_transition	1 if transition regime installed	Dummy
polity_foreign_interrupt	1 if foreign interruption installed	Dummy
k_con	Structures Capital Stock per worker in Construction sector	Thousands of constant 2005 PPP
k_tsc	Transport Equipment Capital Stock per worker in Transport, S&C sector	US dollars per worker
k_utl	Total Installed Power Capacity per worker in Utilities sector	Thousand kW per worker

Labor Productivity in 10 Sectors

1. Agriculture
2. Mining
3. Manufacturing
4. **Construction**
5. **Utilities (EG&W)**
6. **Transport (TC&S)**
7. Commercial Services
8. Financial Services
9. Government Services
10. Social Services



Econometric Approach

- The size of the panel requires some methodological strategy given the dimension and the specific questions we want to address.
- An unrestricted VAR model of productivities across sectors and countries leads not only to a more unknowns than observations (as in the GVAR modeling problem) but also to difficulties to interpret effects.
- Izquierdo *et al* (2018) suggest performing an unrestricted VAR for each country in order to find evidence of infrastructure productivity effects on other sectors of the economy.
- An alternative strategy, that we follow in this paper, is to run time-series cross-section panel models for each sector across economies.

Econometric approach

- To study which kind of infrastructure productivity is relevant on which sector labor productivity at world level we estimated panel data models for each sector.
- Our sample with with N= 25 countries and T= 44 years allowed us to estimate dynamic Time Series-Cross Section models (by OLS) to disentangle long-run and short-run effects (taken into account the possibility of integrated data).
- A main feature of our approach is that *Autometrics* (Doornik, 2009 and Hendry and Doornik, 2014) helped us select the relevant variables in a kind of fixed effect panel model, similarly as in Ahumada and Cornejo (2015).

Econometric approach

- *Autometrics* uses an efficient tree search to discard paths rejected as reductions of the initial unrestricted model based on ordered squared t-statistics, given a target p-value provided by the researcher.

- Given T.N= 1075 observations after lagging variables we used 0,001 for outlier dummies and 0,01 for economic explanatory variables and country effects.

Econometric approach

The starting unrestricted model for the s sector,

$$\Delta y_{s,it} = \alpha_i + \gamma_t + \phi_s \; y_{s,it-1} + \beta_{s,utl} \; y_{utl,it-1} + \beta_{s,con} \; y_{con,it-1} + \beta_{s,tsc} \; y_{tsc,it}$$

$$+ \; \delta_{s,utl} \; \Delta y_{utl,it-1} + \delta_{s,con} \; \Delta y_{con,it-1} + \delta_{s,tsc} \; \Delta y_{tsc,it-1}$$

$$+ \theta_{s,utl} \; k_{utl,it-1} + \; \theta_{s,con} \; k_{con,it-1} + \theta_{s,tsc} \; k_{tsc,it-1}$$

$$+ \; \lambda_{s,utl} \; \Delta k_{utl,it-1} + \lambda_{s,con} \; \Delta k_{con,it-1} + \lambda_{s,tsc} \; \Delta k_{tsc,it-1}$$

$$+ \mathrm{x'}_{it-1} \; \psi_s + \Delta x'_{it-1} \; \tau_s + \epsilon_{s,it} \qquad \epsilon_{s,it} \; \sim N(0, \sigma^2_\epsilon)$$

where $i = 1, ..., N$ denotes country and $t = 1, ..., T$ year; $\alpha_i = z_i'\alpha$ where z_i includes the constant and the country dummies.The long-run effects of labor productivities are given by the negative value of $\frac{\beta_{s,utl}}{\phi_s}, \frac{\beta_{s,con}}{\phi_s}, \frac{\beta_{s,tsc}}{\phi_s}$.

Econometric approach

- Including country dummies unrestrictedly (instead of demeaning variables as usually for fixed effect estimation) we can evaluate country heterogeneity by observing the dummies selected by the algorithm,that is productivity differentials across countries not accounted by the variables of the models.
- Other heterogeneities due to elasticity differences can be evaluated by including multiplicative dummies (say for a specific region like LAC).
- Given the persistent behavior of the data we considered the possibility of unit roots formulating the models of labor productivity for their log-differences and the explanatory variables in log-levels and log-differences, as in Westerlund (2007) and Smith and Fuertes (2010).

The log functional form of the last equation allows us to obtain the infrastructure capital productivities as well when we do not reject the hypothesis

$H_0 : \beta_{s,j} = -\theta_{s,j}$ for $j = utl, cons, tsc$.

When it holds the corresponding effects becomes,

$$\begin{aligned}\beta_{s,j} y_{j,it-1} - \theta_{s,j} k_{j,it-1} &= \beta_{s,j}(y_{j,it-1} - k_{j,it-1}) \\ &= \beta_{s,j}((logY - logL)_{j,it-1} - (logK - logL)_{j,i} \\ &= \beta_{s,j}(logY - logK)_{j,it-1}\end{aligned}$$

and therefore, the estimates of $\beta_{s,j}$ become the short-run elasticity with respect to capital productivity of the j infrastructure sector.

Econometric approach

The advantage of estimating this type of model is that it can be easily reparameterized as an error correction (EC) model.

For example, when there is only a long-run effect of one infrastructure sector j on a given s sector productivity, the EC representation is,

$$\Delta y_{s,it} = \alpha_{s,i} + \phi_s \left[y_{s,it-1} - \beta^*_{s,j} \; y_{j,it-1} \right] + \delta_{s,j} \; \Delta y_{j,it-1} + \epsilon_{s,it}$$

where $\beta^*_{s,j} = -\dfrac{\beta_{s,j}}{\phi_s}$

If the variables were first order integrated, we can test whether or not the long-run relationship is a cointegration vector evaluating the significance of the t-statistic of the lagged explained variable (of the estimated coefficient of ϕ).

Econometric approach

- Although the distribution of this statistic is non-standard when there is no cointegration, the critical values derived from the response function in the Monte Carlo study of Ericsson and MacKinnon (2000) can be used to test cointegration (An useful approximation is given by $-3 - 0.2K - 0.3(d-1)$)

- These critical values can be applied to panel data models in the case of OLS estimates using data pooling with country dummy effects.

Econometric approach

We initially assumed:

- There is no residual (country) cross-dependence.
- There is no cross-sectors effects at world level.
- The explanatory variables, in particular the infrastructure sector labor productivities, are all weak exogenous (for our parameters of interest).

  To evaluate how valid these assumptions are for the selected models, we performed the following post-estimation checks.

Econometric approach

- Regarding cross- country dependence we calculated Driscoll-Kraay SE (which allows for heteroscedasticity, cross-section and time dependence for large T panels) to compare significance according to them (using xtscc in STATA, Hoechle,2007). We also report HCSE (White,1980) trying to separate changes in significance due to heteroscedasticity from cross-dependence.

- To test long-run sectors interdependence we augmented the selected model for a sector by the other sectors lagged levels and tested their significance (by *Autometrics* at 1%).The augmented equation for each sector could be considered as an equation of a VEC sectors system with the relevant infrastructure productivities and capitals as exogenous.

Econometric approach

- To analyse weak exogeneity

Our main identification assumption is that the capital per worker of the infrastructure sectors are exogenous and therefore, they can be used as valid instruments.

However, we can note (see Hendry, 2008) that in the case of variables with unit-root representation we can have different sources of no exogeneity.

Econometric approach

In the simple case of the conditional model which assumes cointegration of the labor productivity of the s sector with that of the infrastructure j, the marginal model for the j infrastructure sector could be,

$$\Delta y_{j,it} = \alpha_{j,i} + \rho_j \left[y_{s,it-1} - \beta^*_{s,j} \; y_{j,it-1} \right] + \omega_j \; \Delta y_{s,it-1} + \epsilon_{j,it}$$

While ω_j is associated with Granger- Causality of the s sector on the infrastructure sector j,for weak exogeneity $\rho_j = 0$ is needed, that is, the EC term does not enter the marginal model (often named LR weak exogeneity as in Juselius, 2006).

Econometric approach

Given that we started with a conditional model (a system approach could require estimations of 125 systems) $\rho_j = 0$ requires that the effect of $y_{s,it-1}$ should be not significant in the last equation.

Therefore, no level of the sector s enters into each equation of infrastructure sector j which has effects on sector s. This evaluation was performed when modelling the different sectors.

Since, apart from these long- run effects, there may be that $E[\epsilon_{s,it}.\epsilon_{j,it}] \neq 0$, we used IVE to have consistent estimates from a single equation when a contemporaneous effect of an infrastructure sector was in the estimated model.

Results: Agriculture

Table 1: Agriculture Sector (OLS)

Modelling DLy_agr by OLS	Coefficient	t-SE	t-HCSE	t-prob	t-DK	t-prob
Constant	0.0778	3.87	3.47	0.001	3.21	0.004
Ly_agrL1	-0.0155	-4.35	-4.24	0.000	-4.31	0.000
OECD*Ly_tscL1	0.0189	4.17	4.25	0.000	4.94	0.000
NoOECD*Ly_utlL1	0.0086	2.58	2.69	0.007	3.34	0.003
Ly-Lk_conL1	0.0154	3.83	3.55	0.000	3.01	0.006
DLtrade_share	0.0628	2.68	2.90	0.004	2.60	0.016
COL+IND+MEX+PER+PHL	-0.0279	-4.48	-5.34	0.000	-6.24	0.000

no.of observations	1075
no. of parameters	14
sigma	0.071
Adj. R^2	0.182

$$y_{agr} = const + 1.21 OECDy_{tsc} + 0.55 NoOECDy_{utl} + 0.99 (y-k)_{con}$$

Results: Mining

Table 2: Mining Sector (OLS)

Modelling DLy_min by OLS	Coefficient	t-SE	t-HCSE	t-prob	t-DK	t-prob
Constant	0.0533	1.89	2.27	0.023	2.11	0.046
Ly_minL1	-0.0335	-6.51	-6.13	0.000	-6.79	0.000
Lk_tscL1	0.0214	3.89	4.45	0.000	4.15	0.000
Ly-Lk_utl*Ly_fireL1	0.00159	1.70	2.16	0.031	2.76	0.011
DNK+NLD	0.0980	4.49	3.95	0.000	4.42	0.000
SGP	-0.0731	-2.47	-1.63	0.103	-1.83	0.079

no.of observations	1075
no. of parameters	13
sigma	0.164
Adj. R^2	0.194

$$y_{min} = const + 0.65k_{tsc} + 0.05(y-k)_{utl} * y_{fin}$$

Results: Manufacturing

Table 3: Manufacturing Sector (OLS)						
Modelling DLy_man by OLS	Coefficient	t-SE	t-HCSE	t-prob	t-DK	t-prob
Constant	-0.0063	-0.63	-0.58	0.564	-0.77	0.447
Ly_manL1	-0.0198	-5.45	-5.32	0.000	-5.29	0.000
Ly_conL1	0.0084	2.78	2.76	0.006	4.34	0.000
DLy_utl	0.0508	4.46	2.89	0.004	3.24	0.003
DLy_con	0.1213	6.93	5.20	0.000	3.45	0.002
DLy_tsc	0.2735	10.54	7.46	0.000	8.07	0.000
DLtrade_share	0.0948	5.78	5.01	0.000	6.77	0.000
Ltrade_gdpL1	0.0069	2.76	2.60	0.009	4.40	0.000
LhcL1	0.0393	4.37	3.72	0.000	2.64	0.014
y09-10	-0.0409	-5.90	-4.11	0.000	-15.24	0.000

no.of observations	1075
no. of parameters	15
sigma	0.048
Adj. R^2	0.318

$$y_{man} = const + 0.42y_{con} + 0.35(trade - gdp) + 1.98hc$$

Results: Manufacturing IVE

Table 4: Manufacturing Sector (IVE)

Modelling DLy_man by IVE	Coefficient	t-value	t-prob
Constant	-0.0056	-0.56	0.578
Ly_manL1	-0.0191	-5.16	0.000
Ly_conL1	0.0080	2.62	0.009
DLy_utl	0.0297	1.68	0.093
DLy_con	0.1351	4.11	0.000
DLy_tsc	0.3914	8.08	0.000
DLtrade_share	0.0850	4.94	0.000
Ltrade_gdpL1	0.0059	2.33	0.020
LhcL1	0.0384	4.21	0.000
y09-10	-0.0379	-5.35	0.000

sigma	0.049
no.of observations	1075
no. of parameters	15
Specification test: Chi^2(4)	7.67 [0.105]
Testing beta = 0: Chi^2(14)	395.4 [0.000]**

Additional instruments: DLk_con; DLk_tsc; DLk_utl; Lk_conL1; Lk_tscL1;Lk_utlL1; DUMMULTlk_utll1

Results:Finance, Insurance and Real Estate

Table 5: Finance, Insurance & Real Estate Sector (OLS)

Modelling DLy_fire by OLS	Coefficient	t-SE	t-HCSE	t-prob	t-DK	t-prob
Constant	0.0063	0.45	0.43	0.669	0.46	0.652
Ly_fireL1	-0.0137	-4.56	-5.13	0.000	-4.36	0.000
Ly_conL1	0.0142	3.92	3.43	0.001	3.15	0.004
DLy-DLk_con	0.2022	6.62	4.80	0.000	4.55	0.000
DLk_tsc	0.1219	3.87	3.14	0.002	3.67	0.001
CHN	0.0671	4.62	4.00	0.000	4.27	0.000
ITA	-0.0489	-3.33	-3.01	0.003	-3.20	0.004

no.of observations	1075
no. of parameters	16
sigma	0.083
Adj. R^2	0.366

$$y_{fire} = const + 1.04\ y_{con}$$

Results:Finance, Insurance and Real Estate IVE

Table 6: Finance, Insurance & Real Estate Sector (IVE)			
Modelling DLy_fire by IVE	Coefficient	t-value	t-prob
Constant	0.0078	0.55	0.580
Ly_fireL1	-0.0137	-4.56	0.000
Ly_conL1	0.0135	3.55	0.000
DLy-DLk_con	0.1667	2.48	0.013
DLk_tsc	0.1274	3.88	0.000
CHN	0.0663	4.54	0.000
ITA	-0.0489	-3.33	0.001

sigma	0.083
no.of observations	1075
no. of parameters	16
Specification test: Chi^2(1)	0.085 [0.771]
Testing beta = 0: Chi^2(15)	597.3 [0.000]**

Additional instruments: DLk_con ; Lk_con

Results: Trade, Restaurants and Hotels

Table 7: Trade, Restaurants & Hotels Sector (OLS)						
Modelling DLy_trh by OLS	Coefficient	t-SE	t-HCSE	t-prob	t-DK	t-prob
Constant	-0.0271	-4.35	-3.92	0.000	-3.72	0.001
Ly_trhL1	-0.0233	-7.40	-5.67	0.000	-5.86	0.000
Ly_conL1	0.0225	6.44	5.09	0.000	4.44	0.000
DLy_con	0.0979	4.95	3.17	0.002	2.38	0.025
DLy_tsc	0.3278	13.198	9.07	0.000	9.12	0.000
DLy-DLk_con	0.0777	3.63	2.67	0.008	1.91	0.068
DLtrade_share	0.1008	6.27	4.31	0.000	4.26	0.000
LhcL1	0.0242	3.24	3.30	0.001	2.44	0.022
COL	-0.0223	-3.00	-2.69	0.007	-3.15	0.004
CHN+JPN+SGP	0.0256	5.40	4.57	0.000	3.89	0.001

no.of observations	1075
no. of parameters	16
sigma	0.045
Adj. R^2	0.369

$$y_{trh} = const + 0.97\ y_{con} + 1.03\ hc$$

Results: Trade, Restaurants and Hotels IVE

Table 8: Trade, Restaurants & Hotels Sector (IVE)

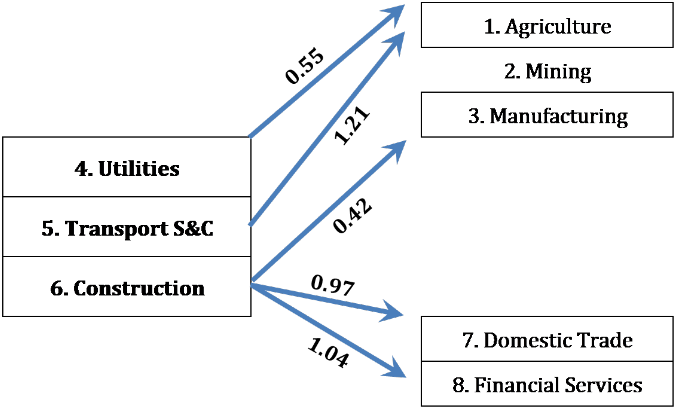
Modelling DLy_trh by IVE	Coefficient	t-value	t-prob
Constant	-0.0322	-4.93	0.000
Ly_trhL1	-0.0239	-5.81	0.000
Ly_conL1	0.0239	4.22	0.000
DLy_con	0.133	1.73	0.084
DLy_tsc	0.4845	5.8	0.000
DLy-DLk_con	0.1219	1.12	0.264
DLtrade_share	0.0713	1.67	0.096
LhcL1	0.0237	2.84	0.005
COL	-0.0201	-2.33	0.019
CHN+JPN+SGP	0.0245	3.89	0.000

sigma	0.047
no.of observations	1075
no. of parameters	16
Specification test: Chi^2(1)	5.0729 [0.024]*
Testing beta = 0: Chi^2(15)	454.59 [0.0000]**

Additional instruments: DLk_con ; Lk_conL1;DLk_tsc ; Lk_tscL1

Summing up results

Long-run labor productivity elasticities: which on which?



Summing up results

- Elasticity estimates show the effect of infrastructure is not uniform across sectors and illustrates the relevance of a *which on which* approach to understand infrastructure effects on productivity.
- Utilities labor productivity has effects on agriculture (outside the OECD), mining (only when interacting with financial services) and manufacturing (only in the short run).
- Transport labor productivity has effects on agricultural sector (for OECD countries), and on domestic trade (only in the short run).
- Construction labor productivity and capital have effects on all the sectors we studied but enter the models in different ways (short and/or long run; labor or capital productivities).

Summing up results

- The effects of infrastructure capital stocks should be taken into account when modeling infrastructure productivity as drivers of other sectors.
- Other control variables such as trade intensity and human capital enter in some equations. However, other institutional or political proxies were not detected as having significant effects in our data sample.
- The estimated models provide a framework to test region or country heterogeneity for the different sectors.
- The estimates also allow us to evaluate how productivity improvements in infrastructure-related sectors affects aggregate productivity through direct and indirect effects.

Final remarks: future extensions

- Qualifications regarding data and methods which we take as a positive aspect to motivate improving this line of research.
- A two digit sectoral database is still too-aggregated and more measurement efforts are needed to obtain more disaggregated productivities.
- Even with this aggregation level, we could benefit from the spread of the KLEMS methodology across regions, which would allow to study TFP.
- The econometric approach can move towards exploiting more interactions across sectors and countries.