Data Sources and Results for "Skyscrapers and Global Connectivity: Which Causes Which?"

Jason Barr

March 26, 2022

I. Data Sources

- The Standardized Globalization Connectivity Network (SGCN) data was generoulsy provided by GaWC project. The data set gives the global connectivity scores for the years 2010, 2013, 2016, 2018 and 2020. They are adjusted so as to make them comparable across years. See DeRudder et al. (2010) for more information
- Skyscraper counts and heights comes from the Emporis.com database. For each year, two dependent variables were created: ln(1+# 90m or taller buildings) for each city and log of the height of tallest building city (in meters).
- City population and GDP for each city were as of 2015 and was from the Global Human Settlement database, https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php . It also includes data on urban elevations, which are used as instruments for some regression given below. The details about the data are here: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC115586/ghs_stat_ucdb2015mt_globe_r2019a_v1_0_web_1.pdf
- Data on average bedrock depths for urban areas is form https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2016MS000686, and which are also used as instruments.

II. Descriptive Statistics & t-test

Summary of Standardized GNC values

year	Mean	Std. dev.	Freq.
2010	16.466091	15.176477	526
2013	15.960901	15.435644	525
2016	13.346086	14.256514	707
2018	14.981555	14.399563	708
2020	15.532747	14.667141	707
Total	15.148097	14.763962	3,173

Variable	Obs	Mean	Std. dev.	Min	Max
SGCN	3,173	15.1481	14.76396	0	103.8559
Max height	1,656	174.8519	102.8942	11.1	828
# skyscrapers	1,656	94.92089	323.8503	0	417
Avg Bedrock Depth(cm)	2,295	1900.812	1326.093	3.192308	13339.35
Avg Elevation (cm)	2,200	341.8763	554.6335	-405.5	4267
GDP 2015	2,300	4.07e+10	9.18e+10	9697784	1.01e+12
Pop 2015	2,300	2495099	4019334	51718.52	3.63e+07

This t-test compares ***change***s in the SGCN between 2010 and 2020 for those countries (SGCN) with and without the countries tallest building in 2010:

Two-sample t test with unequal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	368	2.449377	.2875238	5.515662	1.883976	3.014777
1	144	3.177865	.3476579	4.171894	2.490652	3.865077
Combined	512	2.654264	.228905	5.17953	2.204553	3.103975
diff		-.7284882	.4511495		-1.615856	.1588798

```
diff = mean(0) - mean(1)                                   t =  -1.6147
H0: diff = 0                  Satterthwaite's degrees of freedom =  342.994

Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
Pr(T < t) = 0.0536         Pr(|T| > |t|) = 0.1073          Pr(T > t) = 0.9464
```

The above table barely rejects equal growth, and just seems to not reject larger growth for the cities with the tallest buildings. It suggests a somewhat larger growth rate in the SGNC for cities with a country's tallest building in 2010. However, the table below suggest that in the last few years, cities with the tallest building in the respective country have not continued to grow as fast. It thus appears there is some convergence over time.

Two-sample t test with unequal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	550	.6182724	.0955333	2.240455	.4306169	.8059279
1	157	.2412633	.1990374	2.493932	-.1518929	.6344195
Combined	707	.5345519	.0866008	2.302669	.3645259	.7045778
diff		.3770091	.2207771		-.0579688	.811987

```
    diff = mean(0) - mean(1)                                t =   1.7076
H0: diff = 0                  Satterthwaite's degrees of freedom =  232.649

    Ha: diff < 0               Ha: diff != 0               Ha: diff > 0
 Pr(T < t) = 0.9555        Pr(|T| > |t|) = 0.0890        Pr(T > t) = 0.0445
```

III. Regression Results

Dep Var: Ln(Max)

OLS of Ln(tallest) on lag of SGCN

	(1) lnMax	(2) lnMax	(3) lnMax	(4) lnMax
$SGNC_{t-1}$	**0.0156*** **	**0.00408** **	**0.00271**	**0.00992*** **
	(19.00)	**(3.14)**	**(1.98)**	**(4.68)**
lnSkyCount		0.187***	0.190***	

		(5.87)	(4.82)	
lnPop15			-0.0374	0.0889
			(-0.94)	(1.67)
lnGDP15			0.0539	0.0341
			(1.57)	(0.80)
lnAvgBedrock			0.00808	-0.0166
			(1.33)	(-0.86)
lnAvgEleva~n			0.0118	0.0242
			(0.93)	(1.91)
_cons	3.707***	4.090***	3.521***	2.202*
	(31.08)	(18.74)	(11.47)	(2.53)
N	999	999	899	899
R-sq	0.760	0.875	0.881	0.791
adj. R-sq	0.727	0.858	0.865	0.762
AIC	84.44	-569.9	-569.1	-59.26
BIC	109.0	-545.4	-545.1	-35.25

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. All equations have country dummy variable (i.e. country fixed effects) and year dummies

The above table regresses ln(Max) on the lag of the global network (which are taken in 2-3 year intervals). Presumably the lag is exogenous. Below is results from xtabond estimation in Stata which gives results for panel regression with lagged dependent variable. Including a lag dependent variable is good way to control for omitted variables, but it then introduces endogeneity. The xtabond procedure corrects for this. The table shows the estimates to be similar as the OLS one.

xtabond regression

	(1) D.lnMax	(2) D.lnMax	(3) D.lnMax
LD.lnMax	0.276*	0.0883	0.287*
	(1.91)	(0.63)	(1.75)
D.timevar	0.0204***	0.00494	
	(2.88)	(0.91)	
D.SGCN	**0.00370*** **	**0.00207***	**0.0148*** **
	(2.86)	**(1.80)**	**(1.99)**
D.lnSkyCount		0.168***	
		(8.43)	
D.lnPop15			0.00374
			(0.50)
_cons			0.0208***
			(2.66)
_cons	0	0	0.00403***
	.	.	(2.86)

N	596	596	596

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01

Here are regressions for the number of skyscraper completion in each city each year.

Dep Var.: ln(1+# skyscrapers). OLS Regression

	(1) lnSkyCount	(2) lnSkyCount	(3) lnSkyCount	(4) lnSkyCount
$SGNC_{t-1}$	**0.0619*** **	**0.0619*** **	**0.0379** **	**0.0379** **
	(15.24)	**(15.24)**	**(3.42)**	**(3.42)**
lnPop15			0.664	0.664
			(1.88)	(1.88)
lnGDP15			-0.104	-0.104
			(-0.38)	(-0.38)
lnAvgBedrock			-0.130	-0.130
			(-1.46)	(-1.46)
lnAvgEleva~n			0.0649	0.0649
			(1.31)	(1.31)
_cons	-2.050***	-2.050***	-6.935*	-6.935*
	(-4.38)	(-4.38)	(-2.27)	(-2.27)
N	999	999	899	899
R-sq	0.765	0.765	0.826	0.826
adj. R-sq	0.734	0.734	0.802	0.802
AIC	2700.8	2700.8	2172.1	2172.1
BIC	2720.5	2720.5	2196.1	2196.1

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. All regression includes city fixed effects and year dummies.

xtabond regression

	(1) D.lnSkyCount	(2) D.lnSkyCount	(3) D.lnSkyCount	(4) D.lnSkyCount
LD.lnSkyCount	0.451***	0.422***	0.438***	0.363***
	(5.29)	(4.26)	(4.70)	(2.75)
D.timevar	0.0603***	0.0676***	0.0672***	
	(3.59)	(3.88)	(4.07)	
D.SGCN	**0.00570** **	**0.00508***	**0.00499**	**0.0527*** **
	(2.13)	**(1.67)**	**(1.51)**	**(2.97)**
D.lnPop15				0.0323*
				(1.86)
_cons				0.0788***
				(3.41)
_cons	0	0	0	0.00706*
	.	.	.	(1.74)
N	770	596	547	596

----

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01

Regressions for Standardized Global Network Connectivity Index on skyscraper measures.

OLS Regressions. Dep. Var.: SGNC

	(1) standardiz~c	(2) standardiz~c	(3) standardiz~c	(4) standardiz~c	(5) standardiz~c	(6) standardiz~c
$lnSkyCount_{t-1}$	7.711***	6.871***	5.411***	4.467***	5.089***	4.215***
	(11.35)	(9.30)	(6.14)	(6.64)	(4.55)	(5.41)
$lnMax_{t-1}$	-1.774*	-1.624*	-0.737	-0.507	-0.700	-0.516
	(-2.41)	(-2.35)	(-1.08)	(-0.78)	(-1.27)	(-0.96)
$havecountrtallest_{t-1}$		10.86**		12.72***		11.45**
		(2.92)		(4.03)		(3.21)
lnGDP15			5.032***	4.779*	5.620**	4.431**
			(4.88)	(2.53)	(2.97)	(3.20)
lnAvgEleva~n			-1.095	-1.031	-0.899	-0.890
			(-1.47)	(-1.61)	(-1.41)	(-1.53)
lnPop15				0.229	-1.582	-0.0723
				(0.12)	(-0.69)	(-0.04)
lnAvgBedrock				-0.195	-0.128	-0.230
				(-0.13)	(-0.08)	(-0.16)
have300m sky $_{t-1}$					10.84*	9.455
					(2.36)	(1.99)
_cons	22.87***	13.07**	-90.48**	-94.19**	-95.50**	-88.64**
	(5.95)	(3.70)	(-3.46)	(-3.23)	(-3.67)	(-3.52)
N	706	705	639	638	637	637
R-sq	0.710	0.728	0.801	0.823	0.819	0.837
adj. R-sq	0.665	0.686	0.768	0.794	0.789	0.809
AIC	5086.0	5033.0	4381.3	4298.7	4308.0	4242.9
BIC	5108.8	5055.8	4403.6	4321.0	4330.2	4265.2

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01 All regressions have country and year dummies.

The above results show impact from height of tallest building, but positive impact from having a 300 meter or tall building or the country's tallest. Perhaps ln(Max) is endogenous. Below are regressions for IV estimation.

Instrumental Variable Regressions. Dep. Var.: SGCN

	(1)	(2)	(3)	(4)	(5)	(6)
$lnSkyCount_{t-1}$	**12.52*** **	**6.330*** **		**6.276*** **		
	(15.44)	**(4.24)**		**(3.82)**		
lnSkyCount			**5.994*** **			
			5.71)			
lnMax					**28.78*** **	**49.17*** **
					(9.40)	**(12.37)**
lnGDP15		4.280***	4.411***	4.270**	3.665***	
		(2.76)	(3.40)	(2.57)	(5.28)	

	(1)	(2)	(3)	(4)	(5)	(6)
_cons	6.617***	-66.41**	-102.5***	-61.35	-221.3***	-233.0***
	(15.53)	(-2.11)	(-3.26)	(-1.47)	(-17.60)	(-12.82)
N	878	878	1249	880	1249	1249
R-sq	0.582	0.751	0.764	0.750	0.727	0.515
adj. R-sq	0.529	0.719	0.739	0.718	0.699	0.465
OVERID (p-val)	0.1736	0.613	0.1895	0.6038	0.9121	0.1033
ENDOG	0.00	0.36	0.7989	0.3592	0.0148	0.0000
FIRST (Min Eigen)	119.1	16.6	25.47	22.1411	15.1198	152.73

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01 Year and country fixed effects in all regressions IV reg instruments include ln(avg bedrock depth), ln(avg elevation), latitude (degrees), lnPop and lnGDP (but different set of IVs for different regressions).

The lnMax variable is now statistically significant and it appears that the IVs are exogenous.

Here the xtabond results for SGNC scores. The results support that skyscraper construction positively effect the scores.

	(1)	(2)	(3)	(4)
	D.standard~c	D.standard~c	D.standard~c	D.standard~c
LD.standar~c	0.386***	0.435***	0.457***	0.402***
	(4.80)	(5.52)	(5.76)	(5.63)
D.timevar	0.614***	0.649***	0.830***	0.628***
	(5.41)	(5.12)	(7.20)	(5.05)
D.lnSkyCount	**2.272*** **	**2.521*** **		**2.195*** **
	(4.11)	**(4.39)**		**(3.45)**
D.lnMax	**1.419**		**4.405*** **	**2.070**
	(1.11)		**(3.09)**	**(1.33)**
D.havecoun~t	-0.480		-1.321	-1.232
	(-0.64)		(-1.42)	(-1.37)
N	831	646	646	646

t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01