NBER WORKING PAPER SERIES

SUBWAYS AND URBAN GROWTH:
EVIDENCE FROM EARTH

Marco Gonzalez-Navarro
Matthew A. Turner

Working Paper 24996
http://www.nber.org/papers/w24996

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2018

We are grateful to Fern Ramoutar, Mahdy Saddradini, Mohamed Salat, and Farhan Yahya for their assistance compiling the subway data. We are also grateful to seminar participants at Brown University, University of Chicago, Georgia Tech,ITAM, LSE,PERC, UC Berkeley and the World Bank, and to Victor Aguirregabiria, Dwayne Benjamin, Gilles Duranton, Emilio Gutierrez, Walker Hanlon, Frank Kleibergen, Andreas Kopp, Joan Monras, Peter Morrow and Ken Small for helpful comments and conversations. This paper is part of a Global Research Program on Spatial Development of Cities, funded by the Multi-Donor Trust Fund on Sustainable Urbanization by the World Bank and supported by the UK Department for International Development. The project was made possible through financial support from SSHRC, IGC, Ontario Work Study program and Societé du Grand Paris. Turner acknowledges the financial support and hospitality of the Property and Environment Research Center, and the Enaudi Institute of Economics and Finance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Subways and Urban Growth: Evidence from Earth
Marco Gonzalez-Navarro and Matthew A. Turner
NBER Working Paper No. 24996
September 2018
JEL No. L91,R11,R14,R4

ABSTRACT

We investigate the relationship between the extent of a city's subway network, its population and its spatial configuration. For the 632 largest cities in the world we construct panel data describing population, measures of centralization calculated from lights at night data, and the extent of each of the 138 subway systems in these cities. These data indicate that large cities are more likely to have subways but that subways have an economically insignificant effect on urban population growth. Our data also indicate that subways cause cities to decentralize, although the effect is smaller than previously documented effects of highways on decentralization. For a subset of subway cities we observe panel data describing subway and bus ridership. For those cities we find that a 10% increase in subway extent causes about a 6% increase in subway ridership and has no effect on bus ridership.

Marco Gonzalez-Navarro
Agricultural and Resource Economics
UC-Berkeley
207 Giannini Hall
Berkeley, CA 94720-3310
and University of California, Berkeley
marcog@berkeley.edu

Matthew A. Turner
Department of Economics
Brown University
Box B
Providence, RI 02912
and NBER
matthew_turner@brown.edu

1. Introduction

We investigate the relationship between the extent of a city's subway network and its population, transit ridership and spatial configuration. To accomplish this investigation, for the 632 largest cities in the world we construct panel data describing population, total light, measures of centralization calculated from lights at night data, and the extent of each of the 138 subway systems in these cities. For a subset of these subway cities we also assemble panel data describing bus and subway ridership.

These data suggest the following conclusions. First, while large cities are more likely to have subways, subways have a precisely estimated near zero effect on urban population growth. Second, subways cause cities to decentralize, although this effect appears to be small relative to the decentralization caused by radial highways. Third, a 10% increase in subway extent leads to about a 6% increase in subway ridership and does not affect bus ridership. A back of the envelope calculation suggests that only a small fraction of ridership increases can be accounted for by decentralized commuters. Together with the fact that little new ridership can be attributed to population growth, this suggests that most new ridership derives from a substitution from other modes of travel towards subways.

Subway construction and expansion projects range from merely expensive to truly breathtaking. Among the 16 subway systems examined by Baum-Snow and Kahn (2005), construction costs range from about 25m to 550m USD2005 per km. On the basis of the mid-point of this range, 287m per km, construction costs for the current stock are about 3 trillion dollars. These costs are high enough that subway projects generally require large subsidies. To justify these subsidies, proponents often assert the ability of a subway system to encourage urban growth.[1] Our data allow the first estimates of the relationship between subways and urban growth. That subways appear to have almost zero effect on urban growth suggests that the evaluation of prospective subway projects should rely less on the ability of subways to promote growth and more on the demand for mobility. Our data also allows the first panel data estimates of the impact of changes in system extent on ridership and therefore also make an important contribution to such evaluations.

Understanding the effect of subways on cities is also important to policy makers interested in the process of urbanization in the developing world. Over the coming decades, we expect an enormous migration of rural population towards major urban areas, and with it demands for urban infrastructure that exceed the ability of local and national governments to supply it. In order to assess trade-offs between different types of infrastructure in these cities, understanding the implications of each for welfare is clearly important. Since people move to more attractive places and away from less attractive ones (broadly defined), our investigation of the relationship between subways and population growth will help to inform these decisions. In particular, if the

[1] A statement by the agency responsible for Toronto's transit expansion is typical: "Expanding transportation can help create thousands of new green and well-paid jobs, and save billions of dollars in time, energy and other efficiencies." (`http://www.metrolinx.com/en/regionalplanning/bigmove/big_move.aspx`) (accessed July 28, 2014).

objective of policymakers is to increase a city's population or to decentralize economic activity, highways seem more promising. On the other hand, in a related companion paper, Gendron-Carrier et al. (2017) show that if the objective is to reduce pollution, then subways can be effective.

Finally, an active academic literature investigates the effect of transportation infrastructure on the growth and configuration of cities. In spite of their prominence in policy debates, subways have so far escaped the attention of this literature. This primarily reflects the relative rarity of subways. Most cities have roads so a single country can provide a large enough sample to analyze the effects of roads on cities. Subways are too rare for this. A statistical analysis of the effect of subways on cities requires data from, at least, several countries. An important contribution of this paper is to assemble data that describe all of the world's subway networks. In addition, with few exceptions, the current literature on the effects of infrastructure is static or considers panel data that is too short to investigate the dynamics of infrastructure's effects on cities. Because our panel spans the 60 year period from 1950 until 2010, we are able to investigate such dynamic responses to the provision of subways.

To estimate the causal effects of subways on urban growth and urban form, we must grapple with the fact that subway systems and stations are not constructed at random times and places. This suggests two potential threats to causal identification. The first could occur if subway expansions systematically take place at times when a city's population growth is slower (or faster) than average. For example, if construction crews leave the city when new subway expansions are complete or if subway expansions tend to occur when some constraint on a city's growth begins to bind. The second results from omitted variables. For example, suppose that cities expand their bus networks in years when they do not expand their subway networks and that bus and subway networks contribute equally to population growth. Then any regression of population growth on subway growth that omits a measure of the bus network will be biased downward. Briefly, we address the problem of confounding dynamics by showing that the null population growth result is invariant to using first differences, instrumented first differences, second differences and dynamic panel data models. The instrument we propose takes advantage of the fact that larger subway systems grow more slowly and this allows us to predict subway growth using long lags of subway system size. We address the omitted variables issue by showing that the null effect of subways on population is not masking heterogenous effects by measures such as congestion, road supply, bus supply, institutional quality, city size, or size of network, among others.

2. Literature

A *Subways*

With a few exceptions that we describe below, the literature that analyzes the effects of subways on cities consists entirely of analyses of a single city. Nevertheless, this literature is large and we here focus our attention on the small set of papers which attempt to resolve the problem of non-random assignment of subways. More complete surveys are available in Billings (2011) and Gibbons and Machin (2005).

Gibbons and Machin (2005) examine housing prices in London during the periods 1997-1999 and 2000-2001, periods that bracket two expansions of the London underground. Gibbons and Machin (2005) calculate various difference-in-differences estimates of the effect of these transit expansions on housing prices and find that moving one km away from a subway station decreases house values by about 2% for the first two km, and about zero thereafter. Billings (2011) conducts a similar exercise for a new light rail line in Charlotte, North Carolina.[2] Like Gibbons and Machin (2005), Billings (2011) estimates the effect of subways on housing prices using a difference-in-differences estimator. Despite differences in milieu and method, Billings (2011) arrives at estimates quite close to those of Gibbons and Machin (2005): single family houses within 1.6km of the transit line see their prices increase by about 4% while condominiums see their prices rise by about 11%. Like Gibbons and Machin (2005), Billings (2011) observes that changes result from subway construction over the course of just a few years.[3]

Each of these papers makes a credible attempt to overcome the fact that subway systems are not located randomly within cities. However, neither provides us with much information about the relationship between subways and city-level growth. If subways affect the growth of cities, then they may affect it everywhere, both near and far from a station. By construction, a differences-in-differences methodology cannot measure such citywide effects. Therefore, while the existing literature makes some progress on the problem of non-random assignment of subways to places, it does so at a high cost. The difference-in-differences methodology cannot tell us about the effect of changes in the overall level of activity within a city. Unless we are specifically interested in reorganizing economic activity across neighborhoods within a city, it is changes in the overall level which are of primary policy interest and which are the object of our investigation.

Finally, in an important contribution Ahlfeldt, Redding, Sturm, and Wolf (2015) estimate a structural model of how a subway network can restructure a city, rather than just whether subways attract development. Given this, it is closest in spirit to our decentralization exercise. With this said, Ahlfeldt *et al.* (2015) use time series variation from just one city, so their ability to investigate the effect of subways on urban growth relies heavily on the assumptions underlying their model.

There are few studies considering cross city or panel data on transit and city level outcomes. Pang (2017) and Gendron-Carrier *et al.* (2017) are rare exceptions. Pang (2017) uses US data to investigate the effect of public transit on employment rates for low-skilled workers, and finds that low skilled workers are more likely to be employed as their access to subways improves. Gendron-Carrier *et al.* (2017) is a companion paper to this one, and uses an event study methodology to investigate the relationship between airborne particulates and subway system opening. In

[2]The Charlotte light rail system is not completely isolated from pedestrian and automobile traffic and so does not appear in our data as a subway.

[3]Manelici (2017) investigates the interaction between terrorism and proximity to London subway stations and finds that a terrorist attack in London in 2005 disproportionately affected real estate prices near subway stations.

a sample of about 40 cities all over the world, it finds that subway openings cause economically important reductions in pollution.

Apart from these two, the only studies to investigate the effects of subways on city level outcomes are primarily or completely interested in ridership.[4] On the basis of a single cross-section of about 50 cities, Gordon and Willson (1984) conduct a city level regression to predict riders per mile of track as a function of city population density and country level per capita GDP. They find that these two variables are excellent predictors of ridership - the relationship being positive and negative, respectively. Finally, Baum-Snow and Kahn (2005) provide evidence from 16 US cities for a similar relationship between density and transit use, although their small sample size limits the precision of their results. They also show that ridership shares in catchment areas for new stations attain almost the same level as in the catchment areas of old stations over their 30 year study period. Consistent with the finding in Gordon and Willson (1984) that ridership decreases with income and increases with density, Baum-Snow and Kahn (2005) find that most US transit expansions have only small effects on ridership, a conclusion echoed in Gomez-Ibanez (1996) for time series data on the use of Boston's transit system. Our results on the relationship between subway extent and ridership are the first to exploit city level panel data. Barnes (2005) provides evidence from a few cities in the US that people are more likely to take transit for trips to a central business district than for trips to other locations.

B *Other infrastructure*

Redding and Turner (2015) survey the literature relating roads and highways to urban growth. This literature has developed rapidly over the past several years and suggests the following conclusions.

First, Duranton and Turner (2012) find that the stock of highways in a city contributes to the growth in city population in the US between 1980 and 2000. This effect is small in an absolute sense, though it is economically important as a share of the total growth rate. Using a similar research design, Garcia-López, Holl, and Viladecans-Marsal (2015) finds that highways cause about the same rate of population growth in Spanish cities.[5]

Second, that radial highways can have dramatic effects on the internal structure of cities. Baum-Snow (2007) investigates the effect of radial highways on population decentralization for a sample of large US cities between 1950 and 1990. He finds that, over the whole 40 year course of his study period, a single radial highway causes about a 9% decrease in central city population. This large decentralizing effect of highways is confirmed for China by Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017) and for Spain by Garcia-López (2012).

[4]We note the large literature on modal choice using individual level data. This important literature is only tangentially related to our present inquiry. A survey is available in Small and Verhoef (2007).

[5]Related to this, Blonigen and Cristea (2015) and Campante and Yanagizawa-Drott (2018) investigate the role of airports in urban growth and argue for a causal relationship between airport traffic and urban growth.

Finally, Duranton and Turner (2011) and Hsu and Zhang (2014) find that vehicle kilometers traveled increase about proportionately to increases in the extent a city's road network, and that increases to non-commute driving appear to be the most important contributor to this increase. All of these responses, decentralization, growth and driving, can be detected over a 5-20 year time horizon, much shorter than our 60 year study period.

In contrast, we find that the effects of subways on urban growth are tiny. We find a much larger effect of subways on the configuration of cities. The effect of subways on ridership is large, though probably smaller than the effect of roads on driving. Finally, we will present indirect evidence to suggest that only a small fraction of the increase in ridership reflects decentralized commuters. More likely, commuters shift their mode of transportation towards subways.

3. Results vs. Theory

Anticipating our results, our data indicate that marginal changes to a city's subway network have the following effects. First, subways have approximately zero effect on a city's population. Second, subways cause cities to decentralize in a way that is qualitatively similar to the way that roads cause cities to decentralize. Third, the cities in our sample grow at about 2% per year. In addition, from Gendron-Carrier *et al.* (2017), we know that subway openings cause reductions in air pollution that have an estimated value that is of about the same order of magnitude as construction costs (though the effects of expansions are probably smaller). Finally, what evidence we have on the matter, e.g., Baum-Snow and Kahn (2005), suggests that subways are very expensive to build and that fare revenue does not fully cover operating costs.

These facts are consistent with basic theory. To see this, consider a simple linear city. Each identical agent consumes a unit of land at distance x from the center and commutes to $x = 0$ where she receives wage, w_t. The unit cost of travel is τ_t and land rent, $R_t(x_t)$, varies with distance to the center. Subscripts index two periods, $t \in \{0,1\}$, an initial period where the subway is smaller or not present, and a later period when the subway is more extensive.

Denote the most remote occupied location as $\overline{x}_t$ and suppose that land rent is zero beyond this boundary. Because each agent consumes exactly one unit of land, $\overline{x}_t$ describes both the physical extent of the city and its population. Agents derive utility from consumption, c_t, and from a city specific amenity, A_t. They pay a tax T_t to fund the subway. Agents have the choice to reside in the city of interest, or at some alternative that provides utility u_t^*.

A representative agent solves the following problem,

$$\begin{aligned} &\max c_t + A_t \\ &\text{s.t. } w_t = c_t + \tau_t x + T_t - R(x). \end{aligned}$$

With free mobility, this implies that the boundary of the city is determined by the following condition,

$$\overline{x}_t = \frac{w_t + A_t - T_t - u_t^*}{\tau_t}.$$

Our finding that subways do not change city population means that $\overline{x}_0 = \overline{x}_1$, and hence that

$$\frac{w_0 + A_0 - T_0 - u_0^*}{\tau_0} = \frac{w_1 + A_1 - T_1 - u_1^*}{\tau_1}. \tag{1}$$

Because cities decentralize with subways, we conclude that they drive down unit transportation cost. In the context of this model, this means that $\tau_1 < \tau_0$. Together with equation (1), this requires that

$$w_0 + A_0 - T_0 - u_0^* > w_1 + A_1 - T_1 - u_1^*. \tag{2}$$

If our identification strategy is successful at isolating the effects of quasi-random variation, then a city with a subway expansion faces the same outside option as a city without. That is, $u_1^* = u_0^*$. In this case, our results suggest that whatever beneficial effects subways have on transportation costs, wages and amenities are about offset by the local share of costs.[6]

On the other hand, if cities either invest in subways or in some substitute, then we might think that $u_1^* > u_0^*$. In this case, our results need to be understood as relative, not to the status quo, but to the improving outside option. This means that the benefits of subways minus the local share of costs is no better that the alternative investment.

This model is deliberately stylized and so welfare interpretations should be regarded with care. In particular, the model omits the possibility that people consume more space as transportation costs fall. This is surely valuable, but statements about residential density are beyond the reach of our data and so we omit this margin of adjustment from our model.

In addition, this simple model is based on the assumption that improvements to the subway network reduce transportation costs everywhere. In fact, as we will show, subways overwhelmingly serve central cities. Since mode transfers are costly, one can therefore imagine that subways could reduce the cost of commuting within their central service area, while leaving commute costs from more remote locations more or less unchanged. In this case, changes to the subway system would not affect the condition determining the edge of the city. Relative to the model articulated above, such model has the advantage of greater realism and of predicting widely observed increases in land rent in subway catchment areas. With this said, the same basic intuition holds. In order for population to remain constant with increases in subways, we require that the increased tax burden offset improvements in local wages or amenities.

[6] To our knowledge, there is little systematic evidence about subway funding arrangements. Official reports for Mexico City's subway system suggest that farebox revenue only accounts for about half of operating costs, the rest being financed by subsidies from general city funds. In Toronto's TTC a third of operating costs are covered by the city's property tax. Gomez-Ibanez (1996) reports that capital costs for Boston's subway from about 1965-1990 on were substantially funded by the federal government.

4. Data

To investigate the effect of subways on the evolution of cities' population, spatial structure and transit ridership, we require data for a panel of cities. We construct such data from four principal sources. Our population data are the UN World Cities Data. Our subway data are the result of primary data collection, as is our ridership data. Our description of urban spatial structure derives from satellite lights at night data.

A *Population data*

Our data are organized around the UN World Cities Data.[7] Produced by the United Nations, Department of Economic and Social Affairs, Population Division, these data describe population counts for all cities whose population exceeds 750,000 at any time during 1950-2010.

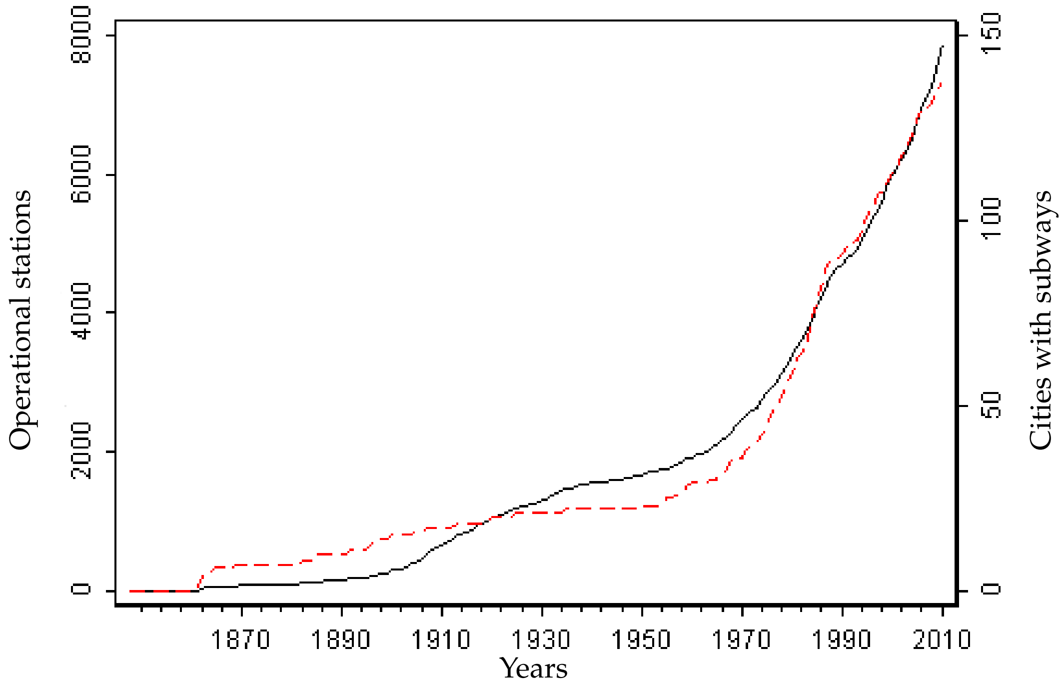
Constructing international data describing city level population is subject to two difficulties. First, population data are generally, but not always, available from decennial or quinquennial censuses, but do not synchronize neatly across countries. To resolve this problem, the UN World Cities Data interpolate across available censuses to construct annual values. Therefore, because few countries conduct censuses more often than every five years, successive annual population changes must sometimes reflect linear interpolation of the same proximate census years. To avoid making inferences from such imputed population changes, we restrict attention to observations drawn every fifth year (e.g., 1950, 1955, ...) and refer to each such observation as a 'city-year'. This decreases the likelihood that sequential city-years are calculated by interpolation from the same two underlying censuses. In fact, for some countries, census data is available less often than every five years, so we also experiment with observations drawn every 10 years and with even longer periods.

A second difficulty arises because metropolitan areas and census units are not defined at the same scale in all countries. To overcome this problem, the UN World Cities Data is based on population counts at the most geographically disaggregated administrative unit available from every country. Once equipped with these data, metropolitan areas are defined as a fixed set of smaller administrative units — regardless of whether the smaller units were in the same state for example. This allowed UN researchers to use a consistent definition of metropolitan areas across countries and over time, and captures what we think of as metropolitan areas.

The top panel of Table 1 describes our population data. The data consist of 632 cities, more than half in Asia. In 2010, the mean population of a city in our sample is about 2.4 million. There is little variation in mean population across continents, although cities in South America tend to be larger while cities in Europe tend to be smaller. Between 1950 and 2010, the mean five year growth rate of a city in our sample is about 18%. This rate falls by about 1 percentage point every five years. Not surprisingly, cities in Africa, Asia and South America grow faster than in North America and

[7]Downloaded from `http://esa.un.org/unup/GIS-Files/gis_1.htm`, February 2013.

Figure 1: Growth of world subway systems



Note: The dashed line indicates the number of cities with a subway system (right axis) and the solid line indicates the total number of operational stations (left axis).

Europe. European cities are the obvious outlier and grow more slowly than cities elsewhere. The growth rate of cities is declining on all continents and this decrease is somewhat slower in Europe.

The bottom panel of Table 1 describes our population data for the 138 cities in our sample with a subway in 2010. At 4.7m people on average, these cities are about twice as large as non-subway cities. Cairo is the single African city with a subway, and so the Africa column in the bottom panel of table 1 is really a 'Cairo column'.[8] Asian and South American subway cities are larger than those in North America and dramatically larger than those in Europe. The five year growth rate for an average subway city is about 11%, slower than in the whole sample. As for the whole sample, European subway cities are growing more slowly than other subway cities. Also similar to the whole population of cities, growth rates between 1950 and 2010 are declining by about 1% every five years and this decrease is somewhat slower in Europe.

B *Lights data*

Lights at night data are collected by earth observing satellites that measure the intensity of visible light every night in 30 arc second cells (about one kilometer square) on a regular grid covering

[8]Australia contains few large cities and has no subways in 2010. To simplify the exposition, we have consolidated Asia and Australia.

Table 1: Descriptive statistics for the world's cities and cities with subway systems in 2010

	World	Africa	Asia	Europe	N. America	S. America
All cities						
N	632	71	347	57	99	56
Mean population	2,427	2,091	2,509	1,921	2,441	2,825
Mean log(Pop.)	14.3	14.3	14.3	14.2	14.3	14.4
Mean Δ_t log(Pop.)	0.18	0.24	0.20	0.05	0.14	0.19
Mean Δ_t^2 log(Pop.)	-0.010	-0.013	-0.008	-0.005	-0.013	-0.015
Mean light gradient	-0.79	-0.85	-0.78	-0.72	-0.69	-0.96
Mean light intercept	11.0	10.5	10.8	10.8	10.8	12.7
Cities with subway in 2010						
N	138	1	53	40	30	14
Total stations	7,886	51	2,977	2,782	1,598	478
Total route km	10,672	56	4,210	3,558	2,219	627
Mean stations	57	51	56	70	53	34
Mean route km	77	56	79	89	74	45
Mean subway lines	4.5	2.0	4.1	5.8	4.7	2.6
Δ_t Stations	3.5	3.9	4.2	3.8	2.5	2.2
Mean log(Stations)	3.60	3.95	3.55	3.90	3.38	3.30
Mean Δ_t log(Stations)	0.23	0.30	0.26	0.22	0.21	0.23
Mean population	4,706	11,031	5,950	2,259	4,813	6,300
Mean log(Pop.)	14.93	16.22	15.15	14.37	15.05	15.34
Mean Δ_t log(Pop.)	0.11	0.12	0.14	0.04	0.12	0.17
Mean Δ_t^2 log(Pop.)	-0.011	-0.014	-0.012	-0.005	-0.013	-0.017
Mean light in 25km disk	122	212	117	95	170	109
Corr. lights & pop.	0.67		0.67	0.69	0.78	0.91
Mean light gradient	-0.72	-0.62	-0.78	-0.71	-0.58	-0.80
Mean light intercept	11.2	11.0	11.8	11.0	10.2	11.9

Note: Population levels reported in thousands. Lights data are based on radiance calibrated lights at night imagery. All entries describing levels report 2010 values. Entries describing changes are averages over the period from 1950 to 2010.

the entire world. Most extant applications of the lights at night data in economics rely on the "DMSP-OLS Nighttime Lights Time Series".[9] These data are available annually from 1992 until 2012. Each of these lights at night images is a composite constructed from many raw satellite images and the value for each cell reflects average light intensity, over all cloud free images, on a scale of 0-62 with 63 used as a topcode. Since most large cities, particularly in the developed world contain large topcoded regions near their centers, these data are of limited use for studying the internal structure of the large wealthy cities where most subways are located. We instead exploit 'radiance calibrated lights at night data',[10] collected during times when the satellite sensor was set to be less sensitive. These data are less able to distinguish dim light sources, but are able to measure variation in light within regions that are topcoded in DMSP-OLS version. Fewer cross-sections of

[9] Available from `http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html` (October 2014).

[10] Downloaded in October 2014 from `http://ngdc.noaa.gov/eog/dmsp/download_radcal.html`. We are grateful to Alexi Abrahms for drawing our attention to these data.

the radiance calibrated lights are available but the available cross-sections (ca. 1995, 2000, 2005 and 2010) match up neatly with the last four cross-sections of our population data.

Lights at night data are of interest as a check on our population data. The lights at night data are measured consistently across cities and we can calculate city level measures of total light without reference to administrative boundaries. That is, the lights at night data are not subject to either of the two problems that we are concerned about for our population data. Since people light the places they live and work, more densely populated and more productive places are often brighter. More concretely, Henderson and Storeygard 2012 use the topcoded version of lights at night data to show that country level mean light intensities are a good proxy for GDP, a result that Storeygard (2017) confirms at the regional level for China.

The bottom panel of table 1 shows the correlation of the mean 2010 light intensity within 25km of a city center and 2010 population in subway cities. It is clear that lights provide some information about population, although this information is imperfect. Finally, we note that the lights at night data are difficult to interpret. While we can be confident that lights at night data are telling us something about the location of economic activity, we cannot know whether places are brighter because the people living there are richer or because the place is more densely populated.

C *Centralization*

We also use the lights data to describe urban centralization. The resolution of the radiance calibrated lights data we use is about 1km square. This is small enough to provide information about the way that cities are laid out, and inspection of figure 2 shows that the lights data reflect broad patterns of urban density.

In order to describe the 'centralization' of each city, we follow a long tradition in urban economics of calculating density gradients (e.g., Clark, 1951; Mills and Peng 1980). In our case, we estimate a light intensity gradient for every city-year to measure the rate at which density decays with distance from the center. To do this, we first calculate mean light intensity, for disks with radius 1.5km, 5km, 10km, 25km and 50k, around each city's centroid. These disks describe a series of doughnuts surrounding the center of each city. Let $x_i \in \{0.75\text{km}, 3.25\text{km}, 7.5\text{km}, 17.5\text{km}, 37.5\text{km}\}$ be the radii of the circles lying halfway between the inner and outer border of these doughnuts. For example, $x_i = 3.25$ lies halfway between the inner and outer radius of the doughnuts that extends from 1500m to 5km from a city's center. For each such doughnut, let y_i denote the average light intensity in the doughnut.[11] All together, for each city, we now have 5 pairs of light intensity and distance, (y_i, x_i).

To characterize the centrality of each city, we estimate the following regression

$$\ln y_i = A + B \ln x_i + \epsilon_i. \tag{3}$$

[11]We note that we do not make any adjustments for geographic features such as mountains or surface water when performing this calculation.

The coefficient B in this regression is the rate at which light decays with a change in distance from the center, and will be our measure of centrality for each city in each year. All else equal, a city with a more negative value of B sees its density decrease more quickly with distance from the center, and is therefore, 'more centralized'.

Table 1 reports sample mean values of A and B for the sample of all cities and subway cities. We see that the gradient for an average city is 0.79. Thus, density falls by 79% with a doubling of distance. Not too surprisingly, cities in Africa and South America are more centralized, while cities in North America are less centralized. Subway cities are slightly less centralized than cities without subways. For these cities, density falls by 72% with a doubling of distance. Interestingly, North American subway cities are particularly spread out, with a density gradient of 0.58.

D *Subways data*

We define a 'subway' as an electric powered urban rail that is completely isolated from interactions with automobile traffic and pedestrians. This excludes most streetcars, because they interact with vehicle traffic at stoplights and crossings, although we include underground streetcar segments. In order to focus on intra-urban subway transportation systems, we also exclude heavy rail commuter lines. We do not distinguish between surface, underground or aboveground subway lines as long as the exclusive right of way condition is satisfied. For the most part, our subways data describe public transit systems that would ordinarily be described as 'subways', e.g., the Paris metro and the New York city subway, and only such systems. As with any such definition, the inclusion or exclusions of particular marginal cases in our sample may be controversial.

On the basis of this definition, we assemble data describing the latitude, longitude and date of opening of every subway station in the world. We compiled these data manually between January 2012 and February 2014 using the following process. First, using online sources such as `http://www.urbanrail.net/` and links therein, together with links on wikipedia, we complied a list of all subway stations worldwide. Next, for each station on our list, we record opening date, station name, line name, terminal station indicator, transfer station indicator, city and country. Latitude and longitude for each station were obtained from GOOGLE maps. This process leads us to enumerate subway stations in 161 cities. Of these, 138 are large enough to appear in the UN World Cities Data and are the main subject of our analysis.[12]

We use our data to construct three measures of subway extent for each city-year. First, we count the number of operational stations in each year. Second, we count the number of operational subway lines in each city in each year. Finally, by connecting stations on each subway line by the shortest possible route, we approximate the route of each subway line. Taking the union of all such lines in a city approximates each city's network and calculating the length of this network gives us

[12]The 23 cities with subways in 2010 that do not occur in our population data because their population is too small are: Bielefeld, Bilbao, Bochum, Catania, Dortmund, Duisburg, Dusseldorf, Essen, Frankfurt, Genova, Hannover, Kitakyushu, Kryvyrih, Lausanne, Mulheim, Naha, Nuremberg, Palma, Perugia, Rennes, Rouen, Seville and Wuppertal.

the length of each system. In this way we arrive at our three primary measures of subway extent for each city-year; operational stations, operational lines and route kilometers.

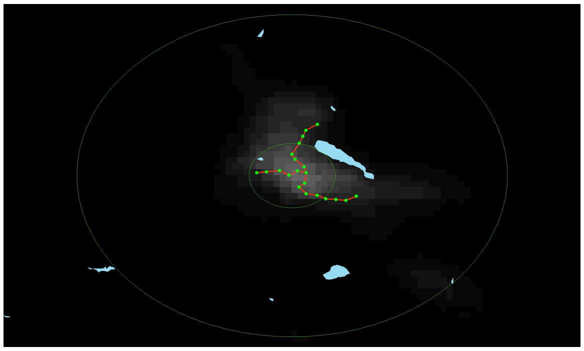
Figure 2 illustrates our subway data for six cities. The figure shows all stations operational prior to 2010 as dots. The network maps, on which the 2010 calculation of route km is based, are shown as connecting lines. In each panel of the figure, the large(small) circle or ellipse describes a circle of 25(5)km radius to show scale. This circle is distorted in Northerly cities as a consequence of our map projection. To show the configuration of each city, the background shows lights at night in 2010. In the top row, with 2010 populations of 1.1m and 0.9m Tibilsi (Georgia) and Toulouse (France) are among the smallest cities in our sample to have subways. In 2010 their subway systems consist of 21 and 37 stations, and 27 and 28 route km. In the middle row, Boston and Singapore have populations of 4.7m and 5.1m, near the 4.7m mean for subway cities. Their subway systems consist of 74 and 78 stations and of 88 and 111 route km, which makes both systems somewhat larger than both world and the relevant continental averages. The bottom row of figure 2 shows two of the largest cities in our sample, Mexico City and Beijing. The population of Mexico City in 2010 was just over 20m against about 15m for Beijing. Their subway systems contained 147 and 124 stations and consisted of 182 and 209 route kilometers.

Figure 2 reveals that in each of the six cities only a small portion of the city is within walking distance of a subway and the catchment area of the subway is centrally located. This is typical. An average city in our sample has about 57 stations. Of these, about 9% are within 1500m of the center, about 29% are between 1500m and 5km of the center, about the same share lie between 5 and 10km and between 10 and 25km. Just 7% of stations are beyond 25km from the center. Since the area to be served expands quadratically, this means that subways per square kilometer decreases rapidly with radial distance. In an average subway city, there are 0.67 stations per km^2 within 1500m of the center, 0.22 stations per km^2 between 1500m and 5km from the center, 0.07 stations per km^2 between 5 and 10km from the center, and 0.001 stations per km^2 between 10 and 25km from the center. Thus, in an average city, the preponderance of the subway system is located within 10km of the center and station density decreases rapidly with distance from the center. This is consistent with the argument that public transit preponderantly serves downtown cores in Glaeser, Kahn, and Rappaport (2008).

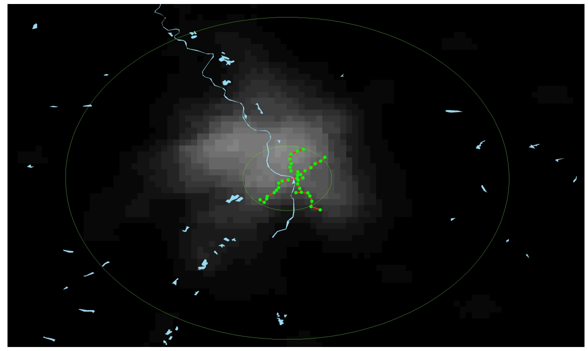
Close inspection of the network maps in figure 2 suggests that our networks probably diverge slightly from the actual network. The algorithm that we use to construct network maps connects all open stations on a subway line by the shortest possible route. Therefore, our measure of length is a measure of the route kilometers required to serve operational stations in each year rather than a literal measure of the length of track in the system.[13] While we regard the route kilometers

[13]Our algorithm will produce routes that diverge from the actual routes for four reasons. First, if pairs of stations are connected with curving track, the actual route will diverge from our straight line network. Second, if intermediate stations on a line open after the end points, then the algorithm will not include the intermediate stations on the network until they open. Third, we may mis-attribute stations to subway lines. Fourth, if a route is served by two or more sets of tracks — such as in New York city — then this replication is invisible to us.

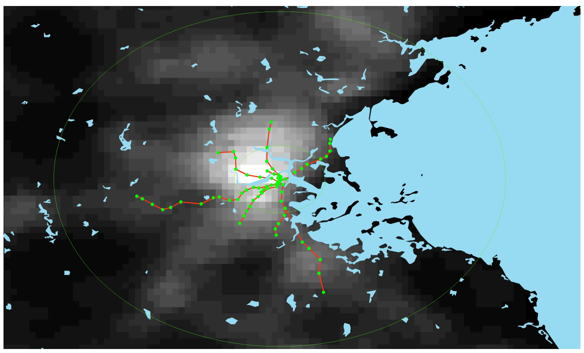
Figure 2: Lights and subways in 2010 for six cities



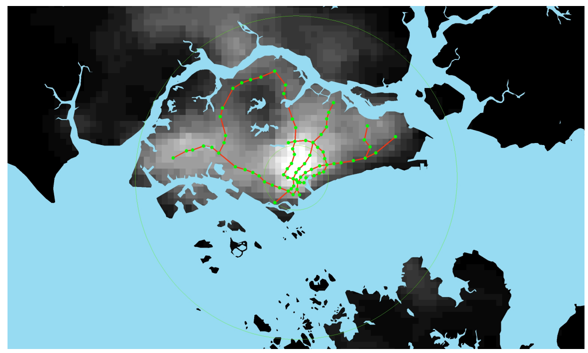
Tibilisi: 1.1m pop, 21 stations



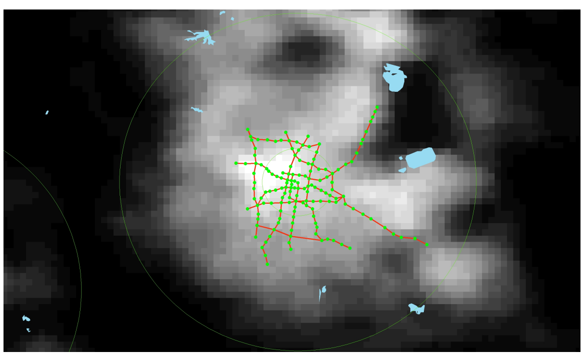
Toulouse: 0.9m pop and 37 stations



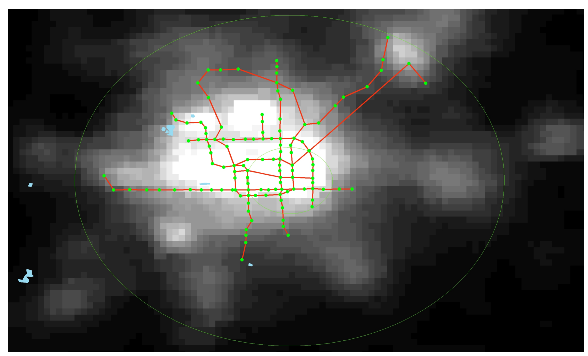
Boston: 4.7m pop, 74 stations



Singapore: 5.1m pop, 78 stations



Mexico City: 20.1m pop and 147 stations



Beijing: 15m pop and 124 stations

Note: Images show 2010 radiance calibrated lights at night, 2010 subway route maps, and all subway stations constructed prior to 2010. The gray/green ellipses in each figure are projected 5km and 25km radius circles to show scale and light gray/blue is water.

measure as being of considerable interest, we suspect it is a noisier measure of subway extent than is the count of operational stations. Given this, our investigation relies primarily on the count of operational stations to measure system extent, although our results are robust to the choice of subway measure.

Table 1 describes the world's subway systems in 2010. In 2010 in our sample of cities, there were 7,886 operational subway stations and 10,672 route kilometers of subways, divided across 138 operational systems. Of these 138 subway cities, 53 are in Asia, 40 in Europe, 30 in North America, 14 in South America and one in Africa. Asia, Europe, North America and South America account for 38, 35, 20 and 6 percent of all operational stations in 2010. The corresponding percentages of route kilometers are 39 for Asia, 33 for Europe, 21 for North America and 6 for South America. Thus, Asia has more systems than Europe, but a typical system in Europe has more stations and route kilometers. North America accounts for a small share of subway stations and route km, it contains a small number of systems and the average extent of these systems is between that of Asian and European systems.

Table 1 reveals substantial differences in the availability of subways across continents. Of the 347 large cities in Asia only 53, about 15%, have subway systems. In Europe, more than two thirds of large cities have subways, while in North America it is just less than one third. South America is a bit lower at 25%. Conditional on being in a subway city, the level of service also varies widely by continent. Cities are smaller and subway systems larger in Europe where there are 25,000 people per route km and 32,000 per station. These service levels are higher than those in North America and Asia and higher still than those in Asian and South American subway cities. Interestingly, although the share of North American cities with subways is much higher than in Asia, people per station and people per route km in subway cities are close for the two continents.

Two features of table 1 stand out. First, the huge gap in subway provision between Europe and the rest of the world. Second, the weak connection between mean city size and subway extent. In particular, Asia is home to the preponderance of the world's large cities while South America's cities are larger, on average, than those elsewhere. However, neither South America nor Asia is well provided with subways relative to Europe and North America. Indeed, Europe's cities are the smallest and slowest growing, and it is by far the best provided with subways.

Figure 1 illustrates the expansion of the world's subway systems over the past century. There were four subway systems in operation prior to or during 1860; Liverpool, Boston, London and New York. The "L" opened in Chicago in 1892 and The Paris Metro opened in 1900. Both the aggregate world data and the continental data, except for Asia, show a first wave of subway construction between the two world wars and a second wave beginning in the 1970s and continuing to 2010. The growth of Asian subways begins in the 1970s and has accelerated since. Except for North America, expansion of subway systems and increases in the number of subway cities track each other closely. In 2010, the 1,169 subway stations operating in the US were spread across 21 cities. However, 489 of these stations were in New York. Chicago is the second largest system at 142 stations. On average, the remaining 19 US subway cities have just 29 stations each, just over half the sample average.

Table 2: Public transit ridership (2010)

	Annual ridership (millions of rides)				Annual ridership per capita (rides per person per year)				Population (millions)		
	Mean	Std. dev.	0.10	0.90	Mean	Std. dev.	0.10	0.90	Mean	Cities	Countries
Subway	377	640	18	1,110	69	76	8	127	5.6	77	34
Bus	242	343	26	697	67	80	12	170	4.0	40	17
Bus $\mid$ Subways>0	256	315	36	584	74	86	14	145	4.5	31	17

Source: American Public Transportation Association, public transit agencies, municipal and state-level statistics agencies, and railway companies.

E ***Public transit ridership data***

We collected panel data on public transit ridership for the cities in our database from publicly available sources and reports. We were able to obtain data on 77 subway systems and 40 bus transit systems.[14] Table 2 shows ridership descriptive statistics for subways and buses in 2010. Bus systems provide on average 240 million trips per year, whereas subways provide on average 380 million trips per year. In per capita terms (columns 5-8), subways and buses are about equally important in terms of rides per person per year. This is true not only when comparing averages, but also when comparing cities for which both types of ridership information are available.

5. The relationship between subways and population

We now turn to a description of the relationship between subways and population. Figure 3 shows the relationship in 2010 between city size and the incidence of subway systems for all of the cities in our sample excluding Tokyo.[15] The horizontal axis gives city population by 0.5m bin and the vertical axis gives the proportion of cities with subways for each bin. We split our sample of cities into rich and poor country cities on the basis of the IMF advanced economy list for 2012.[16] Grey squares and black triangles indicate the share of rich and poor country cities with subways. The markers are spaced irregularly along the horizontal axis because some population bins are empty. The solid line is a smoothed plot of subway frequency in rich country cities and the dashed line is the corresponding plot for poor country cities.[17]

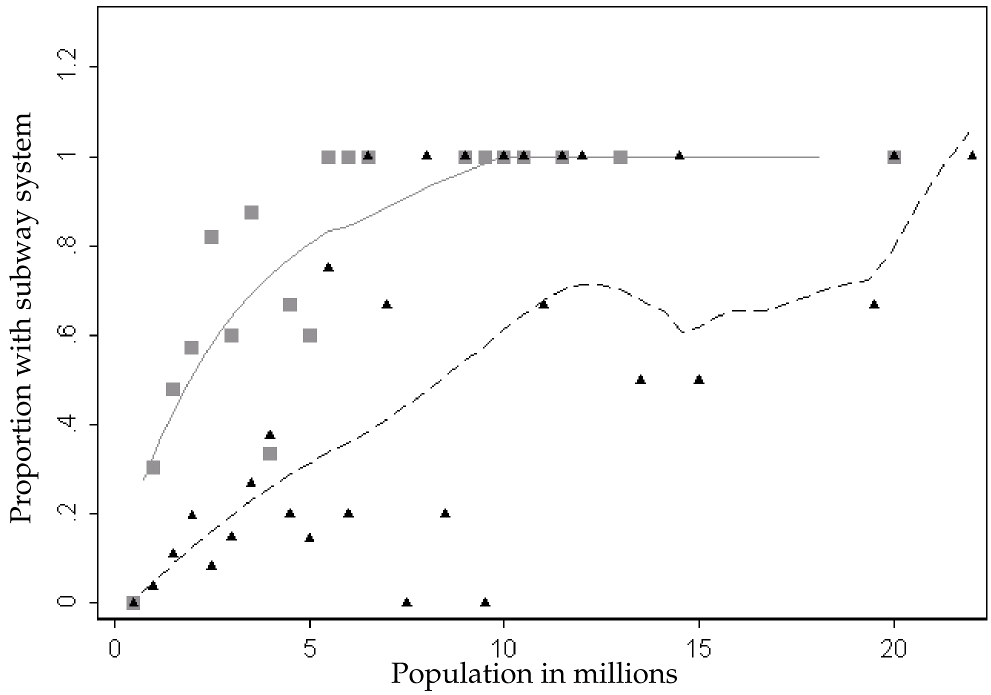
[14] Information on bus ridership by year is only reported by integrated transit systems, something that is not common in developing countries. In particular, we have no bus ridership data for cities in Africa and South America.

[15] At 36 million people, Tokyo is nearly twice as large as the second largest city. We omit it from the figure to improve legibility.

[16] These rich countries are: Australia, Japan, New Zealand, the United States, Canada, Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Israel, Italy, Netherlands, Norway, Portugal, Singapore, South Korea, Spain, Sweden, Switzerland and the United Kingdom.

[17] More specifically, both lines are kernel weighted local polynomial regressions.

Figure 3: Proportion of cities with subways systems by population for two income classes



Note: Gray squares correspond to rich country cities and black triangles to poor country cities. See footnote 16 for the list of countries.

There are no rich country cities with population above 5m without a subway system and subways are common even among rich country cities with populations in the 1m-5m range. Subways are relatively rare among developing country cities with populations less than about 5m and their frequency increases more or less smoothly with city size.

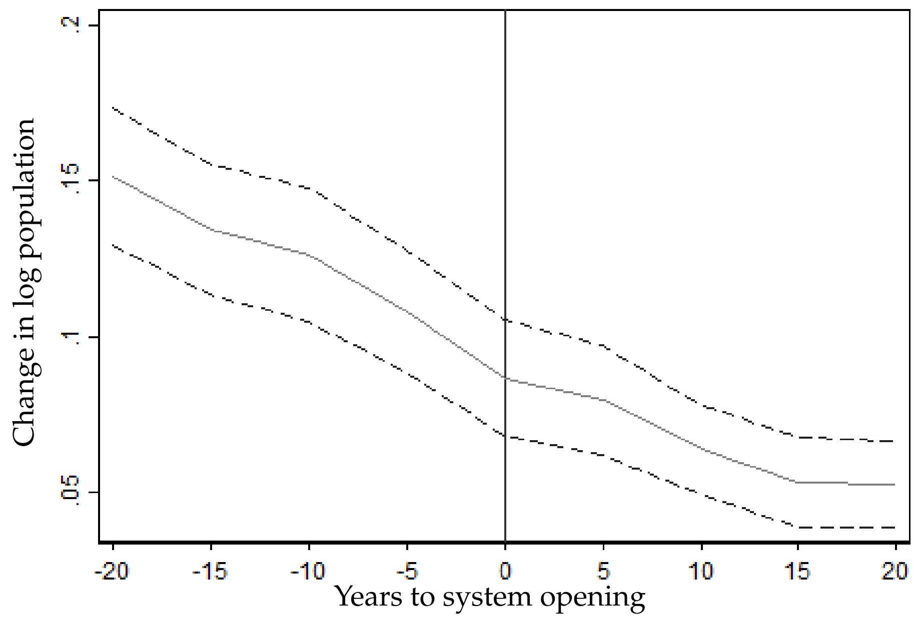
Table 3 describes the largest 90 cities in our sample as of 2010. For each city, the table reports population, the count of operational stations and the number of stations per 100,000 of population. Despite the strong relationship between city size and the presence of a subway system that we see in figure 3, table 3 suggests that the relationship between population and subways is nuanced. In particular, none of the four cities larger than New York has even half as many subway stations. Looking down the list, we see that such reversals are common and do not simply reflect rich and poor country differences. Consistent with this, the raw correlation between operational stations and population in 2010 is about 0.58. While subways are clearly more common in big cities, the relationship between system extent and city size is noisy. Because some of the world's largest cities have no subway system to speak of, table 3 suggests that subway capacity may not be a binding constraint on city size.

Table 3: Population and subway stations for the world's 90 largest cities as of 2010.

City Name	Pop.	Stations	Stations pp.	City Name	Pop.	Stations	Stations pp.
Tokyo	36,933	255	0.69	Ho Chi Minh City	6,189	.	.
Delhi	21,935	128	0.58	Miami	5,971	22	0.37
Mexico City	20,142	147	0.73	Santiago	5,959	93	1.56
New York	20,104	489	2.43	Baghdad	5,891	.	.
Sao Paulo	19,649	62	0.32	Philadelphia	5,841	64	1.10
Shanghai	19,554	239	1.22	Nanjing	5,665	54	0.95
Mumbai	19,422	.	.	Haerbin	5,496	.	.
Beijing	15,000	124	0.83	Barcelona	5,488	137	2.50
Dhaka	14,930	.	.	Toronto	5,485	69	1.26
Kolkata	14,283	23	0.16	Shenyang	5,469	22	0.40
Karachi	13,500	.	.	Belo Horizonte	5,407	19	0.35
Buenos Aires	13,370	76	0.57	Riyadh	5,227	.	.
Los Angeles	13,223	30	0.23	Hangzhou	5,189	.	.
Rio de Janeiro	11,867	35	0.29	Dallas-Fort Worth	5,143	.	.
Manila	11,654	43	0.37	Singapore	5,086	78	1.53
Moscow	11,472	168	1.46	Chittagong	5,069	.	.
Osaka	11,430	125	1.09	Pune	4,951	.	.
Cairo	11,031	51	0.46	Atlanta	4,875	38	0.78
Istanbul	10,953	12	0.11	Xi'an, Shaanxi	4,846	.	.
Lagos	10,788	.	.	Saint Petersburg	4,842	63	1.30
Paris	10,516	299	2.84	Luanda	4,790	.	.
Guangzhou	10,486	123	1.17	Houston	4,785	.	.
Shenzhen	10,222	47	0.46	Boston	4,772	74	1.55
Seoul	9,751	360	3.69	Washington, D.C.	4,634	86	1.86
Chongqing	9,732	.	.	Khartoum	4,516	.	.
Jakarta	9,630	.	.	Sydney	4,479	.	.
Chicago	9,545	142	1.49	Guadalajara	4,442	17	0.38
Lima	8,950	16	0.18	Surat	4,438	.	.
London	8,923	267	2.99	Alexandria	4,400	.	.
Wuhan	8,904	25	0.28	Detroit	4,364	12	0.27
Tianjin	8,535	36	0.42	Yangon	4,356	.	.
Chennai	8,523	.	.	Abidjan	4,151	.	.
Bogota	8,502	.	.	Monterrey	4,100	32	0.78
Kinshasa	8,415	.	.	Ankara	4,074	12	0.29
Bangalore	8,275	.	.	Shantou	4,062	.	.
Bangkok	8,213	51	0.62	Salvador	3,947	.	.
Hyderabad	7,578	.	.	Melbourne	3,896	.	.
Lahore	7,352	.	.	Porto Alegre	3,892	17	0.44
Tehran	7,243	54	0.75	Phoenix	3,830	.	.
Dongguan	7,160	.	.	Montreal	3,808	68	1.79
Hong Kong	7,053	54	0.77	Zhengzhou	3,796	.	.
Madrid	6,405	239	3.73	Johannesburg	3,763	.	.
Chengdu	6,397	16	0.25	Brasilia	3,701	27	0.73
Ahmadabad	6,210	.	.	Recife	3,684	28	0.76
Foshan	6,208	.	.	San Francisco	3,681	48	1.30

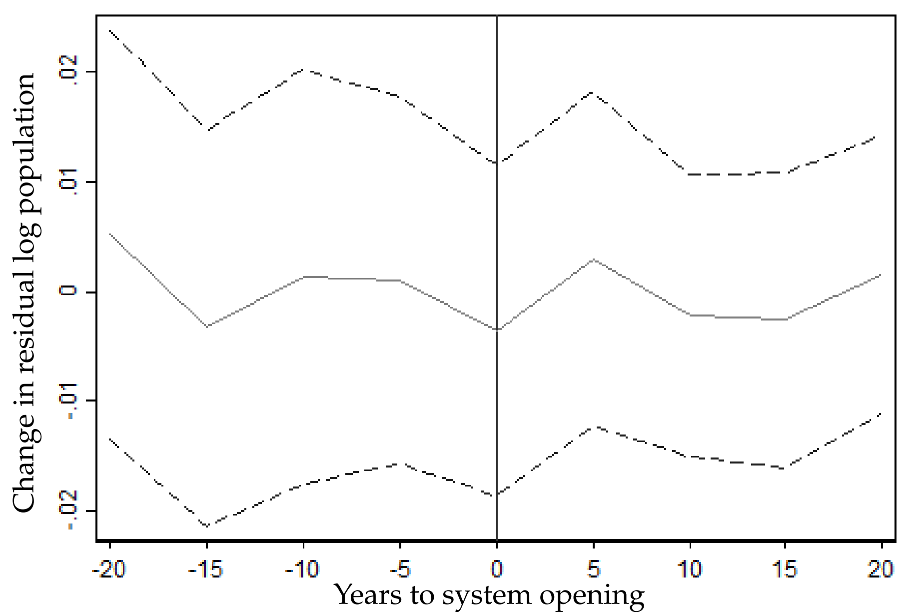
Note: Populations in thousands. Subway stations per person is per 100,000 residents.

Figure 4: Subway system opening and population growth (constant sample)



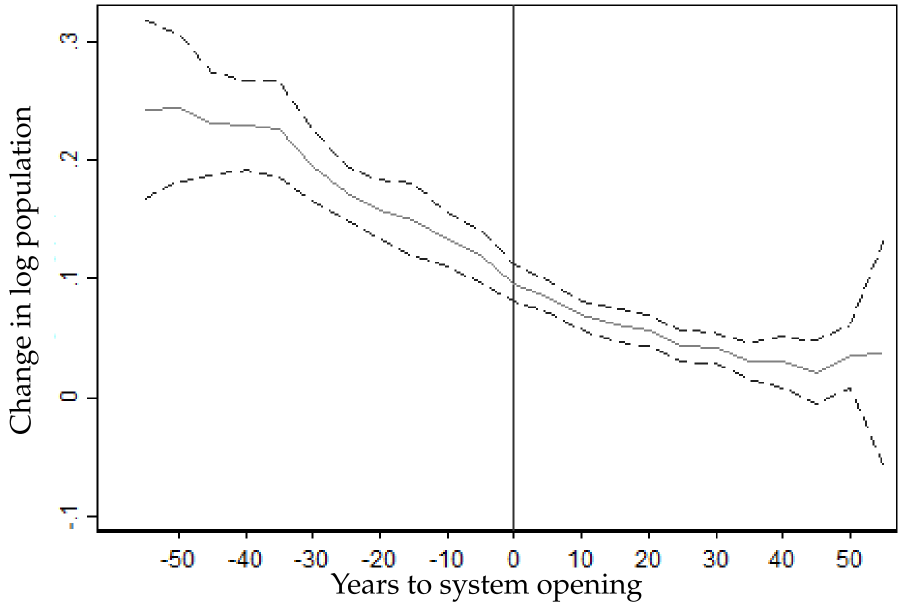
Note: The graph depicts mean change in city log population according to time to system opening. $t = 0$ which indicates the year in which a city's subway system was inaugurated. We impose a constant sample of cities on either side of $t = 0$. Graph based on constant sample of 61 cities.

Figure 5: Subway system opening and population growth (constant sample)



Note: The graph depicts residuals from a regression of change in city log population against continent and year fixed effects using the same sample of cities as Figure 4. Residuals from the regression are averaged conditional on time from subway opening and shown in the graph. $t = 0$ indicates the year in which a city's subway system was inaugurated.

Figure 6: Subway system opening and population growth (non-constant sample)



Note:The graph depicts mean change in city log population according to time to system opening. $t = 0$ indicates the year in which a city's subway system was inaugurated. Graph is based on a sample of 115 cities.

Table 4: Mean city-year population growth rates by time to a subway expansion

	5 year period $\Delta log(population)$					
Event type	Two periods before expansion	Period before expansion	Subway expansion period	Period after expansion	Two periods after expansion	N
Panel a:	Raw growth rates					
1			0.063	0.054***		138
2			0.078	0.067**	0.064**	60
3		0.090***	0.073			204
4	0.120**	0.107***	0.083			141
5		0.075***	0.061	0.052**		64
Panel b:	Growth rates relative to expansion period controlling for continent and year fixed effects					
1				−0.001		138
2				−0.001	−0.009	60
3		0.006*				204
4	0.013*	0.012*				141
5		0.009*		−0.007		64

Notes: Each row in panel (a) shows growth rates of cities in consecutive time periods. Event type 1 is a period of subway expansion (in the middle column) followed by a period with no expansion. Event type 2 is a period of expansion followed by two consecutive 5 year periods with no expansion. Event type 3 is a period with no subway expansions leading up to a period with expansions, and so on. Each row in panel (b) shows the difference in growth rates of cities (relative to a period of expansion) in consecutive time periods from a regression controlling for continent and year fixed effects. Stars indicate a significant difference of growth rate compared to period an expansion period. *** 1%, ** 5%, * 10% significance respectively.

We now turn to an investigation of what happens to a city when its subway system is inaugurated. Figure 4 presents three panels describing the relationship between changes in population and the introduction of a subway system in a city using event study graphs.

The top panel of figure 4 shows the average population growth rate of cities as a function of the time since their subway system opened.[18] This figure is based on data describing the 61 cities that opened their subway between 1970 and 1990, the set of cities for which we can calculate population growth rates both for 20 years before and after their subway opens. This figure shows that the average population growth rate during the five years following the opening of a subway system is about 8%. During the five year period preceding a subway opening by five years, the average population growth rate is about 12%. During the 20 years before and after a subway opening, the average city in our sample sees its growth rate decrease and there is no obvious change in this trend around the opening of the subway system.

The decrease in population growth rates visible in the top panel reflects a sample-wide decrease in growth rates. It may be that this downward trend masks increases in growth rates associated with subway system openings. Figure 5 investigates this possibility by controlling for each period's mean growth rate. Using the same sample as in the top panel, for each year we calculate each city's residual growth rate from a regression of growth rates on continent and year dummies. We next calculate the average of these residuals conditional on time from subway opening. Unsurprisingly, this process removes the downward trend that we see in the first three panels. Perhaps more surprisingly, it still does not show a systematic change in growth rates following subway system inaugurations.

Figures 4 and 5 show that city population growth rates do not increase during the 20 year period following the opening of a subway system. As we discuss in section 2, the literature documents effects of subways on within city outcomes over much shorter periods and the effects of other types of infrastructure on city level outcomes over a 10-20 year horizon. Thus, the 40 year period illustrated in Figures 4 and 5 should be long enough to reveal whether growth rates respond to a subway system opening. Nevertheless, in figure 6 we use our entire sample of cities and investigate population growth rates over the longest time period that our 60 year sample allows, 55 years. This figure suggests that the pattern we see in figure 4 extends nearly 55 years before and after a subway opening, although our estimates become noisier as the time from the subway opening approaches 55 years.

To check for differences across regions in the relationship between urban growth and subways, we produce analogous figures continent by continent (not shown). Remarkably, each of the continents shows a similar pattern. Urban population growth rates decrease in the period around

[18] The horizontal axis of each panel is time in years since a subway system in a city is inaugurated, with negative values indicating years prior and conversely. The vertical axis indicates the mean change in log population — the population growth rate — for all cities during the five year period ending *t* years before or after the subway opening. The solid line plots the mean growth rate and dashed lines give upper and lower 95% confidence bounds. These are local bounds constructed by connecting upper and lower 5% bounds at each year.

subway openings and there is no obvious sign of a change in this trend at the time a subway opens. The only qualification of this statement applies to Europe, where there is a statistically insignificant positive deviation from trend around the opening of a subway system. We also produced analogs to figures 4 and 5 where we restrict attention to cities with population above 1m in 1970. This eliminates the small fast growing cities that qualify for the sample late in the sampling period. The resulting figures are difficult to distinguish from those presented here in figure 4.

Figures 4, 5 and 6 describe population growth rates as time varies relative to the date of a subway system *opening*. In Table 4 we turn our attention to the relationship between subway *expansions* and growth rates. The top row of panel (a) describes 138 city-year pairs where a city-year with a subway expansion is followed by a city-year without a subway expansion (recall that we use observations every five years so technically the table reflects quinquennial city-periods). On average, the growth rate in city-years with an expansion is 0.063, and in the subsequent city-year, without an expansion, it is 0.054. A t-test of the difference between the two means indicates that they are statistically different with high probability. In short, population growth rates are lower following a subway expansion than during one.

The remaining three rows of panel (a) of table 4 perform similar calculations for slightly different sets of city-years. In row two we consider the 60 city-year triples for which we observe a subway expansion followed by two city-years without an expansion. As for row 1, we see that growth rates decrease following a subway expansion and that the decrease in growth rate is statistically different from zero. In the third row we consider the 204 pairs of city-years where a subway expansion follows a city-year without an expansion. The mean growth rate for city-years preceding a subway expansion is larger than for city-years with an expansion, and this difference is statistically different from zero. The fourth row of table 4 considers the 141 triples of city-years where a subway expansion is preceded by two years without an expansion. Again, we see that city growth rates decrease in the years leading up to a subway expansion. The last row of panel a in table 4 considers the 64 triples of city-years for which a subway expansion follows and precedes city-years without expansions. The pattern of the other rows is preserved. Population growth rates are higher before a subway expansion and lower after, and this trend is statistically different from zero.

Similarly to the middle of figure 4, panel (b) of table 4 replicates the results of panel (a), but controls for continent and year fixed effects. Specifically, the values reported in panel (b) of table 4 are regression coefficients β from the regression,

$$\Delta log(\text{Pop}_{it}) = \alpha_t + \phi_j + \sum_{k=-2}^{2} \beta_k \cdot I(\text{Time to Expansion Indicators}_{it} = k) + \epsilon_{it},$$

where ϕ_j refers to continent dummies and the excluded category for the time to expansion indicators is $k = 0$. Standard errors are clustered at the city level, and we use the same samples as in the top panel. We test whether the various time to expansion coefficients are different from the year zero coefficient using a robust F-test. Panel (b) of the table shows that even after we control for year and continent fixed effects, subway expansions are not associated with a measurable increase in population growth rates.

6. Econometric model

The descriptive evidence presented so far indicates a positive cross-sectional relationship between the extent of a city's subway network and its population. Larger cities have more extensive subway networks. On the other hand, time series evidence suggests that changes to subway networks do not affect the population of cities. These facts suggest that large cities build and expand subway networks but that these networks do not cause changes in subsequent population growth. To establish this causal interpretation of the patterns we see in the raw data, we must address two main inference problems, confounding dynamics and omitted variables.

A *The problem of confounding dynamics*

Confounding dynamics arise if subway extent and population evolve such that subways open or expand in years that are, on average, different from other years. Many examples are possible. Cities may tend to build and open subways as some constraint to their growth begins to bind and their growth is slowing. In this case, these cities might have seen a dramatic decrease in growth had they failed to construct a subway but manage to maintain their growth by adding to their networks. Alternatively, city population may naturally decrease when subways open and construction workers leave, and positive effects of subways on growth just offset this loss.

More generally, this class of problems arises when there is some series of population shocks that systematically precedes an expansion of the subway network and confounds naive estimates of the relationship between subway expansion and growth. Describing the problem in this way suggests two possible responses. The first is simply to control for the history of population growth in the period leading up to a subway expansion. In this way, we can estimate the effect of subways, holding constant their population growth during the preceding periods. The second is to find an instrument that predicts subway expansions but is conditionally orthogonal to the hypothetical sequence of confounding population shocks.

As we will see, subway systems grow along a predictable trajectory (see appendix figure A.1) and so long lags of subway extent are good predictors of current subway growth (See figure A.2).[19]

[19]Indeed, the growth of subway systems is surprisingly predictable. We can only speculate as to why this might be. One explanation that would lead to the pattern we observe is that every city's administration tries to show competence by adding a subway line to the system. Another possibility is suggested by Gomez-Ibanez's (Gomez-Ibanez, 1996) history of the Boston subway. In this history, Gomez-Ibanez documents a series of expansions, partly motivated by the need to expand the tax base on which to draw for subsidies for the system.

By construction, long lags of subway extent pre-date the hypothetical confounding recent history of population growth, and hence should satisfy the relevant exclusion restriction.[20]

In the remainder of this section we develop an econometric model that allows us to make this intuition precise and will form the basis for subsequent estimations. To begin, index the set of observed cities by i and the set of observed years by t. Let y_{it} denote an outcome of interest for city i in year t. Depending on context, y will be population, mean light intensity within 25km of the city center, centrality or ridership. Let s_{it} denote a measure of subway extent in city i in year t, usually the number of operational stations but sometimes the number of operational subway lines or route kilometers. Let x_{it} denote a vector of time varying city level covariates, most often country level population, GDP per capita and continent specific year indicators, and z_i a time-invariant vector of city level controls. The operator Δ denotes first differences, $\Delta x_t = x_t - x_{t-1}$.

We do not have a strong prior over whether subways should affect city population levels or growth rates additively or multiplicatively. However, plots of population growth against subway growth in both logarithms and levels clearly suggest that the logarithmic forms better represent the data. Given this, quantities are typically in logarithms and where necessary we add one to variables to facilitate this transformation. This also allows us to interpret regression coefficients as elasticities.

In light of the differences between the time series and cross-sectional relationship between subways and population growth, we are also concerned that cities have time invariant characteristics correlated with size and subway extent. The following system, while too stark to be defensible, formalizes this problem and allows a discussion of how our lagged subways instrument addresses the problem of confounding dynamics.

$$y_{it} = A_1 s_{it} + c_i + \epsilon_{it} \tag{4}$$

$$s_{it} = B_1 s_{it-k} + d_i + \eta_{it}, \tag{5}$$

where A_1, the "outcome elasticity of subway extent", is the parameter of interest and k is a positive integer. In words, population depends on contemporaneous subways, a city specific intercept and a random disturbance. Subways at t depend on subways at period $t-k$, a city specific intercept and a random disturbance.

Written this way, it is natural to consider using s_{it-k} as an instrument for s_{it}. This is subject to two objections. First, this system of equation commits us to a particular dynamic structure for the relationship between subways and population. It is natural to wonder whether this dynamic structure is correct. In our estimations we consider alternative dynamic structures for our data. Second, unobserved time invariant determinants of subway construction are probably related to unobserved time invariant determinants of growth. That is, $cov(c_i, d_i) \neq 0$. It follows that, because

[20]It is worth pointing out that our use of long lags of subway system status as a source of quasi-random variation is conceptually similar to the use of historical networks as instruments for highways, e.g., Duranton and Turner (2012). The difference is that we here implement a panel data model, which looks quite different from the existing literature on roads, and our 'long lags' are recent relative to the historical network variables used in Duranton and Turner (2012).

s_{it-k} also depends on d_i, we should not expect $cov((c_i + \epsilon_{it}), s_{it-k}) = 0$. That is, the dynamic structure described by equations (4) and (5) requires that s_{it-k} be correlated with unobservables in the population equation, and thus, that it is not a valid instrument in this context.

As a first response to this problem, first difference equations 4 and 5 to get

$$\Delta y_{it} = A_1 \Delta s_{it} + \Delta \epsilon_{it} \tag{6}$$

$$\Delta s_{it} = B_1 \Delta s_{it-k} + \Delta \eta_{it}. \tag{7}$$

Differencing solves two problems. First, and as usual, it removes time-invariant unobservables from the first equation.[21] Second, after removing the city specific intercept from the population equation, the validity of lagged subways as an instrument for current subways hinges on the whether $cov(\Delta s_{it-k}, \Delta \epsilon_{it}) = 0$, or in words, on whether lagged change in subways is uncorrelated with current change in the time varying propensity to grow. This is simply a more technical statement of the intuition that motivates this instrumental variables strategy.[22]

Since $\Delta s_{it-k} = s_{it} - s_{it-k}$ and since the error term in equation 6 no longer includes c_i, as is standard in dynamic panel estimation, the same logic that justifies using Δs_{it-k} as an instrument also justifies using the component levels. In fact, we find that the levels have much better predictive ability in the first stage than do changes, and so we rely on lagged levels of log subways as our instruments.

The discussion above describes an econometric strategy based around using old subway system extent to instrument for current subway system growth. An alternative is to use lagged changes of population to instrument for current changes in subways. The basic logic of this approach is similar to that described above. However, lagged population levels and changes have less ability to predict current changes to subways than do lagged subway variables, so we organize our discussion and analysis around the lagged subways instruments.

The instrumental variable strategy articulated above responds to the possibility that subway construction reflects recent trends in population. A more direct approach to this problem is to simply control for lagged population, which we also do in the results section.

A related problem arises if both population growth and subway growth reflect some unobserved city specific time-varying factor. For example, it may be that poor administrations cause cities to grow slowly and build subway networks. In this case, our estimated effect of subways on population growth confounds the effects of bad municipal government with the effects of subways. To address this possibility, we would like to include fixed effects in the first differences regressions,

[21]While differencing solves one problem, it may create another. If $k = 1$ then both Δs_{it-1} and Δy_{it} involve terms for quantities for time $t-1$. If we are concerned about contemporaneous correlation of errors in the population and subway equations, then this creates an obvious problem. This is a classic problem in dynamic panel data and the conventional approach is to substitute s_{it-2} for Δs_{it-1} or to use longer lags.

[22]We note that the instrumental variables strategy described here is related to the one proposed by Olley and Pakes (1991), while the exogeneity condition of equation 7 is related to ideas developed in Arellano and Bond (1991).

or equivalently, city specific trend in the levels regressions, equations (4) and (5). To implement this estimator, we second difference equation (4).[23]

Summarizing, our econometric investigation will be organized around estimating the following system,

$$y_{it} = A_1 s_{it} + A_2 x_{it} + A_3 z_i + c_i + g_i t + \epsilon_{it} \tag{8}$$

$$s_{it} = B_1 s_{it-k} + B_2 x_{it} + B_3 z_i + d_i + h_i t + \eta_{it}. \tag{9}$$

This generalizes equations 4 and 5 in a number of ways. First, it allows for time-invariant control variables, z_i. Second, it allows for city specific trends and intercepts in both population and subways equations. Third, it allows for time varying controls, lags of y_i in particular. In practice, we predict current changes in subways with 20 or 40 year old subway extent, so that $k = 4$ or 8.

B *The problem of omitted variables*

We are concerned that subway expansions and population growth are correlated with some unobservable. For example, one can imagine that cities experiencing bouts of growth-inhibiting automobile congestion decide to build subways. If this is indeed the case, then we should observe different effects of subways on population growth in congested than in uncongested cities. In particular, we should observe that subway expansions in cities with low levels of congestion attract population but that subway expansions in congested cities do not (or conversely). If we find no heterogenous effects of subways by city congestion levels, this suggests that this particular omitted variable is not biasing our estimations.

A second possibility is that the effect of subways on growth may be heterogenous across fixed city characteristics. For example, subway expansions may attract population to cities that already have a substantial subway network coverage, such as Paris or New York, but not to cities such as Miami with small systems. We can test for this by looking for heterogenous effects by subway network coverage. If we find no heterogenous effects by subway network coverage, we interpret this as suggestive that this type of consideration is not leading to the null result.

More formally, we estimate the following regression

$$\Delta y_{it} = A_1 \Delta s_{it} + A_2(\Delta s_{it} \times x_i) + \Delta \epsilon_{it} \tag{10}$$

where x_i denotes the terminal value of some control variable omitted from our main specification.[24] The particular variables that we consider measure: topography; the terminal stock of roads; capital

[23] In principle, one could also implement our instrumental variables strategy in second differences. We experimented with this but found that lagged subways and population variables do not have much ability to predict current second differences of subways. Consequently, these regressions were not informative.

[24] We do not have a strong prior over whether or not the variable x_i should occur independently in this equation. It is conventional that it should do so, however, since this is a first difference regression and since the x_i's do not vary over time, the first difference of a regression in levels that included an independent x_i term would look like equation (10). As a practical matter, we report estimates of equation (10), but corresponding estimates that include and independent term in x_i do not lead to important differences in our estimates of the effects of subways on population growth.

status; post WWII subway system indicator; degree of centralization; road congestion levels; and an ease of doing business index, among others. The data sources and definitions for these variables are described in the data appendix.

7. Subways and population: Main estimation results

We proceed by estimating successively more complete and complex versions of equations (8) and (9). To begin, in table 5 we estimate equation (8) using OLS on pooled cross-sections. Such estimations result in unbiased estimates only if the time invariant determinants of subways and population are uncorrelated. This condition seems implausible. We expect that unobserved factors affecting the attractiveness of a city also affect its construction of subways, so we regard these estimations as primarily descriptive.

In column 1 of table 5 we regress the log of population on log of the count of operational subway stations. We use the entire sample of 632 cities for which we have population and subway data. Since our panel is complete for these two variables, we have a sample of $13 \times 632 = 8{,}216$ city-years. The subway elasticity of population is large. A 10% increase in a city's count of stations is associated with a 4.8% increase in population. Column 2 replicates this result, but controls for country level GDP and continent-by-year fixed effects, along with several time-invariant controls; a capital city indicator, and distances to the ocean, international boundary and nearest navigable river. We see that the coefficient on subways, while still large, decreases to 0.28. Our sample size decreases to 7,374 in this regression, primarily because a number of the countries covered by our sample, particularly those in the former Soviet Union, came into existence after 1950 and so country level GDP is not available.

Column 3 considers the same regression as column 2 but restricts attention to cities that had subways in 2010. This is the largest sample of cities that could possibly contribute to a first differences estimate of the effect of subways. This reduces our sample to 1,565 city-years but leaves the coefficient of subways almost unchanged. The sample of 137 cities used in column 3 includes some cities that were small in 1950 and grew quickly to cross the 750,000 threshold for inclusion in the UN World Cities Data. To investigate the importance of this sampling problem in column 4 we restrict attention to cities that were already large in 1970 (above 1 million).[25] The estimated coefficient with the sample restricted to large cities changes very little. Columns 5 and 6 replicate column 3, but consider alternative measures of subway extent, route kilometers and log subway lines. Coefficient magnitudes change approximately in proportion to the changes in the standard deviation of the subway measures.

Column 7 reports a regression similar to column 3, where our dependent variable is the logarithm of mean light intensity in a 25 km disk centered on the city. As in column 3, we restrict

[25]We experimented extensively with different sampling rules to investigate whether our results are driven by the small cities that grow rapidly to get into the sample. We could find no evidence that this is the case.

Table 5: Pooled cross sectional estimation

	All cities		Subway cities				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\ln(\text{pop}_t)$	$\ln(\text{pop}_t)$	$\ln(\text{pop}_t)$	$\ln(\text{pop}_t)$	$\ln(\text{pop}_t)$	$\ln(\text{pop}_t)$	$\ln(\text{Lights}_t)$
$\ln(\text{subway stations}_t)$	0.48***	0.28***	0.26***	0.22***			0.17***
	(0.02)	(0.03)	(0.03)	(0.03)			(0.03)
$\ln(\text{route km}_t)$					0.23***		
					(0.03)		
$\ln(\text{subway lines}_t)$						0.52***	
						(0.06)	
$\ln(\text{GDP}pc_t)$		0.31***	0.02	-0.04	0.03	0.01	0.37***
		(0.04)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)
$\ln(\text{COUNTRY POP}_t)$		0.17***	0.28***	0.22***	0.29***	0.27***	0.20***
		(0.03)	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)
Geographic controls	No	Yes	Yes	Yes	Yes	Yes	Yes
YearXContinent dummies	No	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. variable	13.35	13.44	14.48	14.82	14.48	14.48	4.67
Mean of subways regressor	0.38	0.40	1.88	2.18	1.99	0.79	3.06
SD subways regressor	1.15	1.17	1.92	1.96	2.05	0.91	1.49
R-squared	0.18	0.49	0.53	0.58	0.52	0.53	0.54
Number of cities	632	627	137	99	137	137	137
Number of subway cities	138	137	137	99	137	137	137
Number of periods	13	13	13	13	13	13	4
Observations	8216	7374	1565	1155	1565	1565	548

Notes: Dependent variable: Log population of metropolitan area in quinquennial period t (except last column - see (7) below). City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. Geographic controls are capital city dummy, log km to ocean, log km to land border, and log km to navigable river. (1)- Pooled cross section. (2)- Geographic controls, GDP pc control, country population, and year-by-continent dummies. (3)- Restrict sample to cities with subway by 2010. (4)- Restrict sample to large cities in 1970 (population > 1 million). (5)- Log route km of subways as main regressor. Sample is cities with subway by 2010. (6)- Log subway lines in system as main regressor. Sample is cities with subway by 2010. (7)- Dep. var. is log mean radiance calibrated lights in a 25km circle around the centroid of the city in quinquennial period t.

attention to cities with subways in 2010. Our sample of city-years is smaller than for population regressions because we have just four cross sections of lights data. We see that a one percent increase in subways is associated with a 0.17 percent increase in lights. This is close to our results for population and suggests that our population regressions are not driven by problems in the UN World Cities Data. In sum, table 5 confirms the conclusion of figure 3. Cities with more subways tend to be bigger. This relationship is robust to controls, sampling, the particular measure of subway extent and whether we measure city size with lights or population.

We now turn to first difference regressions. Table 6 presents first difference estimates of a version of equation (8) without city specific trends. We note that both first difference and within estimators are consistent estimators for equation (8) if the errors, ϵ_{it}, in each period are not correlated with the regressors in any period conditional on the unobserved fixed effect. Because our approach to estimating equations (8) and (9) revolves around differencing, we prefer the first differences estimator.[26]

Columns 3-6 in table 6 use the same sample of cities as column 3 of table 5, while columns 1 and 2 use the slightly larger sample available when we do not control for changes in GDP. In column 1, we report the results of regressing change in log population on change in the log of the count of operational stations. In column 2 we repeat this regression with continent specific year dummies. Like table 4 we see a negative relationship between subway expansions and population growth when we do not control for continent specific year effects, but that the relationship between subways and population is approximately zero once we include these controls. In column 3 we add controls for country level changes in GDP and population and in column 4 we restrict attention to large cities in 1970 (over 1 million). In every case, we estimate the effect of subways to be less than 0.01 with standard errors around 0.003. These are tiny effects, precisely estimated. In unreported results we estimate these same specifications separately for each continent and find virtually no heterogeneity across continents, indicating that these small coefficients are not masking across-continent heterogeneity.

In column 5 we control for our measure of bus ridership. We do this to address the following concern. Suppose that in every year that a city does not invest in subways, it invests in buses, and that buses and subways substitute perfectly for each other. In this case, years with subway expansions will be identical to years without, even though subways may be having an arbitrarily large positive effect on population growth. Our data allows us to deal with this particular concern by controlling for changes in bus ridership. Since the sample of cities and years for which we observe bus ridership is much smaller than the sample for which we observe subways and population, our sample of years and cities shrinks considerably. However, including this control does not lead to a positive effect of subways on population. In fact, the relationship is slightly negative.

In columns 6 and 7 we measure subway extent using route km and counts of subway lines. We still find very small and statistically insignificant effects. In column 8 we use the average light intensity in a disk of 25km centered on the city as our dependent variable. As with our other regressions, we find a much smaller effect than in the comparable cross-sectional regression, column 7 of table 5, in this case not distinguishable from zero.

In column 9 we replicate column 3 but use 10 year rather than five year intervals to construct our panel, while in column 10 we report a long difference regression where we conduct a cross-sectional regression of long differences of population on long differences of subways. Both point estimates are small negative numbers indistinguishable from zero at ordinary levels of confidence.

[26]The choice between the two estimators hinges on subtle differences in the errors. The first difference estimator is more efficient if ϵ_{it} is a random walk, while the within estimator is more efficient if the ϵ_{it} are i.i.d. (Ch. 10, Wooldridge 2001).

Columns 9 and 10 suggest that our first difference estimates are not an artifact of the frequency with which we sample the data.[27]

Summing up, first difference estimates are dramatically smaller than cross-sectional estimates. Not only are the estimates of the effect smaller than those in the cross-sectional estimates, but they are small in an absolute sense, often well under 1% and precisely estimated.

We now investigate the possibility of confounding dynamics. Columns 1 and 2 of table 7 replicate column 3 of table 6 while controlling for the second and third lag of population change, respectively. Our sample size drops slightly in these specifications because we observe lagged population for fewer city years than contemporaneous population. Like the corresponding first difference regression in table 6, these regressions indicate tiny and precisely estimated effects of subways on population growth. Because the first lag of population is mechanically endogenous in our first difference regressions, columns 1 and 2 of table 7 control for the second and third lags of population. Column 3, instead reports second difference regressions. If there are city specific trends, this regression will account for this. As in the first difference regressions, we see a tiny precisely estimated relationship between subways and population.

In the remainder of table 7 we turn attention to the instrumental variables regressions described in section 6. That is, we replicate the first difference regressions of columns 1 and 2, but use the fourth or eighth lag of subways as an instrument for the current change in subways. The appendix describes the first stage. As we see in appendix figure 1, subway systems grow predictably, and at a decreasing rate. Thus, given the extent of a subway system in any period, we can forecast the future, lower, growth rate quite accurately. This is demonstrated in table A.1 which presents first stage results predicting current subway system growth rate as a function of lagged subway extent and the controls that appear in the first two columns of table 7. We see that our instruments are not weak, and behave as we would expect given the profile of system growth that we see in figures A.1 and A.2. Given that the instrument for subway growth in period t is subway extent 20 or 40 years prior, our instrumented estimate provides a local average treatment effect for which identifying variation is obtained from lower ranked expansions, as the instrument excludes variation from the initial and usually highest priority subway stations. For example, subway stations built during the first decade of a system are on average 6.4km from the city center, whereas this distance increases to 7.6km and 8.5km in the second and third decades respectively.

In columns 4 and 5 of table 7 we replicate columns 1 and 2, but instrument for change in log subways with the fourth lag of log subways. In column 6 we replicate column 1 but instrument for change in subways with the eighth lag of log subways. The IV point estimates of the effect of log subways are slightly larger than the first difference estimates, but never above 2% and never

[27]In fact, the long distance estimates are sensitive to the choice of time period. For example, if we conduct a long difference regression from 1950-2010, we get a statistically significant positive relationship between subways and population. This result is driven entirely by two cities which grew rapidly over the whole period and built large subway systems between 2005 and 2010. Excluding these two cities restores a coefficient of about zero in this regression. For this reason, we regard the long difference estimates as less reliable than other estimates.

statistically distinguishable from zero.[28] As a robustness check, in column 7 we take the specification in column 4 but replace the control for lagged population growth with lagged population level. This results in a negative but insignificantly different from zero instrumented coefficient of subway expansions on city growth. The IV coefficient estimates are not uniformly positive and suggest that the effect of subway expansions on city growth is small and not distinguishable from zero. In sum, table 7 does not support the hypothesis that subways have a large positive effect on population growth that is masked by some confounding dynamic process.

We next consider models that allow for a distributed lag structure in our data. In column 1 of table 8, we replicate column 3 of table 6 and in columns 2-4 we substitute successively older lags of change in subways for the current value. Like the effects of current subways, the effects of lagged subways are tiny and precisely estimated. In column 5 we include the current change of subways and three lags and see that coefficients are virtually identical to those we obtain when we include subway variables one at a time. This suggests that our focus on the relationship between current subway expansions and current population growth is not leading us to miss some longer term effect of subways on population growth. These regressions also suggest that a subway expansion does not affect current or future rates of population growth.

In table 9 we turn attention to the problem of omitted variables using the strategy described in equation (10). In column 1 of table 9 we replicate the first difference regression from column 3 of table 6 for reference. In column 2 we include an interaction between subways and an indicator for above median mean slope within 25km of the city center. If we think that cities build subways when some topographical constraint on their development begins to bind, then we should expect cities more subject to such topographical constraints to respond differently to changes in subways than other cities. The results in column 2 do not support this intuition. Column 3 replicates column 2, but in place of the average slope, measures topographical constraints with the elevation range within 25km of the city center. Like column 2, the results in column 3 do not suggest that subways affect cities with difficult topography differently than flatter cities.

In column 4 we interact subway growth with an indicator for above median kilometers of highways in a 25km circle around the city. That the coefficients on the main effect and the interaction are zero suggests that subway growth does not have a differential impact depending on whether the city is well served by highways. In column 5 we include an interaction between an indicator for above median traffic congestion and subways. If we think that cities tend to build subways as traffic congestion begins to constrain their growth, then we should see congested and uncongested

[28] Although point estimates from IV and first difference estimates of the effect of subways are not distinguishable from zero at conventional levels, they are close, e.g., compare columns 2 and 5 of table 7. To the extent that IV estimates are larger than the first difference estimates, they suggests that subways assigned to city-years by the equilibrium process are assigned to city-years that are growing slightly more slowly than city-years selected at random.

Table 6: First differences estimation

Time periods:	Quinquennial panel								Decennial panel	Long Difference
Dependent variable:	$\Delta\ln(\text{pop}_t)$							$\Delta\ln(\text{Lights}_t)$	$\Delta\ln(\text{pop}_t)$	$\Delta\ln(\text{pop}_{50-00})$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\Delta\ln(\text{subway stations}_t)$	-0.011**	0.001	-0.002	0.006**	-0.022*			0.024	-0.007	
	(0.004)	(0.003)	(0.003)	(0.003)	(0.012)			(0.015)	(0.005)	
$\Delta\ln(\text{subway stations}_{50-00})$										-0.060
										(0.058)
$\Delta\ln(\text{route km}_t)$						-0.001				
						(0.003)				
$\Delta\ln(\text{subway lines}_t)$							-0.002			
							(0.008)			
$\Delta\ln(\text{Bus ridership}_t)$					0.035					
					(0.039)					
$\Delta\ln(\text{GDP}pc_t)$			0.201***	0.176***	0.006	0.201***	0.201***	0.643***	0.222***	
			(0.042)	(0.032)	(0.103)	(0.042)	(0.042)	(0.106)	(0.038)	
$\Delta\ln(\text{COUNTRY POP}_t)$			0.951***	1.121***	1.222***	0.951***	0.949***	0.260*	0.911***	
			(0.118)	(0.138)	(0.214)	(0.117)	(0.118)	(0.144)	(0.150)	
YearXContinent dummies	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. variable	0.113	0.113	0.111	0.098	0.057	0.111	0.111	0.027	0.110	1.234
Mean of subways regressor	0.25	0.25	0.26	0.26	0.10	0.27	0.10	0.36	0.27	0.12
SD subways regressor	0.69	0.69	0.71	0.69	0.38	0.75	0.26	0.82	0.72	0.24
R-squared	0.00	0.29	0.42	0.55	0.51	0.42	0.42	0.46	0.39	0.40
Number of cities	138	138	137	99	31	137	137	137	137	138
Number of subway cities	138	138	137	99	31	137	137	137	137	138
Number of periods	12	12	12	12	8	12	12	3	6	1
Observations	1656	1656	1428	1056	63	1428	1428	411	730	138

Notes: Sample is subway cities. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. (1)- No controls. (2)- Add year-by-continent dummies. (3)- Add change in log gdp and change in log country pop. controls. (4)- Restrict sample to large cities in 1970 (population > 1 million). (5)- Control for change in log bus ridership. (6)- Use change in log route km as main regressor. (7)- Use change in log subway lines as main regressor. (8)- Dep. var. is change in log mean radiance calibrated lights in a 25km circle around the centroid of the city. (9)- Ten year panel analysis. (10)- Long difference regression 1950-2000, control for $\Delta\ln(\text{GDP}pc_{50-00})$ and $\Delta\ln(\text{COUNTRY POP}_{50-00})$, not shown.

Table 7: Robustness to confounding dynamics

Time periods:	Quinquennial panel						
Estimation:	OLS			IV			
Dependent variable:	$\Delta \ln(\text{pop}_t)$		$\Delta^2 \ln(\text{pop}_t)$	$\Delta \ln(\text{pop}_t)$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln(\text{subway stations}_t)$	-0.006	-0.006		0.018	0.016	0.014	-0.036
	(0.004)	(0.003)		(0.011)	(0.010)	(0.015)	(0.035)
$\Delta^2 \ln(\text{subway stations}_t)$			-0.003				
			(0.002)				
$\Delta \ln(\text{pop}_{t-2})$	0.553***	0.599***		0.545***	0.600***	0.546***	
	(0.052)	(0.113)		(0.053)	(0.119)	(0.053)	
$\Delta \ln(\text{pop}_{t-3})$		-0.059			-0.068		
		(0.082)			(0.087)		
$\ln(\text{pop}_{t-2})$							-0.040***
							(0.012)
$\Delta \ln(\text{COUNTRY POP}_t)$	0.465***	0.446***		0.434***	0.415***	0.438***	0.868***
	(0.058)	(0.045)		(0.061)	(0.049)	(0.061)	(0.123)
$\Delta \ln(\text{GDP}pc_t)$	0.128***	0.124***		0.126***	0.122***	0.126***	0.204***
	(0.025)	(0.023)		(0.024)	(0.022)	(0.024)	(0.045)
$\Delta^2 \ln(\text{COUNTRY POP}_t)$			0.301**				
			(0.100)				
$\Delta^2 \ln(\text{GDP}pc_t)$			0.067**				
			(0.022)				
YearXContinent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. variable	0.098	0.091	-0.010	0.098	0.091	0.098	0.106
Mean of subways regressor	0.29	0.31	0.02	0.29	0.31	0.29	0.27
SD subways regressor	0.74	0.77	1.01	0.74	0.77	0.74	0.72
R-squared	0.61	0.60	0.11	0.59	0.58	0.60	0.46
Number of cities	137	137	137	137	137	137	137
Number of subway cities	137	137	137	137	137	137	137
Number of periods	10	9	11	10	9	10	11
F-stat excluded instrument				132.36	147.51	153.49	216.75
Observations	1235	1124	1291	1235	1124	1235	1344

Dependent variable: Change in log population of metropolitan area in a 5 year period. Sample is subway cities. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. (1)- First differences controlling for $\Delta \ln(\text{pop}_{t-2})$. (2)- First differences controlling for $\Delta \ln(\text{pop}_{t-2})$ and $\Delta \ln(\text{pop}_{t-3})$. (3)- Second differences regression. (4)- Instrument $\Delta \ln(s_t)$ with $\ln(s_{t-4})$ controlling for $\Delta \ln(\text{pop}_{t-2})$. (5)- Instrument $\Delta \ln(s_t)$ with $\ln(s_{t-4})$ controlling for $\Delta \ln(\text{pop}_{t-2})$ and $\Delta \ln(\text{pop}_{t-3})$. (6)- Instrument $\Delta \ln(s_t)$ with $\ln(s_{t-8})$ controlling for $\Delta \ln(\text{pop}_{t-2})$. (7)- Instrument $\Delta \ln(s_t)$ with $\ln(s_{t-4})$ controlling for $\ln(\text{pop}_{t-2})$.

Table 8: First Differences - Distributed Lag Models

Time periods:	Quinquennial panel				
Dependent variable:	$\Delta \ln(\text{pop}_t)$				
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln(\text{subway stations}_t)$	-0.002				-0.002
	(0.003)				(0.004)
$\Delta \ln(\text{subway stations}_{t-1})$		-0.002			-0.002
		(0.003)			(0.003)
$\Delta \ln(\text{subway stations}_{t-2})$			-0.003		-0.003
			(0.003)		(0.003)
$\Delta \ln(\text{subway stations}_{t-3})$				-0.005	-0.006
				(0.003)	(0.004)
$\Delta \ln(\text{GDP}pc_t)$	0.201***	0.201***	0.200***	0.200***	0.200***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
$\Delta \ln(\text{COUNTRY POP}_t)$	0.951***	0.948***	0.946***	0.945***	0.944***
	(0.118)	(0.119)	(0.119)	(0.119)	(0.118)
YearXContinent dummies	Yes	Yes	Yes	Yes	Yes
Mean of dep. variable	0.11	0.11	0.11	0.11	0.11
Number of cities	137	137	137	137	137
Number of subway cities	137	137	137	137	137
Number of periods	12	12	12	12	12
Observations	1428	1428	1428	1428	1428

Notes: Dependent variable: Change in log population in a 5 year period. Sample is cities with subway in 2010. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.

Table 9: Robustness to confounding unobservables

Time periods: Quinquennial panel Estimation: OLS Dependent variable: $\Delta \ln(\text{pop}_t)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$\Delta \ln(s_t)$	-0.002	-0.001	-0.004	0.002	0.003	-0.000	-0.003	-0.029	0.044*	-0.000	-0.000	0.005	-0.002
	(0.003)	(0.006)	(0.005)	(0.007)	(0.004)	(0.004)	(0.006)	(0.032)	(0.027)	(0.005)	(0.004)	(0.004)	(0.006)
(25 km slope > median)$X\Delta \ln(s_t)$		-0.000											
		(0.006)											
(25 km elevation range > median)$X\Delta \ln(s_t)$			0.005										
			(0.006)										
(25km highways > median)$X\Delta \ln(s_t)$				-0.004									
				(0.008)									
(TomTom congestion > median)$X\Delta \ln(s_t)$					-0.012								
					(0.008)								
(Capital)$X\Delta \ln(s_t)$						-0.004							
						(0.006)							
(Good doing business)$X\Delta \ln(s_t)$							0.003						
							(0.009)						
(System built post WW II)$X\Delta \ln(s_t)$								0.028					
								(0.031)					
Centralization$X\Delta \ln(s_t)$									-0.045				
									(0.028)				
(Subway coverage > median)$X\Delta \ln(s_t)$										-0.002			
										(0.004)			
(City pop. 1950 > median)$X\Delta \ln(s_t)$											-0.003		
											(0.007)		
(Coastal city)$X\Delta \ln \ln(s_t)$												-0.015**	
												(0.006)	
(Bus ridership pc > median)$X\Delta \ln(s_t)$													0.002
													(0.007)
$\Delta \ln(\text{GDP}pc_t)$	0.201***	0.201***	0.201***	0.201***	0.299***	0.201***	0.195**	0.202***	0.189**	0.201***	0.201***	0.203***	0.082
	(0.042)	(0.042)	(0.042)	(0.042)	(0.080)	(0.042)	(0.066)	(0.043)	(0.058)	(0.042)	(0.042)	(0.042)	(0.061)
$\Delta \ln(\text{COUNTRY POP}_t)$	0.951***	0.951***	0.955***	0.952***	0.807***	0.952***	0.749***	0.947***	0.776***	0.947***	0.949***	0.956***	0.840**
	(0.118)	(0.117)	(0.115)	(0.117)	(0.111)	(0.119)	(0.092)	(0.117)	(0.061)	(0.117)	(0.117)	(0.109)	(0.366)
YearXContinent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. variable	0.11	0.11	0.11	0.11	0.10	0.11	0.15	0.11	0.07	0.11	0.11	0.11	0.10
Number of cities	137	136	136	137	84	137	63	137	137	137	137	137	40
Number of subway cities	137	136	136	137	84	137	63	137	137	137	137	137	40
Number of periods	12	12	12	12	12	12	12	12	4	12	12	12	12
Observations	1428	1416	1416	1428	937	1428	579	1428	541	1428	1428	1428	453

Notes: s_t is short for subway stations$_t$. Sample is subway cities in 2010. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.

cities respond differently to subways. Column 5 does not support this intuition.

In column 6 we include an interaction of subways with a capital city indicator. If we think, for example, that capital cities are more likely to be the beneficiary of public expenditure than other cities, then we might expect such spending to have a lower return in capital cities than elsewhere. Column 6 does not support this intuition. In column 7 we interact an indicator of an index of institutional quality with subways. If we think that a city's response to subways depends on its ability to reorganize private sector employment, then we might expect cities with a low score on this index to respond differently to subways than those with a high score. The data also do not support this idea.

In column 8 we interact subways with an indicator for whether the subway system predates the second world war — the time when cars became ubiquitous. If we think that older cities are laid out in a way that is more conducive to public transit, then we might expect to see such older cities respond differently to subways than other cities. We do not. In column 9 we interact subways with a measure of city centralization defined as the absolute value of the city light gradient in 1995. The point estimate on main effect is positive and marginally significant at the 10% level and suggests that subways have slightly larger effects on population in more decentralized cities — since the interaction coefficient is negative and of about same magnitude.

Column 10 investigates whether the subway network extent is important. To accomplish this, we calculate the share of all light within 25km of the center that is within 2km of a station. If cities respond differently to subways that serve a larger fraction of their economic activity and population, then we should expect to see a significant coefficient on the interaction of this variable with subways. Our data do not support this intuition. Column 11 investigates whether cities that were large in 1950 respond differently to subways. They do not. In column 12 we see that coastal cities grow slightly less fast in response to subways than do other cities, but this effect is tiny.

Finally, in column 13 we ask whether cities with an effective bus network respond differently to subways than those that do not. The data suggest that they do not. This is consistent with the first difference regression in column 9 of table 6, where we see that controlling for bus ridership in a first difference regression does not lead to a positive estimated effect of subways.

We have now presented five types of results, cross-sectional, first difference, IV, second difference and first differences including a variety of interaction effects. Consistent with descriptive evidence presented in section 1, cross-sectional estimates are much larger than first differences estimates. Results based on metropolitan area light intensity are qualitatively similar to those based on population. Once we add continent specific year effects in column 3 of table 5 the cross-sectional estimate of the effect of doubling subway stations is a 26% increase in population. In first differences, the corresponding estimate is less than 1% and is indistinguishable from zero. Our attempts to deal with confounding dynamics and with omitted variables do not change this conclusion.

Broadly, formal econometric results support the conclusion suggested by the descriptive evidence. That is, that big cities build subways and that these subways subsequently have little or no effect on the population in these cities. Our most favorable IV regressions indicate that doubling a subway system will increase population by less than 2%, although these estimates are never distinguishable from zero and first difference estimates of the effect of subways on population are often an order of magnitude smaller.

8. Subways and urban form

In this section, we use the lights data to investigate the relationship between urban centralization and subway extent. We are interested in determining if the light gradient changes with subway expansions. We follow our previous empirical approach using the light gradient and light intercept in a city-year as our dependent variables. These variables describe, respectively, the rate at which light decays with distance from the center, and brightness at the center.

More specifically, we regress our estimate of the light slope B and the intercept A in equation (3) respectively for each city-year on a measure of subways using the various regression specifications employed previously to analyze subways and population.

Table 10 reports our results. Panel (a) shows results using the light gradient B as the dependent variable, while panel (b) shows results using A as the dependent variable. We first discuss panel (a) at the top of the table. Column 1 shows the pooled OLS estimate. In the cross section, the elasticity of light gradient to subway extent is 0.034. Given that the light gradient is negative, this indicates that cities with larger subway systems have a flatter light gradient and are less centralized. Column 2 presents the first difference regression result in which we find an elasticity estimate of 0.023. In column 3 we control for the second lag of population growth, and find virtually the same coefficient as in column 2. Columns 4 and 5 present our instrumented first difference estimates and show that we find a statistically significant elasticity of 0.060. We experimented with a number of different indexes of centralization, for example, the ratio of light within 5km of the center to light between 5 and 25km. Our estimates of the effects of subways on decentralization are broadly similar across indexes.

The bottom part of the table, panel b, shows results using the light intercept A as the dependent variable. Mean light at the origin is 12.1 log points and is 0.17 log points lower in cities with subways (column 1). Column 2 shows the first difference regression result in which the elasticity of light at the origin to subways is -0.20. Controlling for lagged population growth does not change the estimated coefficient (column 3). Columns 4 and 5 show our instrumented estimates which are also negative but larger and statistically significant at the 5% level in column 6 which controls for lagged population growth. Taking together the results in panels (a) and (b) suggests that subways decentralize activity (flatter light slopes and lower intercept) from the center to the peripheral areas of the city, and are consistent with the absence of population growth documented in section 7.

These results allow us to reject the claim that subways lead to a concentration of activity in the downtown core. While this may seem surprising, decentralization in response to a decrease in transportation costs is an almost universal feature of theoretical descriptions of cities. It is also consistent with established empirical results about the effects roads (Baum-Snow (2007), Baum-Snow *et al.* (2017) and Garcia-López (2012)) and with Ahlfeldt and Wendland (2011) who find that commuter rail contributes to the decentralization of Berlin. In our data, we observe that 72% of subway cities have subway stations beyond 10 kilometers from the city center, and 16% of them have stations beyond 25 kilometers. These statistics suggest that subways are built to have some radial capacity that can contribute to decentralization.

One of the most robust findings of the literature using within city variation to study the effects of subways. e.g., Gibbons and Machin (2005) and Billings (2011), is that economic activity becomes relatively concentrated near subways. To confirm that this feature is present in our data, we restricted attention to areas with 2km of a subway station and recalculated light density gradients for each city on the basis of these areas. As expected, density declines much more slowly along subway lines than it does along other rays out from the city center. That is, our lights data confirm the main pattern seen in studies of subways that exploit within city variation.

9. Ridership

Previous literature has provided wide-ranging predictions about travel mode substitution patterns. For example, the Los Angeles subway expansion was opposed by groups representing residents of poor neighborhoods under the argument that funding (and hence the supply) of buses serving these neighborhoods would decrease as a consequence of large operating subsidies to the subway (Grengs, 2002). If this argument holds in general, we should observe that bus ridership decreases when subways expand. On the other hand, some authors have argued that overall public transit ridership should be positively affected by subway expansions since buses and subways complement each other in providing public transportation (*c.f.* Hensher, 2007). As an example of why this would occur they point out that bus lines are redesigned after subway expansions to feed passengers into the subway system. Under this argument bus ridership should increase when subway systems expand. Finally, studies of rail expansions have argued that most subway users were previously bus users (Baum Snow and Kahn, 2005), suggesting that the net effect on overall ridership of rail expansions should be small.

Table 11 shows pooled cross sectional estimates relating subway extent to ridership. Cities with larger subway systems have more transit riders (the elasticity is 0.90 in column 2). Similarly, cities with larger subway systems have more subway riders (the elasticity in column 4 is 1.19) as well as bus riders (the elasticity in column 6 is 0.61). As with Table 5, we view these pooled OLS estimates as mainly descriptive.

Table 12 presents our first difference estimations. In Column 3 we find that the total transit ridership elasticity of subway extent is 0.68 (significant at the 5% level). This suggests that subway expansions lead to increases in total transit ridership.

In columns 4-6 we show that subway ridership elasticity to subway extent is 0.61 and is distinguishable from zero. On the other hand, the effect of subway expansions on bus ridership is close to zero in columns 7-9. This echoes Duranton and Turner (2011) who find that increase to the stock of highway kilometers in a city lead to large increases in driving, and that only a little of this increase reflects diversion of traffic from other roads. The results in columns 4-6 also suggest that subway ridership increases less than proportionally with system extent (e.g., one-sided test p-value=0.044 for column 6). This is interesting for two reasons. First, it suggests that increases in subway extent elicit smaller increases in ridership than the increases in driving that follow from increases in the road network (Duranton and Turner, 2011). Second, it suggests that subway networks may be subject to decreasing returns to scale. This is consistent with findings of decreasing returns to scale in the road network in Couture, Duranton, and Turner (2018).

10. Discussion

A *Subways and growth*

On the basis of figure 3, it is natural to conjecture that subways are important for the growth of cities. Our cross-sectional estimates support this conjecture. With 4.5 lines in an average system, adding a subway line is about a 23% increase in system extent. Using our cross-sectional estimate of the relationship between subway lines and population we have that a new subway is associated with a population increase of about 12%.[29] This is close to a back of the envelope calculation of the population growth that would occur if a new subway line operated at capacity and all of its riders migrated to the city because of the new subway line.[30] Thus, if we compare the cross-sectional estimates with the technical capabilities of subways, the cross-sectional estimates seem feasible, but only barely.

Other estimation strategies tell a different story. Our first difference estimates suggest that doubling the extent of a subway network causes at most a tiny increase in population. While these estimates are consistent with patterns seen in the raw data, the possibility of confounding dynamics or omitted variables are obstacles to a causal interpretation of these estimates. To investigate the

[29]From table 5 column 6, the subway line elasticity of population is 0.52. Thus we have, $0.52 \times 0.23 = 0.12$.

[30]Ten car subway trains can carry about 35,000 people per hour (Transit Capacity and Quality of Service Manual (1999)(ch. 1, part 1, p1-22), Transit Cooperative Research Program) or 87,500 over the course of a 2.5 hour morning commute. Thus, a single new subway line could allow 87,500 new commuters to reach a central city. With a 50% labor force participation rate such a migration could increase a city's population by 175,000. This is 3.7% increase to the 4.7m population of an average subway city in our sample. Since an average subway system has 57 stations and an average subway line has 13.2 stations, adding a new subway line is a 23% increase in the extent of an average subway network. Dividing, this suggests that doubling the extent of an average subway network could lead to a population increase of about $\frac{0.037}{0.23} \times 100 = 16.1\%$.

possibility that subway expansions systematically occur in periods of low population growth, we control for the recent history of population growth, conduct second difference and instrumental variables estimates. These estimates also yield tiny elasticities. To investigate the role of omitted variables we consider a large set of possible control variables. These estimates fail to find evidence for a big hidden effect of subways on growth. The weight of evidence hence suggests that big cities build subways, but that subways have at most a tiny effect on urban population growth.

B *Subways and ridership*

We also investigate the effect of subway expansions on transit ridership. Somewhat surprisingly, we find that subway expansions do not decrease bus ridership. We also find that doubling the extent of a subway network leads to about a 60% increase in ridership. Our estimates are precise enough to allow us to reject the hypothesis of no-effect and also to reject the hypothesis of a 100% effect. Thus, our point estimates are suggestive of a large ridership response to subway expansions, and also of modest decreasing returns to subway extent.

To understand the relationship between our findings for ridership and population, we first calculate the number of immigrant subway commuters that would be required to completely account for the increase in ridership associated with a subway expansion.[31] This calculation suggests that the increases in ridership that follow subway expansions are far too large to consist of immigrant commuters. This suggests, in turn, that increases in ridership must primarily reflect an increase in commute or non-commute trips by current residents.

C *Subways and decentralization*

Our investigation of the effect of subways on urban form finds that subway expansions cause cities to spread out. Our first difference and IV estimations in table 10 indicate that a doubling of the subway network causes the light density gradient to flatten by between 0.02 and 0.06. Using the larger of these two estimates, we can calculate that a doubling of the subway network causes the share of all light within 5km of the center to decrease by about 2.2% in an average city, holding total light constant. At 13.2 stations per line and 57 stations per system, adding an average radial subway line increases system capacity by about 23% and should lead to about 0.5% decrease in the central share of a city's light.

Although this decentralization effect is also seen for radial highways, the effect of subways seems to be smaller. Baum-Snow (2007) finds that a single interstate highway causes about 9% of the population of a US city to decentralize, while Baum-Snow *et al.* (2017) find that a radial highway

[31] An average subway network serves about 377m riders per year. If a dedicated subway commuter rides the subway twice per day, 250 days per year, then an average subway system could serve about 0.75m such commuters. This means doubling the extent of a subway network would require about $0.6 \times 0.75\text{m} = 0.45\text{m}$ new dedicated subway commuters. With 50% labor force participation and average city population of 4.7m, if, hypothetically, new ridership resulting from an expansion is provided by new migrants to the city who are dedicated subway commuters, then city population would increase by about 19% in response to a doubling of system extent.

causes about 5% of the population of a Chinese city to decentralize. These effects are about 10 times as large as those we find for subways. The relative size of the subway effect seems even smaller if we compare the capacity of a subway line with that of a radial highway.[32]

To understand the relationship between our decentralization results and those for population and ridership, suppose that changes in light are exactly proportional to changes in residential population. In this case, doubling the extent of an average city's subway network would lead to a 2.2% decrease in lights within 5km of the center. If changes in lights and changes in population are perfectly proportional, this requires that about 94,000 people move in an average city with population 4.7m. Again assuming 50% labor force participation, this means moving 47,000 workers. If all of these workers use the subway to commute to immobile jobs from their newly remote residences, this subway induced decentralization will give rise to about 47,000 new dedicated subway commuters.

We saw above that the increase in ridership that follows from a doubling of the subway network could serve about 450,000 new dedicated subway commuters. Even under our extreme assumptions, this is about 10 times as many as are implied by the amount of decentralization. Thus, we probably cannot account for the increase in ridership that follows a subway expansion with an increase in commuting by newly decentralized existing residents.

Since we also cannot account for the increase in ridership with new city residents, we conjecture that subways either displace other modes of transport while keeping travel constant, or induce new trips by subway. The evidence provided in Gendron-Carrier et al (2017) that subways lead to lower levels of pollution is consistent with this kind of transport mode substitution.

11. Conclusion

Subway expansions appear to have little or no effect on population growth, they lead to modest increases in ridership, and they have small effects on the configuration of cities. New ridership is unlikely to primarily consist of new commuters and subway expansions probably lead to increases in aggregate city land rent that are small relative to construction costs. These results do not seem to provide a basis for justifying the large subsidies that subway construction and operation often requires.

While we have addressed the effects of subway expansion on population, urban form and ridership, we have not addressed the effect of subway expansions on air pollution - although this is taken up in a companion piece (Gendron-Carrier et al. 2017). With this said, our results so far suggest that the evaluation of subway projects ought to rest on the demand for mobility, farebox revenue, and not on the ability of subways to promote city growth.

[32]As we note in footnote 23, a subway line can carry about 35,000 people per hour at peak capacity. A limited access highway lane carries about 2,200 cars per hour at peak capacity. Thus, a four lane radial highway consisting of two lanes in each direction can carry about 4,400 cars per hour each way, about 12% of the capacity of a subway line (in the US, interstate highways are most often two lanes in each direction).

Table 10: Decentralization - Radiance calibrated light gradient

Panel a - Light gradient

Dependent variable:	Light gradient	ΔLight Gradient			
Estimation:	OLS	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)
$\Delta\ln(\text{subway stations}_t)$		0.023***	0.024***	0.047*	0.060**
		(0.0062)	(0.0062)	(0.025)	(0.024)
$\ln(\text{subway stations}_t)$	0.034***				
	(0.010)				
$\Delta\ln(\text{GDP}pc_t)$		-0.078	-0.079	-0.100*	-0.11*
		(0.053)	(0.053)	(0.056)	(0.058)
$\Delta\ln(\text{COUNTRY POP}_t)$		-0.0051	-0.0014	-0.091	-0.13
		(0.17)	(0.17)	(0.21)	(0.22)
$\ln(\text{GDP}pc_t)$	0.043*				
	(0.024)				
$\ln(\text{COUNTRY POP}_t)$	0.048***				
	(0.014)				
$\ln(\text{pop}_{t-2})$ control			Yes		Yes
Mean of dep. variable	-0.811	0.041	0.041	0.041	0.041
Mean of subways regressor	3.06	0.36	0.36	0.36	0.36
SD subways regressor	1.49	0.82	0.82	0.82	0.82
R-squared	0.35	0.19	0.19	0.17	0.15

Panel b - Light intercept

Dependent variable:	Light intercept	ΔLight intercept			
Estimation:	OLS	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)
$\Delta\ln(\text{subway stations}_t)$		-0.20***	-0.20***	-0.33	-0.46**
		(0.056)	(0.056)	(0.20)	(0.20)
$\ln(\text{subway stations}_t)$	-0.17*				
	(0.085)				
$\Delta\ln(\text{GDP}pc_t)$		1.36**	1.37**	1.48**	1.61**
		(0.47)	(0.48)	(0.50)	(0.52)
$\Delta\ln(\text{COUNTRY POP}_t)$		0.14	0.10	0.62	1.03
		(1.20)	(1.21)	(1.58)	(1.59)
$\ln(\text{GDP}pc_t)$	0.030				
	(0.22)				
$\ln(\text{COUNTRY POP}_t)$	-0.23*				
	(0.13)				
$\ln(\text{pop}_{t-2})$ control			Yes		Yes
Mean of dep. variable	12.135	-0.367	-0.367	-0.367	-0.367
Mean of subways regressor	3.06	0.36	0.36	0.36	0.36
SD subways regressor	1.49	0.82	0.82	0.82	0.82
R-squared	0.27	0.30	0.30	0.30	0.28
Number of cities	137	137	137	137	137
Number of subway cities	137	137	137	137	137
Number of periods	4	3	3	3	3
Observations	548	411	411	411	411

Notes: For each city-year, a linear regression was estimated between the log mean radiance calibrated light intensity in successive rings at 0-1.5km, 1.5-5km, 5-10km, 10-25km and 25-50km and log distance from the city center centroid. Panel a column 1 dependent variable is the slope of the light gradient. Columns 2-5 use as dependent variable the change in slope over a 5 year period. Panel b column 1 dependent variable is the intercept of the light gradient. City-level robust standard errors in parentheses. All regressions include geographic controls and year by continent dummies. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.

Table 11: Log ridership - Pooled cross section

Time periods: Quinquennial panel						
Dependent variable:	ln(All ridership$_t$)		ln(Subway ridership$_t$)		ln(Bus ridership$_t$)	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(subway stations$_t$)	0.66***	0.90***	1.09***	1.19***	0.54***	0.61***
	(0.16)	(0.15)	(0.13)	(0.15)	(0.14)	(0.11)
ln(GDPpc_t)		-1.31***		-0.25		-1.76***
		(0.31)		(0.28)		(0.37)
ln(COUNTRY POP$_t$)		-0.12		-0.09		-0.04
		(0.17)		(0.15)		(0.15)
Geographic controls	No	Yes	No	Yes	No	Yes
YearXContinent dummies	No	Yes	No	Yes	No	Yes
Mean of dep. variable	19.77	19.77	18.82	18.82	18.60	18.60
Mean of subways regressor	4.04	4.04	3.87	3.87	3.67	3.67
SD subways regressor	0.98	0.98	1.04	1.04	1.17	1.17
R-squared	0.32	0.78	0.57	0.74	0.23	0.70
Number of cities	34	34	78	78	45	45
Number of subway cities	34	34	78	78	45	45
Number of periods	10	10	10	10	10	10
Observations	88	88	225	225	117	117

Notes: Dependent variable: Log ridership of subways and buses in metropolitan area in period t. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. Geographic controls are capital city dummy, log km to ocean, log km to land border, and log km to major navigable river. (Odd columns)-Pooled cross section. (Even columns)-Add geographic controls, GDP pc control, country population, and yearXcontinent dummies.

Table 12: Log ridership - First differences

Time periods: Quinquennial panel									
Dependent variable:	$\Delta \ln(\text{All ridership}_t)$			$\Delta \ln(\text{Subway ridership}_t)$			$\Delta \ln(\text{Bus ridership}_t)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta \ln(\text{subway stations}_t)$	0.728**	0.734**	0.678**	0.572**	0.660**	0.613**	-0.001	0.005	-0.011
	(0.238)	(0.261)	(0.299)	(0.213)	(0.198)	(0.224)	(0.044)	(0.060)	(0.050)
$\Delta \ln(\text{GDP}pc_t)$			0.069			0.158			0.271
			(0.229)			(0.228)			(0.276)
$\Delta \ln(\text{COUNTRY POP}_t)$			1.238			1.116			3.181**
			(1.302)			(1.154)			(1.186)
YearXContinent dummies	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Continent dummies	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Mean of dep. variable	0.064	0.064	0.064	0.150	0.150	0.150	0.014	0.014	0.014
Mean of subways regressor	0.06	0.06	0.06	0.11	0.11	0.11	0.10	0.10	0.10
SD subways regressor	0.15	0.15	0.15	0.23	0.23	0.23	0.38	0.38	0.38
R-squared	0.39	0.56	0.57	0.20	0.41	0.42	0.00	0.35	0.39
Number of cities	24	24	24	63	63	63	31	31	31
Number of subway cities	24	24	24	63	63	63	31	31	31
Number of periods	8	8	8	9	9	9	8	8	8
Observations	48	48	48	143	143	143	63	63	63

Notes: Dependent variable: Change in log ridersip of metropolitan area in a 5 year period. Sample is subway cities. City-level clustered standard errors in parentheses. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. (1)- No controls. (2)-Add yearXcontinent dummies (3)-Add log gdp and log country pop. controls.

References

Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm, and Nikolaus Wolf. 2015. The economics of density: Evidence from the Berlin wall. *Econometrica* 83(6): 2127–2189.

Ahlfeldt, Gabriel and Nicolai Wendland. 2011. Fifty years of urban accessibility: The impact of the urban railway network on the land gradient in Berlin 1890-1936. *Regional Science and Urban Economics* 41: 77–88.

Arellano, Manuel and Stephen Bond. 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment. *Review of Economic Studies* 58(2): 277–297.

Barnes, Gary. 2005. The importance of trip destination in determining transit share. *Journal of Public Transportation* 8(2): 1–15.

Baum-Snow, Nathaniel. 2007. Did highways cause suburbanization? *The Quarterly Journal of Economics* 122(2): 775–805.

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, railroads, and decentralization of chinese cities. *Review of Economics and Statistics* 99(3): 435–448.

Baum-Snow, Nathaniel and Matthew E. Kahn. 2005. Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000. *Brookings-Wharton Papers on Urban Affaires: 2005* 1(4): 147–197.

Billings, Stephen B. 2011. Estimating the value of a new transit option. *Regional Science and Urban Economics* 41(6): 525–536.

Blonigen, Bruce and Anca D. Cristea. 2015. Air service and urban growth: Evidence from a quasi-natural policy experiment. *Journal of Urban Economocs* 86: 126–146.

Campante, Filipe and David Yanagizawa-Drott. 2018. Long-range growth: Economic development in the global network of air links. *The Quarterly Journal of Economics* 133(3): 1395–1458.

Clark, Colin. 1951. Urban population densities. *Journal of the Royal Statistcal Society* 114(4): 490–496.

Couture, Victor, Gilles Duranton, and Matthew A. Turner. 2018. Speed. *Review of Economics and Statistics* Forthcoming.

Duranton, Gilles and Matthew A. Turner. 2011. The fundamental law of road congestion: Evidence from US cities. *American Economic Review* 101(6): 2616–2652.

Duranton, Gilles and Matthew A. Turner. 2012. Urban growth and transportation. *Review of Economic Studies* 79(4): 1407–1440.

Garcia-López, Miquel-Ángel. 2012. Urban spatial structure, suburbanization and transportation in Barcelona. *Journal of Urban Economics* 72: 176–190.

Garcia-López, Miquel-Ángel, Adelheid Holl, and Elisabet Viladecans-Marsal. 2015. Suburbanization and highways: When the Romans, the Bourbons and the first cars still shape Spanish cities. *Journal of Urban Economics* 85: 52–67.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew Turner. 2017. Subways and urban air pollution. *NBER Working Paper* (24183).

Gibbons, Stephen and Stephen Machin. 2005. Valuing rail access using transport innovations. *Journal of Urban Economics* 57(1): 148–1698.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport. 2008. Why do the poor live in cities? the role of public transportation. *Journal of Urban Economics* 63(1): 1 – 24.

Gomez-Ibanez, Jose A. 1996. Big-city transit, ridership, deficits, and politics. *Journal of the American Planning Association* 62(1): 30–50.

Gordon, Peter and Richard Willson. 1984. The determinants of light-rail transit demand - an international cross-sectional comparaison. *Transportation Research Part A: General* 18(2): 135–140.

Grengs, Joseph. 2002. Community-based planning as a source of political change: The transit equity movement of Los Angeles' bus riders union. *Journal of the American Planning Association* 68(2): 165–178.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. Measuring economic growth from outer space. *American Economic Review* 102(2): 994–1028.

Hensher, David. 2007. *Bus transport, economics, policy and planning*. JAI Press.

Hsu, Wen-Tai and Hongliang Zhang. 2014. The fundamental law of highway congestion: Evidence from national expressways in Japan. *Journal of Urban Economics* 81: 65–76.

Manelici, Isabela. 2017. Terrorism and the value of proximity to public transportation: Evidence from the 2005 london bombings. *Journal of Urban Economics* 102: 52–75.

Mills, E. S. and J. Peng. 1980. A comparison of urban population density functions in developed and developing countries. *Urban Studies* 62(3): 313–321.

Olley, G. Steven and Ariel Pakes. 1991. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64(6): 1263–1297.

Pang, Jindong. 2017. Do subways improve labor market outcomes for low-skilled workers? *working paper, Syracuse University* .

Redding, Stephen J. and Matthew A. Turner. 2015. Transportation costs and the spatial organization of economic activity. In Gilles Duranton, William Strange, and J. Vernon Henderson (eds.) *Handbook of Urban and Regional Economics Volume 5*. New York: Elsevier, 1339–98.

Small, Kenneth A. and Erik T. Verhoef. 2007. *The economics of urban transportation*. New York (NY): Routledge.

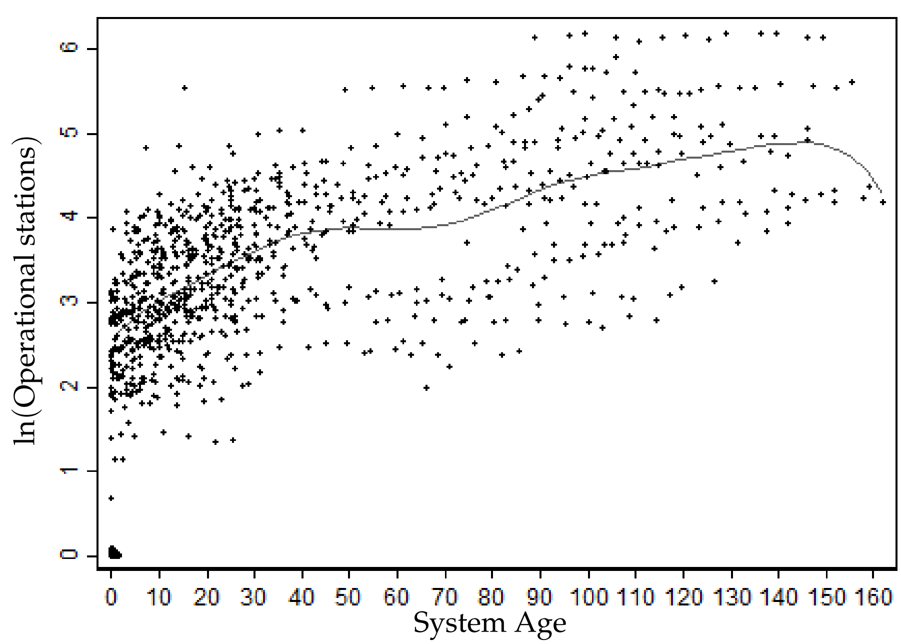
Storeygard, Adam. 2017. Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa. *Review of Economic Studies* .

Wooldridge, Jeffrey M. 2001. *Econometric Analysis of Cross Section and Panel Data*. First edition. Cambridge MA: MIT press.

Appendix: Supplemental results

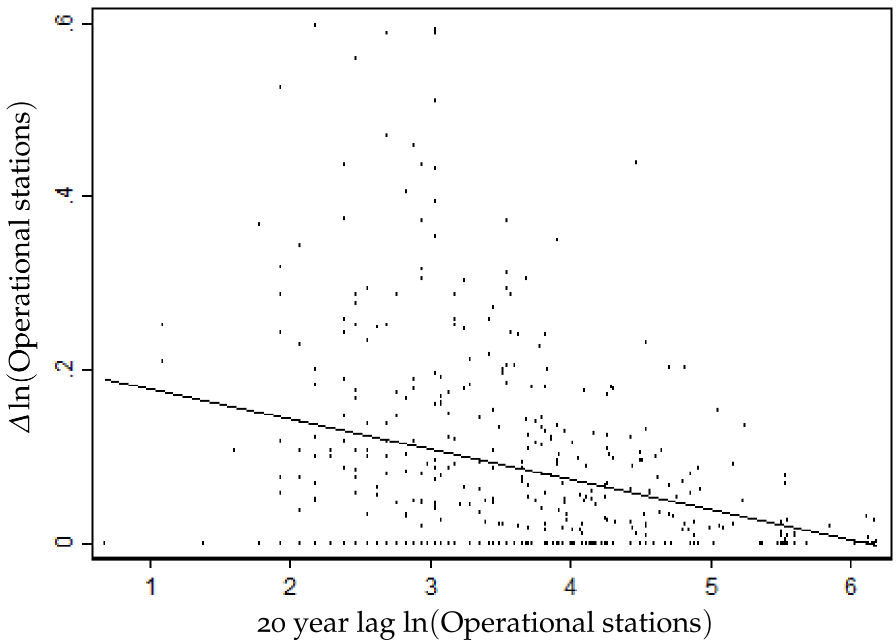
While figure 1 shows the growth of the world's subways, figure A.1 traces out the extent of individual systems as a function of the time since they opened. Each marker in this figure describes a city year, so that there is one marker for each of the city-years in our data where at least one subway station is open. Consistent with figure 1, most of the observations are in the left portion of the graph. This reflects the fact that many subways systems have opened in the past 30 years. On the other hand, markers in the right hand portion of the graph describe the handful of subway systems that date back to the 19th century. The solid line in the figure describes a locally weighted regression of system extent on system age. This figure suggests that the expansion of a city's subway network is predictable. Expansion is rapid during the first 30-40 years after a system opens and slows thereafter. Figure A.2 illustrates the variation that identifies our first stage regression more explicitly. The horizontal axis is the fourth lag of log system extent and the vertical axis is change in current log extent. The negative relationship we would expect from figure A.1 is clear. Table A.1 presents our first stage regressions. These regressions show that the clear negative relationship between lagged level and change that we see in figure A.2 is robust to the inclusion of controls.

Figure A.1: Stations in a subway system by time since system opening



Note: Vertical axis is log of subway stations in a system. Horizontal axis is years since system opening. Dots indicate individual city-years.

Figure A.2: Growth of subways and 20 year lagged subway level



Note: Vertical axis is change in log stations in a system over five years. Horizontal axis is log stations 20 years prior ($t-4$). Linear fit overlaid.

Table A.1: Subways first stage: First difference – lagged subway instruments

	(1)	(2)	(3)
Dependent variable:	$\Delta \ln (s_t)$	$\Delta \ln (s_t)$	$\Delta \ln (s_t)$
$\ln(\text{subway stations}_{t-4})$	-0.094***	-0.100***	
	(0.008)	(0.008)	
$\ln(\text{subway stations}_{t-8})$			-0.067***
			(0.005)
$\Delta \ln(\text{pop}_{t-2})$	0.084	-0.121	0.199
	(0.151)	(0.526)	(0.151)
$\Delta \ln(\text{pop}_{t-3})$		0.251	
		(0.585)	
$\Delta \ln(\text{GDP}pc_t)$	0.024	0.001	0.057
	(0.160)	(0.170)	(0.167)
$\Delta \ln(\text{COUNTRY POP}_t)$	0.905	0.980	1.156*
	(0.660)	(0.662)	(0.613)
YearXContinent dummies	Yes	Yes	Yes
Mean of dep. variable	0.29	0.31	0.29
R-squared	0.13	0.12	0.10
Number of cities	137	137	137
Number of subway cities	137	137	137
Number of periods	10	9	10
Excluded instruments F-stat	132.36	147.51	153.49
Observations	1235	1124	1235

Notes: Dependent variable is the change in log subway stations in a five year period. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01. Sample is subway cities. City-level clustered standard errors in parentheses.

Appendix: Data description

In this subsection we describe the data sources and variable definitions for each of the interaction variables used in Table 9.

Digital elevation maps (DEM) were obtained from the publicly available Shuttle Radar Tomography Mission (NASA-SRTM). The DEM dataset contains elevation as well as land slope at 3 arc-second resolution (about 90 meters) worldwide. The mean slope was calculated within a 25 km disk around the city center. Cities were then partitioned at the median value of the average slope to generate the interaction used in column 2.

The elevation range variable was defined using the SRTM DEM data as the maximum minus the minimum value for terrain elevation within a 25km disk around the city center. Cities were then partitioned at the median value of the elevation range to generate the interaction used in column 3.

Digital data on worldwide highways was obtained from ESRI's roads and highways layer. We used rank 1 roads (highways) and calculated total kilometers of roads within a 25km disk of a city's center. Cities were then partitioned at the median value of kilometers of highways in a city to generate the interaction used in column 4.

Congestion data was downloaded from TomTom (`http://www.tomtom.com/en_ca/trafficindex/#/list`, accessed July 2015) which ranks city traffic conditions in 219 major cities worldwide. Cities were partitioned at the median value of congestion to generate the interaction used in column 5.

Capital city refers to being a country capital. This variable was obtained from the UN cities dataset.

For institutional quality we used the World Bank's Doing Business ranking (`http://www.doingbusiness.org/rankings`, accessed may 2013). The 'Good for doing business' variable used in column 7 indicates that the country is among the top half for ease of doing business, which means the regulatory environment is conducive to the starting and operation of a local firm. The rankings are determined by sorting the aggregate distance to frontier scores on 10 topics, each consisting of several indicators, giving equal weight to each topic.

The centralization variable used in column 9 is defined the absolute value of the city light gradient in 1995. Larger values hence correspond more centralized cities.

The subway coverage variable used in column 10 is a measure of whether the subway system in 1995 provided an above median coverage of total city lights. To create this variable, we first defined 2km radius disks around subway stations operational in 1995. We then calculated the sum of lights within the subway disks in 1995 and proceeded to take the ratio of this value to the sum of lights in a 25km disk around the city center. Cities were partitioned at the median value of subway coverage to generate the interaction used in column 10.

In column 11, a city was classified as coastal if its centroid is located within 20km of the ocean. To provide a concrete example of this, Houston is the city closest to the limit of the cutoff for being coastal using this definition.