2nd Draft: 5/16/2020

Doubts about Density: Covid-19 across Cities and Towns

By

William C. Wheaton
Department of Economics
Center for Real Estate
MIT*
Cambridge, Mass 02139
wheaton@mit.edu

Anne Kinsella Thompson
Center for Real Estate
MIT*
Cambridge, Mass 02139
anniek@mit.edu

The authors are respectively Professor (Emeritus, Economics) and Lecturer at the MIT Center for Real Estate. They remain responsible for all results and conclusions derived there from.

JEL Code: R, I.
Keywords: Coronavirus, Health, Urban Economics
http://ssrn.com/abstract=3586081

Abstract

It has been more than 100 days since the first case of Covid-19 was detected in the US. Until most recently information on the spread of the disease was not available by geographic areas smaller than counties. In an earlier piece we found a significant positive impact of county density on the daily progress of the disease in a dynamic model. Since then, sporadic data on disease incidence at the municipal level has become available in a few states. We use Massachusetts data released on 4/15/20 and then again on 5/6/20 since it is a fully incorporated state. We find that municipalities with greater density and with a greater share of land use in commercial-industrial categories have a significantly higher percapita incidence of the disease. The quantitative impact of these two "urban" variables is particularly large.

I. Introduction.

Historically, dense urban living – most often found in large cities – has had a mixed reputation. Jacobs (1961) reviewed the older literature on health and the squalor of cities and combined it with the newer negative experiences of Post War urban renewal. She suggested that lower density living, albeit with planned mixed use, offered a better human environment. In the more recent literature from Urban Economics, there has been a resurgent interest in urban "increasing returns" (Krugman, 1991). Krugman argued that increased productivity and consumption diversity result when economic activity is agglomerated. The empirical work that followed this tended to support higher urban wages, greater urban productivity and the superior operation of urban labor markets. It is quite extensive and we simply cite the paper by Ciccone and Hall as an example (1996). Most recently, Glaeser (2010) has argued that the economic advantages of dense urban living outweigh any disadvantages – so large, dense cities are in fact civilization's greatest "Triumph".

With the onslaught of the Covid-19 virus, however, the media and medical community have begun to suggest that intrinsically there might be higher infection rates (percapita) in urban as opposed to less dense or rural areas. Living farther apart is certainly a helpful first step in creating greater "social distancing", and the latter appears right now to be the only salve for the virus. Until and if effective vaccines become available – social isolation is the best way to slow or halt the spread of the disease.

In this paper we use a newly released data base of recorded virus cases per capita (as of 5/6/20) across the 351 cities and towns that fully cover the Commonwealth of Massachusetts. There is plenty of accompanying static data about the differences between these municipalities, but we focus on three hypothesized factors that may impact the growth of the disease. The first is percapita income. The second and third relate to how "urban" a community is. In addition to overall population density, our third variable measures the prevalence of commercial and industrial facilities (as a share of land use) in a town. These latter activities are where people tend to work and socialize in close proximity. We find very strong evidence that the disease is more prevalent in dense towns, where there is a large fraction of commercial and industrial land use. The income result is more mixed. In this data, "urban living" seems to have inherent risks that pertain to the current virus, and possibly future other infectious diseases as well.

II. Modeling Covid-19 Infections in a cross section of cities and towns

Our approach in this paper is to provide a very simple cross-section analysis of the percapita infection rate on 4/15/2020 and then again on 5/6/20 for most of the Commonwealth's 351 cities and towns. Unlike our previous paper (SSRN 3570540) we do not have the daily history of cases over time and so cannot relate the disease incidence at a particular time to the days since it first began in a particular town. This omission may not be that serious however if we think that the disease was widely "seeded" during a common window. The disease first appears to have been seeded in numerous places around the Commonwealth - all during the last few weeks of February. The major Bio Tech conference on 2/26/20 quickly created a large number of seeds throughout the Boston area (NYT 4/12/20), but there were also seeds in the western part of the state at that time as well.

Our Model allows for some flexibility in functional form and is shown below in equation (1). If a covariate (X) is thought to be a necessary condition for virus cases (a zero value implies zero cases), then we enter it multiplicatively (or in logs for the estimating equation). Here the estimated value of the λ parameter represents a full elasticity of the variable on the incidence rate. If a covariate (Z) need not be present for the virus, then its exponentiated value can be entered multiplicatively (or linearly in the estimating equation). Here the estimated value of the parameter λ represents a semi-elasticity (the percentage change in the incidence rate given a unit change in the covariate).

To statistically estimate the values of the λ (and λ_0) we take natural logs and transform the equation – as in the second line of (1). If the original source of equation error is multiplicative and normal, then in the transformed equation it will be additive and log normal. The statistical standard error of the equation will likely be proportional to the number of cases percapita, but this is taken care of by using Robust Standard Errors. We make no attempt to investigate the possibility of spatial autocorrelation (which might occur if neighboring towns tend to infect each other).

$$C_i / P_i = \lambda_0 \prod X_i^{\lambda} \exp(\sum \lambda Z_i) \qquad (1)$$

$$\ln(C_i / P_i) = \ln(\lambda_0) + \sum \lambda \ln(X_i) + \sum \lambda Z_i$$

$where:$

$C_i/\mathrm{P_i} =$ Cases per capita in area i at the observation date

$X_i =$ covariate values in area i whose zero value implies zero cases

$Z_i =$ covariate values in area i whose zero value does not imply zero cases

$\alpha, \lambda, \lambda =$ parameters to be estimated

III. Data (4/21/20 with 5/6/20 update)

Prior to 4/15/20 the Commonwealth of Massachusetts released virus case counts only by county – updating these daily. Since April 15 it has released the data additionally by city and town, but with a limit. Any town with 5 or fewer cases would simply display the number 5. The result on 4/15/20 is a sample of 269 cities and towns with positive entries not equal to the truncated number. The data was then updated on 5/6/20 and the sample of non-truncated towns expanded to 290.

We also obtain current (2019) data for each municipality on its population, overall residential density (population/square mile), average per capita income, and finally the share of assessed property value in the commercial and industrial categories – as opposed to residential. Commercially operated apartment buildings are classified as residential. Table 1 gives the municipal level statistics used in the analysis.

Table 1: Mass. Municipal Data

Series	Obs	Mean	Std Dev.	Minimum	Maximum
Cases/100000 pop(4/15)	269	303	211	27	1890
Cases/100000 pop(5/6)	294	687	561	90	5958
Density (pop/sqm)	351	1368	2590	4	19541
Per Capita Income	351	44054	13106	19024	115586
Share Land Use C&I	351	.1157	.0992	.02	.63
Boston MSA	138	.392	0.0	0	1

The virus case rate exhibits wide variation – about 70 fold – on both dates. Density varies even more, so much so that we also identified town's that lie within the greater Boston MSA and will run separate analysis for that subset – to eliminate truly rural areas of the state. In this subsample the minimum density is closer to 160. The share of land use in the commercial-industrial category runs from zero up to a very large 63%. There is a half dozen communities in the Commonwealth where commercial and industrial activity accounts for the majority of the capital stock.

IV. Estimation results

In all regressions, the dependent variable is the log of cases per 100000 population either on 4/15 or 5/6. We experimented with each independent variable, entering it either in logs or linearly, as described in equation (1). The result with the best fit was to enter all three variables in logs. These results are shown in Table 2, both for the full sample of 269 towns (with non-truncated data) as well as the 138 towns that are part of the greater Boston MSA. The impact of density and land use seems quite robust to this choice of sample, but town income is important only when considering the eastern, developed portion of the Commonwealth.

Table 2: MSA Regressions, dependent variable: log (cases/100000 population)

Independent	4/15: All Towns	4/15: Boston	5/6: All Towns	5/6: Boston
R-squared	0.215	0.471	0.440	0.591
Constant	5.901***	8.464***	7.48***	11.118***
	(1.255)	(1.602)	(1.199)	(1.616)
Density (log)	0.191***	0.297***	0.299***	0.312***
	(0.0293)	(0.0374)	(0.0267)	(0.0372)
Income/pop (log)	-.1305	-0.438***	-0.252***	-0.583***
	(0.1214)	(0.1469)	0.117***	(0.1482)
% Land Use (log)	0.112***	.1020**	0.201***	0.2405***
	(0.521)	(0.5603)	(0.047)	(0.0557)
Number of Observations	269	138	294	138

Notes: Standard errors in parentheses. *p 0.10, **p<.05, ***p<.01 (robust standard errors)

VI. Discussion

The “fit” of our model is improving quite significantly over time. Even with additional town data points, the variance in the virus infection rate is better explained on 5/6 than on 4/15. Beyond the statistical significance of the results, it is important to assess the quantitative impact of overall community land use on the prevalence of the Covid-19 disease. Within the subsample of Boston MSA towns, density ranges from roughly 160 to 19000. Using the most recent results for the Boston MSA, this range, when taken to the power of .29, generates a predicted 6-fold increase in cases. In other words, the “slope” of the density coefficient is very “impactful”. Turning to the coefficient on commercial-industrial land use share. The sample range of .02-to-.63 when taken to the power of .245 predicts a 130% increase in the case rate. While this is a significant impact – it is far from the impact of density. The full equation, however, compounds the two effects. So a city with .63 commercial land use share and the maximum density has 14 times the predicted virus cases of a suburban town with the minimum density and only 2% non-residential land use. In the Boston MSA the actual case rate on 5/6 ranges from 260 to 5958 or about 20-fold. This is close to the predicted range.

While significant statistically, percapita income is much less impactful. The full sample range in income (115k down to 19k) increases the virus rate by only 2-fold.

While dense urban living clearly still brings well documented economic benefits, if pandemic virus’ are to become more common, there may be a rethinking of whether these benefits are worth the magnitude of the risks uncovered here.

REFERENCES

Ciccone, Antonio & Hall, Robert E, 1996. "Productivity and the Density of Economic Activity," *American Economic Review*, vol. 86(1), pages 54-70, March.

Glaeser, Edward. (2010) *The Triumph of the City…,* The Penguin Press.

Jacobs, Jane. (1961) *The Life and Death of Great American Cities.* Random House, NY.

Krugman, P. (1991) *Geography and Trade*. MIT Press.

https://www.nytimes.com/2020/04/12/us/coronavirus-biogen-boston-superspreader.html

Wheaton, W., Kinsella, A.K. (2020) “The Geography of Covid-19 Growth in the US: Counties and Metropolitan areas”, MIT Working Paper, SSRN 3570540.