VVoIP Theoretical Geography Debates 13 December 2019

Do We Design Our Understanding of Cities or Do We Discover it?

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty

http://www.complexcity.info/
http://www.spatialcomplexity.info/

Outline

- About Prediction, About Invention
- Induction and Deduction: About Discovery
- Thinking About Cities, About the Scientific Object
- The Question: Theoretical Geography: Discovered or Invented
- Three Examples
  - Cities as Machines, Mechanisms that are Invented not Discovered
  - Search for Urban Patterns: Patterns of Human Behaviour
  - Generalisation: How Can we Demonstrate This Robustness
- Developing Applications- Theory into Design: Present into Future

About Prediction, About Invention

- My key argument: for most systems across all the sciences, it is only possible to predict within quite strict limits, and an extreme version of this is that we cannot predict at all
- We have enough evidence to now show that most of our predictions are flawed, badly so if we go outside certain limits
- There is a host of evidence just in the economy to reveal our very limited abilities in this regard.
- To an extent this is obvious, we are not magicians, we cannot predict the future
- The future is unknowable largely because it is invented by ourselves.
- We invent the future, as individuals and sometimes as collectives but this does not mean we invent a better future

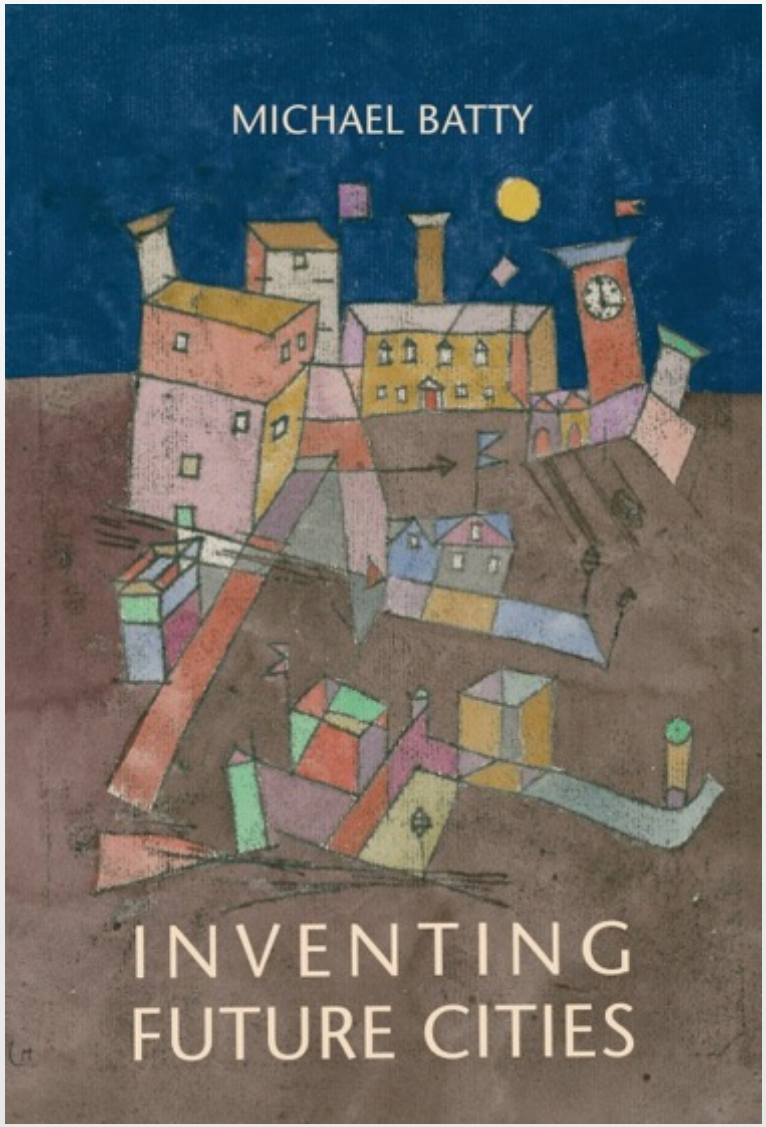
- But if we can't predict the future, we can't predict what we might invent. ***This is the double bind of prediction***.
- Let me take time out to state the logic of our being unable to predict the future. This is Karl Popper's argument. Popper argues that only in entirely closed systems where you have no influences in or out of the system – can you ever have complete certainty. And this is impossible or rather we never know if we have such completeness. We assume we don't
- Thus only will science be predictable where we can pin everything down and in terms of physics, things like the solar system are closest to this sort of closure. Hence Newton's Laws 'work' sort of, not quite, for we need relativity as we scale up.
- And of course as we ourselves are assumed to be in control of our destinies, we can invent or change the future. If we make a prediction and don't like it, then we can change it.
- As well, the future is always subject to uncertainty – the classic example being a ***Black Swan*** event.

- It is worth saying that there are many eminently quotable scholars and scientists who have made the point that even if we cannot predict the future, we can invent it
- And this makes science much closer to design than we usually consider.
- This goes to heart of discovery. Very often with uncertain systems, different theories can account for the same thing. This begs the question that if we can discover more than one explanation, are we engaged in ***discovery*** or ***invention***?
- Now some kinds of discovery are a kind of invention. If we develop a theory even before testing it against independent data to see if it can be validated, in a sense we have invented it – the process of discovery may then begin after the theory has been invented – so invention here is not discovery
- We also need to throw into this nexus the idea that we also are intent on prediction. To take this further let us look at the scientific method

Induction and Deduction: About Discovery

- The scientific method usually pits an idea called a theory (or these days a model) against data. Theory enables us to make a prediction which we then test to see if the prediction is made good against data. This is ***deduction***.
- Now the other way around is that we start with data and we then search for pattern which we then elevate to the status of theory. This is ***induction***.
- We need both to test a theory: where does theory come from – well from ***induction*** – and how do we know if it is any good other than the case from which the theory is derived – well we use ***deduction*** – but on different data. And so the ***inductive-deductive*** loop goes round and round gathering different case studies, situations each time around – possibly modifying the theory but always waiting for when ***it can be falsified.***
- Hence 'Popper' – you can only falsify – you cannot prove truth

- How does discovery, invention and prediction fit into all this?
- Well we need to predict so that we can falsify, but we know we cannot predict because our models and theories are incomplete, inevitably – we need to predict but we can't
- The key question though is can we ***discover patterns*** in the data that we then elevate to theory. In this sense, we often can only develop theory by taking such pattern and then ***inventing a theory*** of how these patterns can be embedded into the theory or hypothesis.
- In all of this then, prediction is thus a mixture of invention and discovery – neither one nor the other
- Now we usually accomplish this in science with respect to the past and the present rather than the future. But our theories can predict any event past, present or future. The future of course is unknowable. Hence there is an obvious asymmetry.
- Of course this is given a different context in terms of our understanding of past cities or our invention of new



MIT Press, Cambridge MA, 2018.

- I develop applications of these ideas in Chapter 1 of my recent book.
- This is hardly about geographical systems but a lot of theoretical geography is based on city systems
- And ideas about the future pertain to the use of such theory in urban planning
- I provide some interesting quotes about whether or not we invent or predict, from people like Alan Kay and Denis Gabor which essentially said that 'we cannot predict the future but only invent it'

Thinking About Cities, About the Scientific Object

- Cities are complex systems that grow from the ground up and are essentially ill-defined across many of their dimensions.
- Our theories tend to be either based on some sort of deductive analogies with physical systems – such as using gravitational models or, or on inferring patterns of locational-spatial behaviour that are inferred and modelled using a variety of multivariate methods.
- In fact most of our theories tend to be non quantitative but my points are much more clearly indicated using quant methods and I will focus on these
- Cities have so many dimensions that our models tend to be gross simplifications of the real thing and as there is no consensus about what the real thing is, then they are inevitably open systems

- In fact, there are so many processes that define a city and need to be simulated by any good theory, that inevitably we leave most things out and we tend to therefore to construct plausible sequences of how we intuitively think a city works.
- There are mechanisms which produce plausible links – we very often can't test them and we can't really observe them but they make sense
- This is not a good criterion for a theory but it is a very widely used strategy as we will see below.
- We often do the same when it comes to searching for patterns in data – often we construct mechanisms that produce good patterns and some of these can be interpreted as being plausible but some enable us to produce good fits with data
- Machine learning often falls into this category
- So my argument is that when we devise theories of how cities work – how geography effects cities, we invent as much as we discover and discovery is often invention and vice versa

The Question: Theoretical Geography: Discovered or Invented

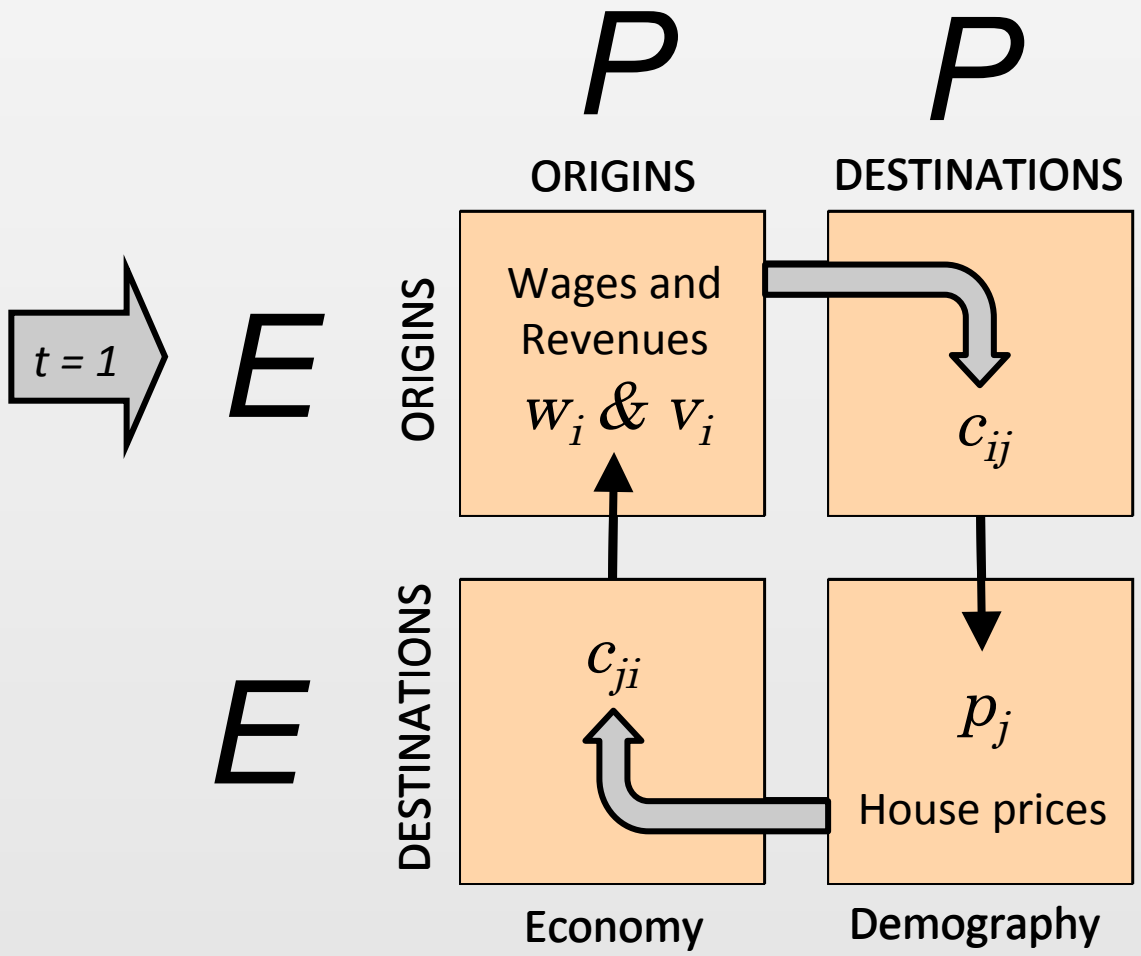
- Before I show some examples from my own work – let me rephrase what I believe in terms of ***discovery*** and ***invention***.
- We are never able to definitively extract a theory of how any bit of a city works from data along
- We always need to add something by way of an auxiliary hypothesis from our experience, and this is tantamount to our invention of some of the theory
- Discovery is the process of extracting meaning from the patterns we see in data and even this might involve invention.
- If developing theory is always a question of extracted pattern & plausible invention, every time we predict, we discover & invent. In fact prediction is discovery **AND** invention, and this is particularly resonant with respect to the future.

Three Examples

Cities as Machines, Mechanisms that are Invented not Discovered

- Now let us build a model of city where employment and population influence one another – to figure out where employment is we need to know population and vice versa
- We link these two sectors with respect to movement patterns – from work to home is the journey to work and from home to work can be see as the demand for work from home, in short
- ***E=f(P)*** and ***P=g(E)*** and if these variables employment ***E*** and population ***P*** are located in different zones the links between them are trips or movements.
- From data we know that distance or travel cost between locations usually follows well defined gravitational attraction-deterrence and this has been validated many times
- Now how do we take around of the simultaneity of ***E*** and ***P***

Here is the model structure – this itself is sort of invented



An Idea of the Model:

Essentially a couple of nested spatial interaction-discrete choice-like models – one for residential, the other for the retail sector, cross sectional static. We can extend these to many other sectors – the structure is sort of invented

$$T_{ij}^{k} = O_i \frac{D_j^{obs} \exp(-\beta^k d_{ij}^k)}{\sum_k \sum_j D_j^{obs} \exp(-\beta^k d_{ij}^k)} \qquad k = 1,\ k = 2,\ k = 3$$

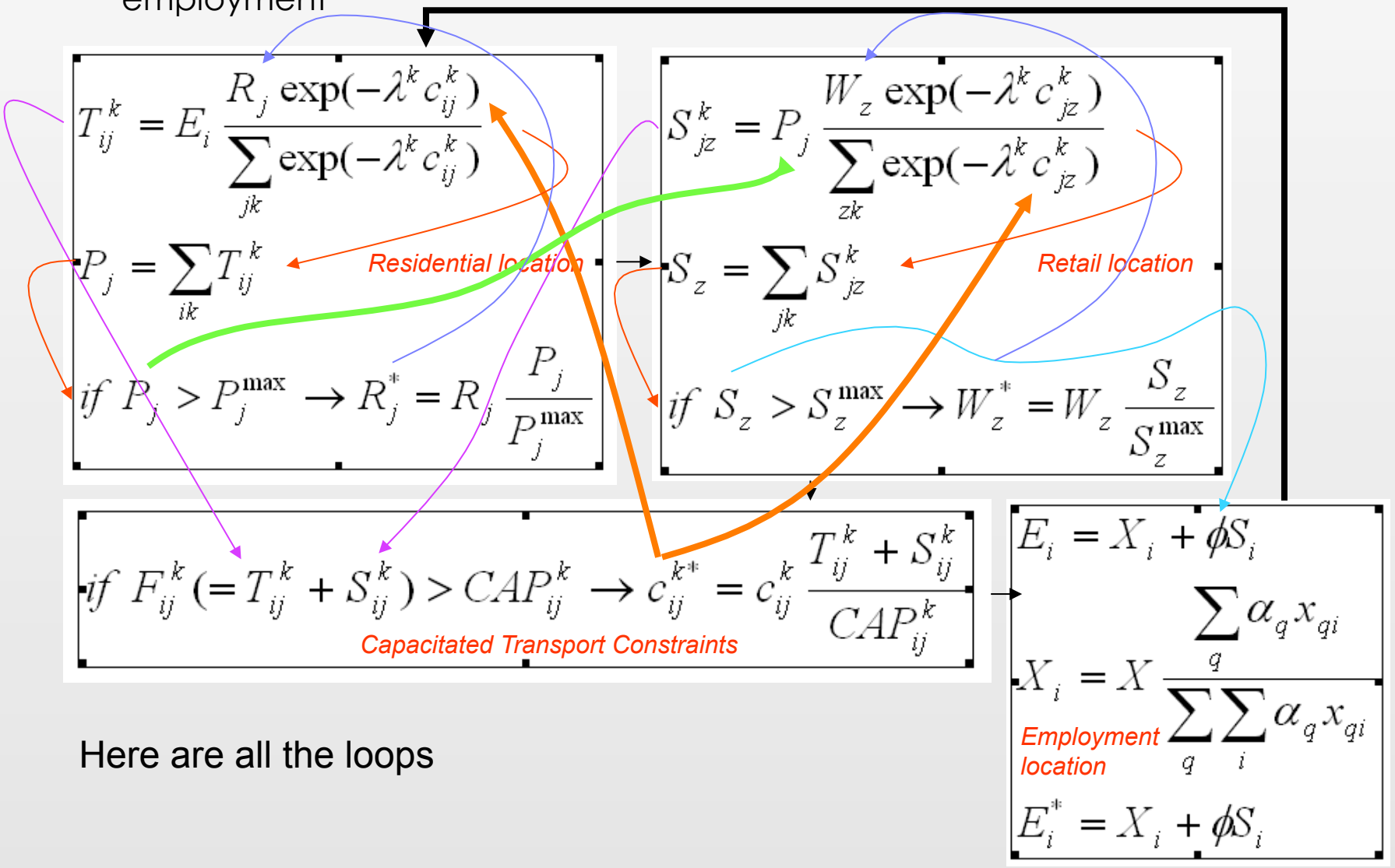
$$road, rail, bus$$

$$P_j = \sum_k \sum_i T_{ij}^k$$

$$S_{j\ell}^k = P_j \frac{W_\ell^{obs} \exp(-\lambda^k d_{j\ell}^k)}{\sum_k \sum_\ell W_\ell^{obs} \exp(-\lambda^k d_{j\ell}^k)}$$

$$O_\ell = \sum_k \sum_j S_{j\ell}^k$$

We begin with the residential, then retail sector, then trips capacities, and finally employment



Here are all the loops

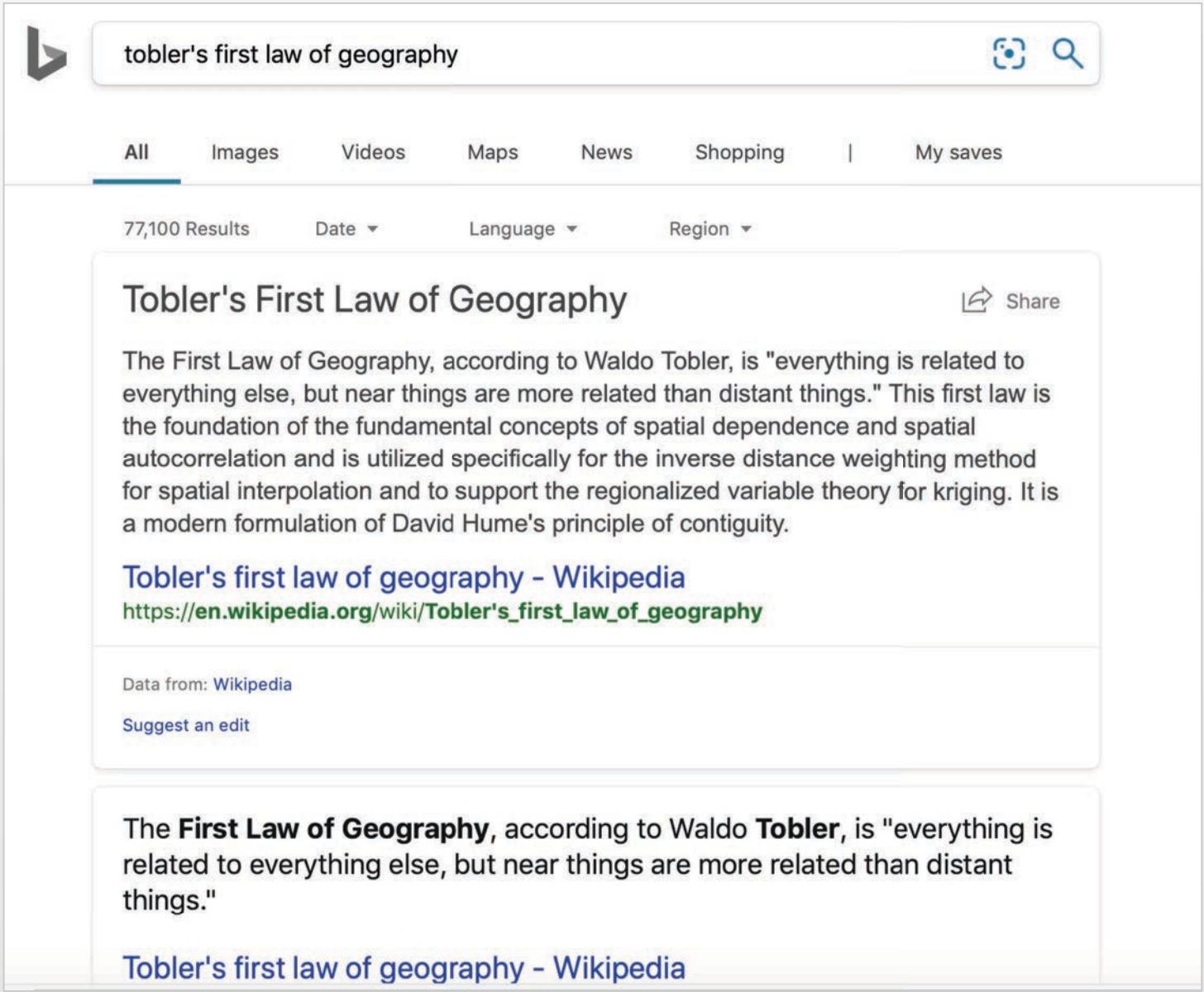
- Now it is well known that if we add some employment, then it will generate population – this is a sort of identity – a default – but then the population would generate some new employment and then so on
- This is the **multiplier effect** – so we can break into the cycle by articulating the input employment as the employment that drives the process of multiplicative increments
- But of course the model is static cross sectional – the multiplier effects summarise what has happened in all the times up to the cross section, so where does the multiplier come from
- It is a concept that was articulated over 100 years ago in economics and popularised by Keynes as a mechanism
- Using it for this type of model, it is a mechanism that enables us to make predictions with the model. It is not observed.
- We can observe it – to an extent it is an invention
- Many of our mechanisms in models are like this. Here is an outline of these sorts of model

Search for Urban Patterns: Patterns of Human Behaviour

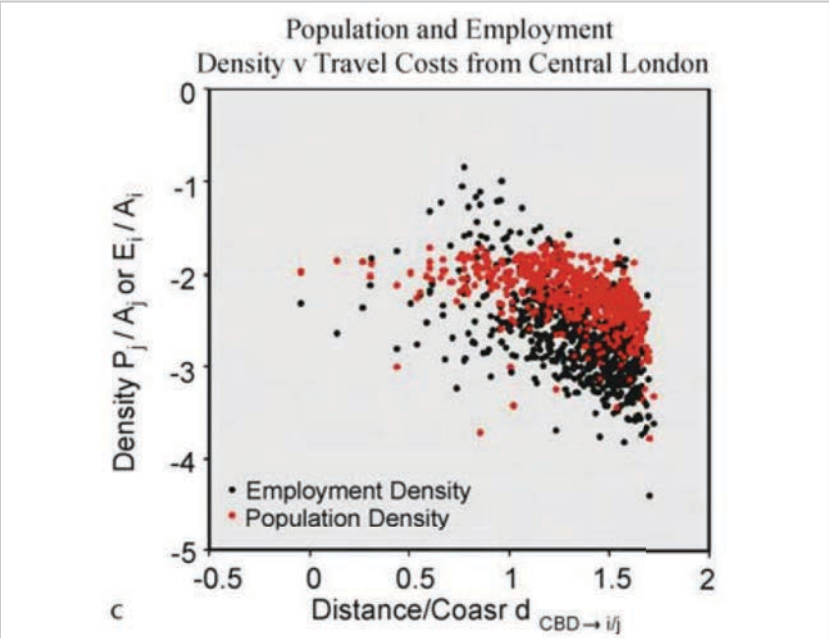
In terms of patterns in data, then there is a long tradition of examining how people respond to distance or some measure of deterrence and essentially all other things being equal – ceteris paribus – trips between any two locations decline with distance between those locations.

There is plenty of empirical evidence that this indeed is the case. It also goes for density in cities and for rents and so on. There is a reasonably well-developed edifice in urban economic theory explaining all of this – or at least being consistent with it.

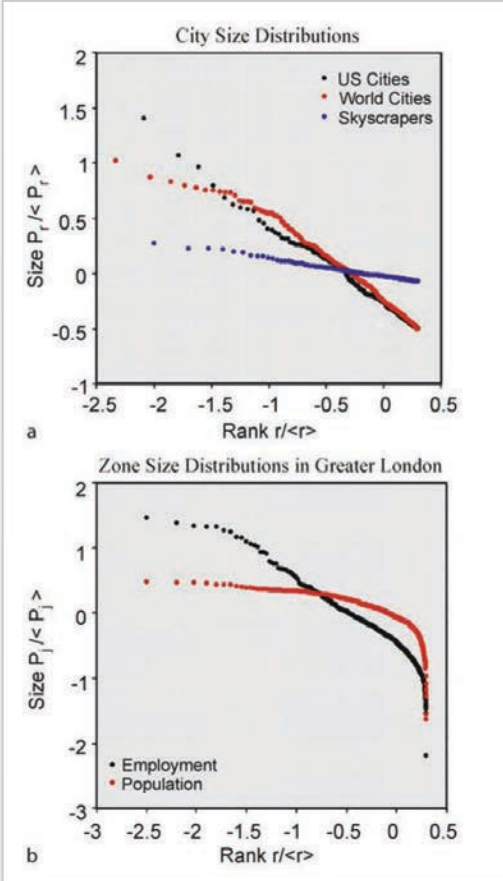
All this is contained in the generalisation which is called Tobler's **First Law of Geography** – accredited to Waldo Tobler but in fact a law as old as the hills – goes back Leonardo and plenty of examples all the way since Newton first articulated it in his second law of motion

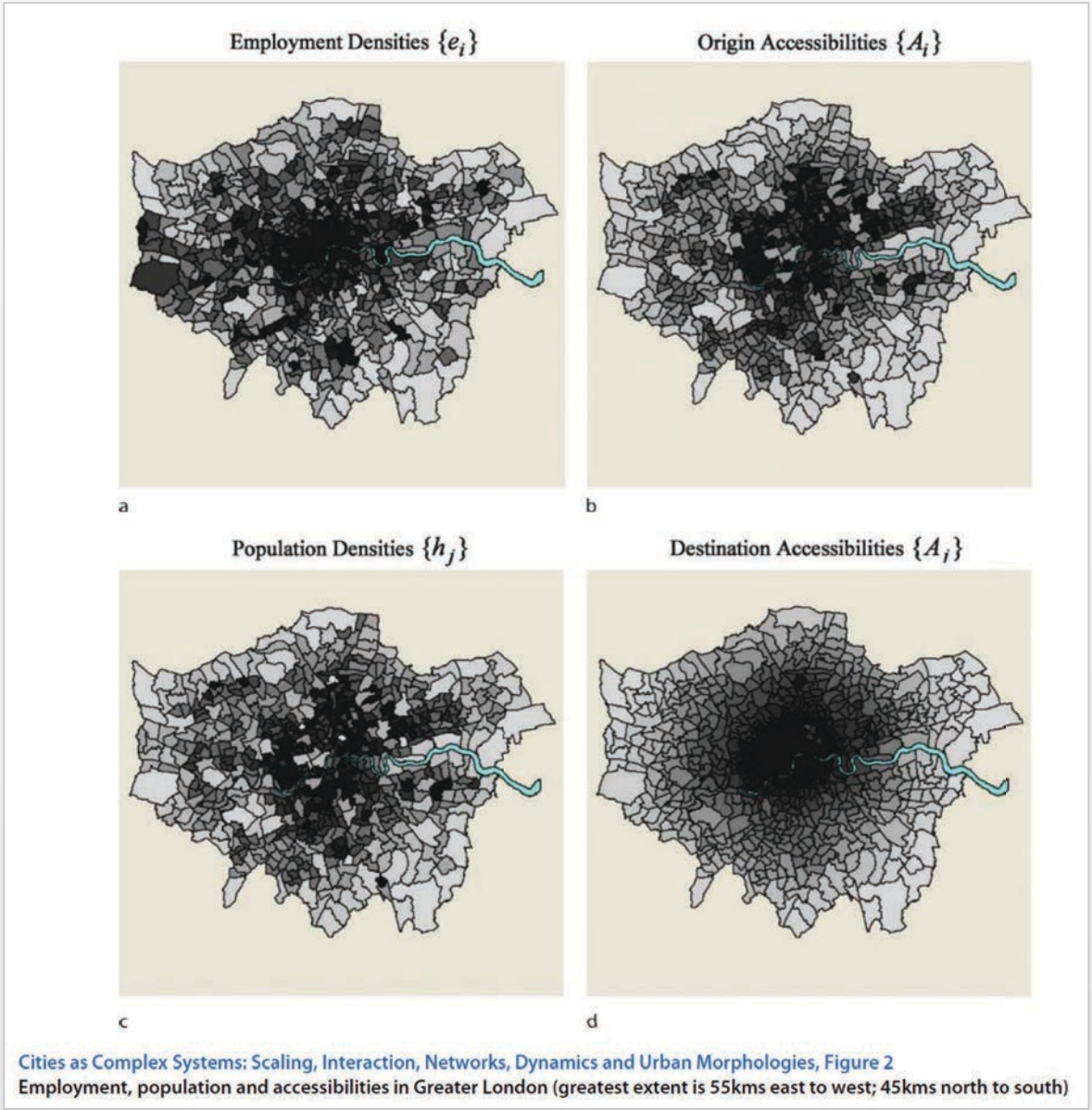

tobler's first law of geography
All
Images
Videos
Maps
News
Shopping
My saves
77,100 Results
Date
Language
Region
Tobler's First Law of Geography
Share
The First Law of Geography, according to Waldo Tobler, is "everything is related to everything else, but near things are more related than distant things." This first law is the foundation of the fundamental concepts of spatial dependence and spatial autocorrelation and is utilized specifically for the inverse distance weighting method for spatial interpolation and to support the regionalized variable theory for kriging. It is a modern formulation of David Hume's principle of contiguity.
Tobler's first law of geography - Wikipedia
https://en.wikipedia.org/wiki/Tobler's_first_law_of_geography
Data from: Wikipedia
Suggest an edit
The First Law of Geography, according to Waldo Tobler, is "everything is related to everything else, but near things are more related than distant things."
Tobler's first law of geography - Wikipedia

Here are some empirical regularities: Basically this is what is observed and we need to articulate these to explain them but these models are invariantly invented. And there are many gravitational variants and developments in statistical physics, economic utility theory and so on



Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies, Figure 1
Scaling distributions in world cities and in Greater London





Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies, Figure 2
Employment, population and accessibilities in Greater London (greatest extent is 55kms east to west; 45kms north to south)

We invent mechanisms to explain the patterns that we see – densities and accessibilities

All these are analogies with social physics and laws of potential and gravitation which have been developed by a mix od induction and deduction over the last 100 years

Generalisation: How Can we Demonstrate This Robustness

My last example – and we could go through all our theories and models in this way – relates to the necessity to continue to defy falsification of the various theories against new data which continue to generalise the models.

Their 'truth' is always contingent and the models continue to be informed by new methods of validation and calibration. They are fit for propose so long as they resits falsification.

The process of generalisation and the continued reflection on their applicability – which in the last analysis is not science for science's sake but applications to improve the quality of life in cities – to improve the spatial system in its wider context –

In all of this, discovery, invention and prediction are all entangled within each other.

Developing Applications- Theory into Design: Present into Future

- We cannot predict, we can only invent and the double bind of predictions means that we cannot predict what we invent.
- We can engage in discovery in all these explorations and analyses and it is important that we recognize that context is all important.
- Our understanding of the past and present is very different from the future and we are left with the conundrum that if our theories are combinations of discovery and invention, to what extent does our invention of the future, interfere with our understanding of the past and the present which is also poorly based on different kinds of discovery and invention