



UNIVERSITY OF EDINBURGH
Business School



Accessing Big (Commercial) Data across a
global research infrastructure
Modelling Consumer Behaviour in China

Ashley D. Lloyd, Mario A. Antonioletti, Terence M. Sloan





A Big Data opportunity for Social Science?

[VIDEO] main points:

- (i) Technology and infrastructure now being developed at a scale that would allow data rates equivalent to all the words ever spoken by humans in less than one month
- (ii) Illustration of recent developments in global Research and Education network infrastructure, focussing on TEIN3 – the first large-scale network for the Asia-Pacific region
- (iii) Illustration of the Edinburgh-Beijing-Perth collaboration that models, predicts and visualises behaviour in global markets so that abstract mathematical models become useful to the creative process from which new products and services emerge
- (iv) This work has had to grapple with some 'big data' issues (next slide)



Big Data Problems

DATA ACCESS: Scientific data <> Socioeconomic Data

Scientific Data = 7-bit & “worthless” (Jim Gray)

Socioeconomic data = multi-byte, contextualised & sensitive

Q: Is access to commercial medical data easier or harder than NHS data?

RESEARCH METHODS: Correlation versus Causation

- Inductive or Deductive research lens?

Q: If we use big computers to find correlations in data, is that less valid than using big computers to test for correlations that we hypothesise?



Big Data Challenge to Social Science: Scale

EDUCATION

The Times - Public Agenda - Page 6: 11 September 2007

2007: Supermarkets oust sociologists

In an age when supermarkets routinely collect vast amounts of detailed information about their customers, traditional research carried out by academics has less of a role, writes **the Times Higher Education Supplement** (Sept 7).

and Roger Burrows, of York University. "Now they occupy an increasingly marginal position," they say.

"Commercial sociology" — such as the research carried out by supermarkets and telecom companies — dwarfs conventional research in terms of scale. And whereas in



"Hey, so what's a nice sociologist like you doing working on a survey like this?"



Big Data Challenge to Social Science: Scale and Scope

2009 (Hammersley): Inheritance of holism from anthropology: “we cannot understand what goes on within particular situations unless we can locate these within a larger picture”

EDUCATION The Times - Public Agenda - Page 6: 11 September 2007
Supermarkets oust sociologists

THERE'S a new endangered species in academia — sociologists, are being left reeling by their numbers.

In an age when supermarkets routinely collect vast amounts of detailed information about their customers, traditional research carried out by academics has less of a role, writes the Times Higher Education Supplement (Sept 7).

Fifty years ago, academic social scientists might be seen as occupying the apex of the social science research “pyramid”, writes Professor Mike Savage of Manchester University and Roger Burrows, of York University. “Now they occupy an increasingly marginal position,” they say.

“Commercial sociology” — such as as the research carried out by supermarkets and telecom companies — dwarfs conventional research in terms of scale. And whereas in



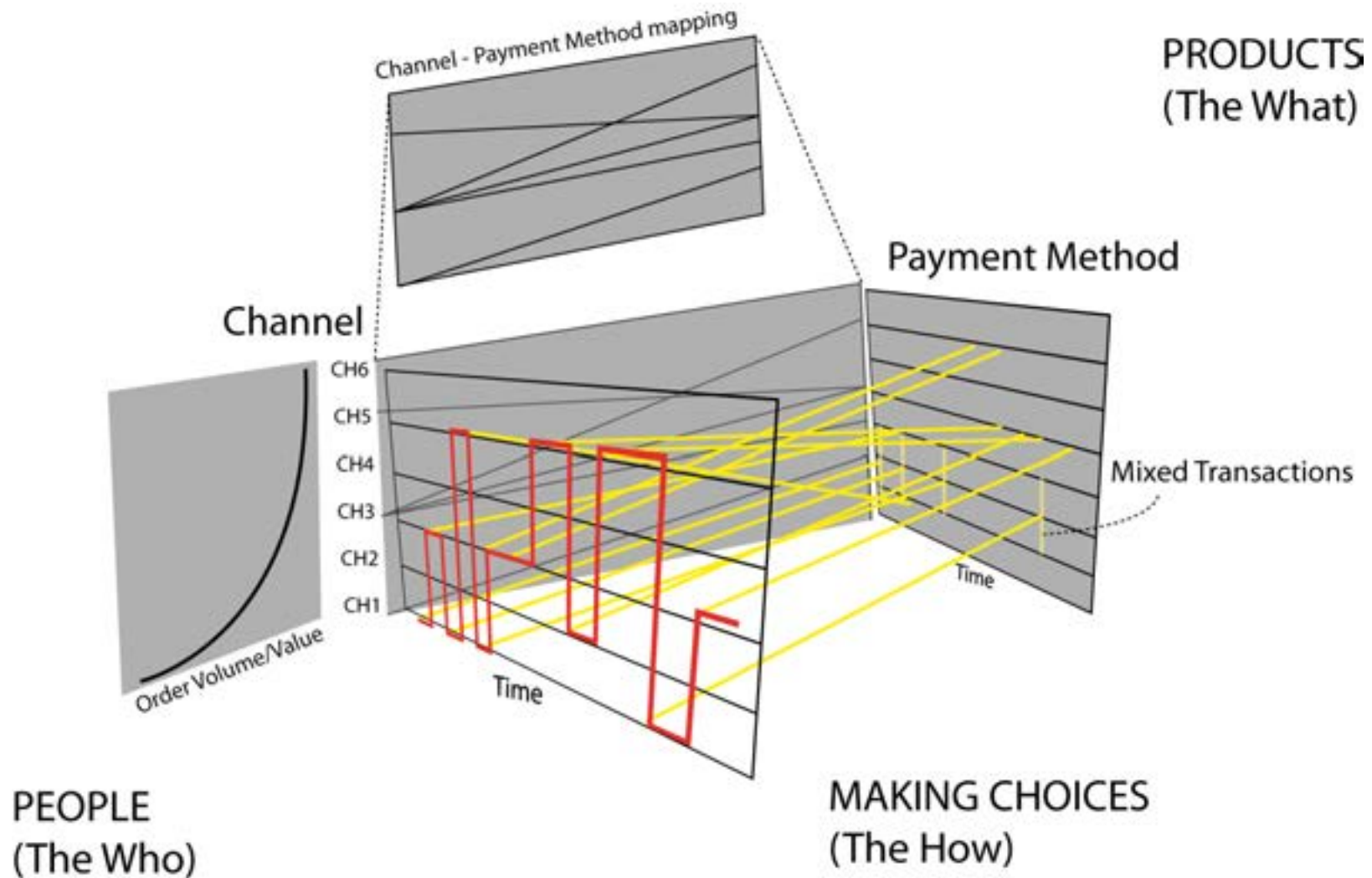
“Hey, so what’s a nice sociologist like you doing working on a survey like this?”

2010 (Sahlins) (Chicago) “anthropology appears to have become little more than the production of “thin” ethnographic accounts of the myriad, dispersed effects of global capitalism”

2014: ‘Virtual Ethnographies’ ignore distance altogether?

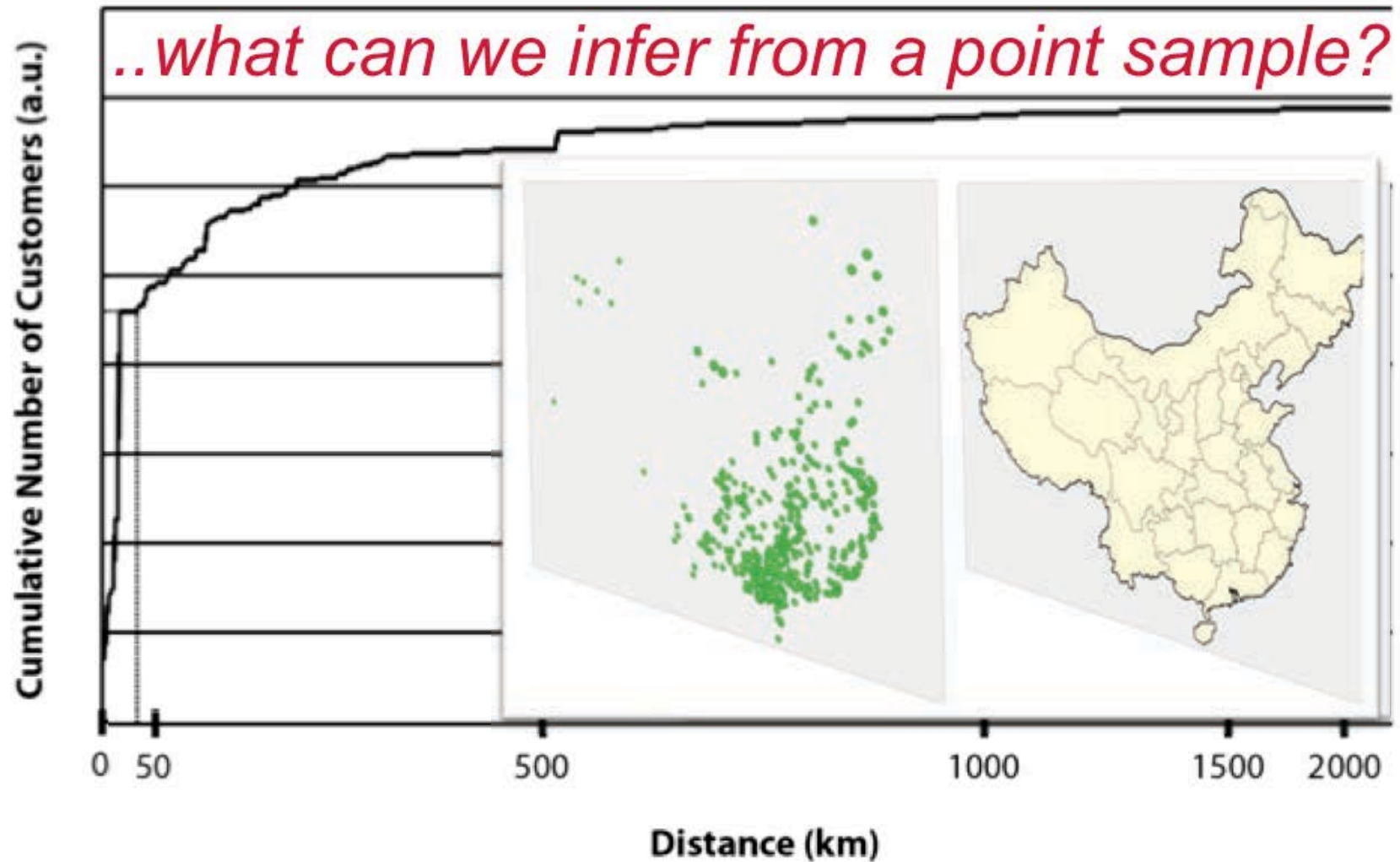


Look at fine-grained (big!) data on regional consumption



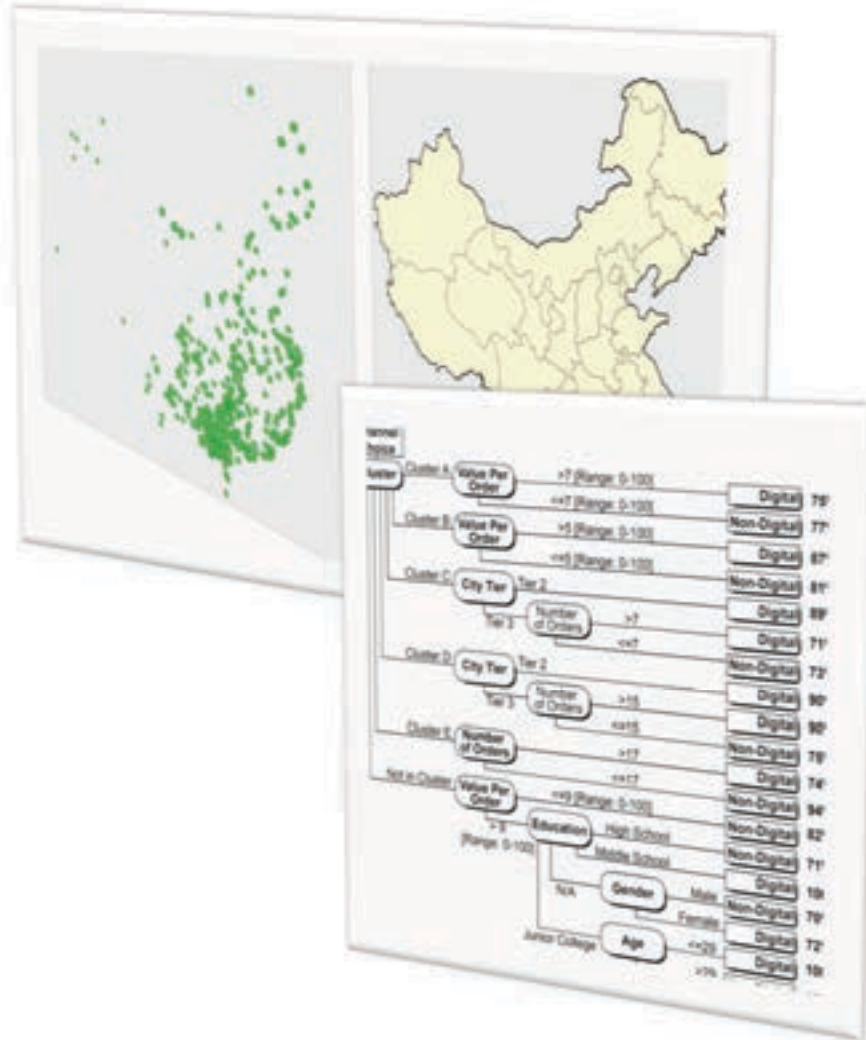


What does it mean to 'live' somewhere?





'Distance from Home' as an individual demographic

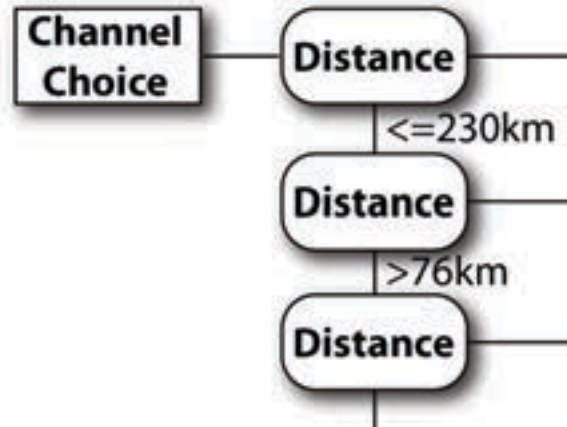


- Add distance as an individual demographic
- Probe two key channels in terms of barriers: Digital versus Physical
- Repeat analysis and re-build tree

Tree Attribute usage	
100%	Number of Order
100%	Value Per Order
23%	Distance from Home
6%	Occupation
2%	Age



Insights from new demographics: 'Distance from Home'



Jeyaraj, A., Rottman, J.W., Lacity M.C. (2006):
A review of the predictors, linkages, and biases in IT innovation adoption research,
Journal of Information Technology 21: 1–23.

.. identified the most significant factors influencing technology adoption and diffusion based on a review of 99 empirical studies. The total number of unique independent variables identified was 135, and the number of relationships was 505.

- ***Distance was not examined in any of the studies!***
- ***Are we tied to old-fashioned ways of explaining behaviour, designing products and media segmentation?***
- ***Is this why the digital divide is actually growing in the US?***

LEAVING THE LAND

China's Great Uprooting: Moving 250 Million Into Cities

Articles in this series look at how China's government-driven effort to push the population to towns and cities is reshaping a nation that for millenniums has been defined by its rural life.

[点击查看本文中文版](#) | [Read in Chinese >](#)

JUNE 15, 2013



250,000,000

Dealing with Data:

Q1: what are the social and economic implications of a correlation-driven global economy when there is a big 'big data' skew in favour of the economically most active?