

What can an Open, Computational Approach Bring to the Social Sciences?

Charlie Rahal

LCDS and Nuffield College, University of Oxford

An invited talk at Hong Kong University, 2023



Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

1. Scientometrics, Bibliometrics, and the Sociology of Science.

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

1. Scientometrics, Bibliometrics, and the Sociology of Science.
2. Intersection of Civic Technology and Public Administration.

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

1. Scientometrics, Bibliometrics, and the Sociology of Science.
2. Intersection of Civic Technology and Public Administration.
3. Inequality, in multiple, hidden forms.

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

1. Scientometrics, Bibliometrics, and the Sociology of Science.
2. Intersection of Civic Technology and Public Administration.
3. Inequality, in multiple, hidden forms.
4. Model Uncertainty, and ways reduce or represent it.

Hello! Wonderful to meet you!

A computational econometrician by training, I *mostly* focus on:

- New Methods
- Fresh Datasets
- Open Science
- Reducing Uncertainty.

This typically falls into five substantive areas:

1. Scientometrics, Bibliometrics, and the Sociology of Science.
2. Intersection of Civic Technology and Public Administration.
3. Inequality, in multiple, hidden forms.
4. Model Uncertainty, and ways reduce or represent it.
5. Machine/Deep Learning in health and social contexts.

In the spirit of ‘Open Science’:

crahal.github.io/files/hku_talk.pdf

and *all* of the research you see here today is at:

github.com/crahal

github.com/OxfordDemSci

github.com/centre-for-care

github.com/uk-third-sector-database

Throughout the talk, I’ll frequently focus on ‘What’s Next?’.

What is ‘Open Science’?

‘Open science ... promotes transparency, accessibility, and collaboration throughout the research process. It involves making scientific knowledge, data, and methodologies openly available to the public, researchers, and other stakeholders, without restrictions or barriers.’

Source: ChatGPT

Pragmatically ‘Open Science’:

1. Allows creation of value and advancement.
2. Allows verification of the scientific record.
3. Allows public disclosure and engagement.
4. Allows science to become accessible/transparent.

Pragmatically ‘Open Science’:

1. Allows creation of value and advancement.
2. Allows verification of the scientific record.
3. Allows public disclosure and engagement.
4. Allows science to become accessible/transparent.

"If I have seen further, it is by standing on the shoulders of giants."

Source: Isaac Newton to Robert Hooke, in February 1676

What motivates a talk about ‘Open Science’?

- Four hard lessons learned as a doctoral student:

What motivates a talk about ‘Open Science’?

- Four hard lessons learned as a doctoral student:
 1. Make sure your coauthors have code.

What motivates a talk about ‘Open Science’?

- Four hard lessons learned as a doctoral student:
 1. Make sure your coauthors have code.
 2. Build robust research designs.

What motivates a talk about ‘Open Science’?

- Four hard lessons learned as a doctoral student:
 1. Make sure your coauthors have code.
 2. Build robust research designs.
 3. Critically analyze everyone else's data.

What motivates a talk about ‘Open Science’?

- Four hard lessons learned as a doctoral student:
 1. Make sure your coauthors have code.
 2. Build robust research designs.
 3. Critically analyze everyone else's data.
 4. Consider using open-source software.

What motivates a talk about 'Open Science'?

- Four hard lessons learned as a doctoral student:
 1. Make sure your coauthors have code.
 2. Build robust research designs.
 3. Critically analyze everyone else's data.
 4. Consider using open-source software.

Disclaimer: Presently the LNL Co-ordinator for UKRN!

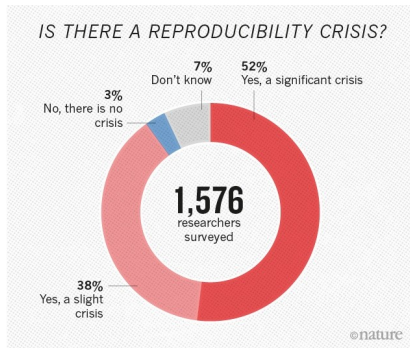
How it started:

Why Most Published Research Findings Are False

John P. A. Ioannidis

Published: August 30, 2005 • <https://doi.org/10.1371/journal.pmed.0020124>

How it's going:



These slides are made with an open source typesetting language!

I_make_this_slide.tex

```
\begin{frame}[]  
  These slides are made with an open source typesetting language!  
  \begin{center}  
    \begin{scriptsize}  
      \begin{minipage}{.95\linewidth}  
        \begin{myenv}{guess$_\_what$__this$__does.py}  
          \begin{minted}{latex}  
            # Hello friends :)  
          \end{minted}  
        \end{myenv}  
      \end{minipage}  
    \end{scriptsize}  
  \end{center}  
\end{frame}
```

- This slide and the next emphasize not just transparency, but the beautiful flexibility of open source tools!

Open source statistical code can come in the form of Python (or R):

guess_what_this_does.py

```
import json, io, imageio, requests
from PIL import Image

def get_cat_bounded(url):
    api_ret = requests.get(url)
    if api_ret.status_code==200:
        text = json.loads(api_ret.text)
        while (text[0]['width']<=400 or text[0]['width']>=600) or \
            (text[0]['height']<=400 or text[0]['height']>=600):
            api_ret = requests.get(url)
            if api_ret.status_code==200:
                text = json.loads(api_ret.text)
        r = requests.get(api_text[0]['url'])
        return Image.open(io.BytesIO(r.content))

url = 'https://api.thecatapi.com/v1/images/search'
images = []
for _ in range(0, 50):
    images.append(get_cat_bounded(url))
imageio.mimsave('CATS.gif', images, fps=1)
```

Thats right: Cats!

Scientometrics

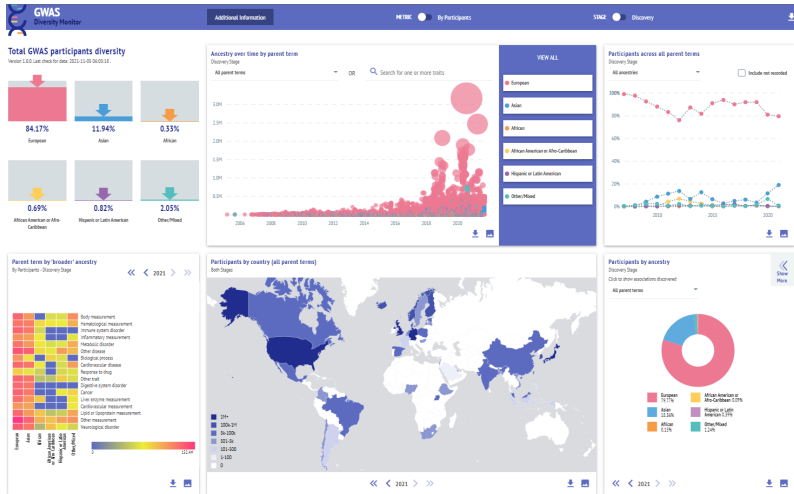
I frequently work within the areas of 'Scientometrics', 'Bibliometrics', and the 'Sociology of Science':

- **Scientometrics**: concerned with measuring and analysing science itself.
- **Bibliometrics**: statistical methods to analyze quantitative information pertaining to research.
- **Sociology of Science**: how society influences discovery, knowledge production, and advancement.

These fields are naturally related to 'Open Science', and are well-facilitated by a computational approach.

Genome Wide Association Studies

- My early scientometric work analyzed GWAS.
 - What is a GWAS?
 - Why is this type of generative inequality so important?
- We take lists of GWAS from the European Bioinformatics Institute, then call scientific meta-data from PubMed.
- We published two papers which:
 1. Conduct a thorough scientometric review.
 2. Launch a substantial interactive dashboard.
- Conducting genomic work ethically of paramount importance.

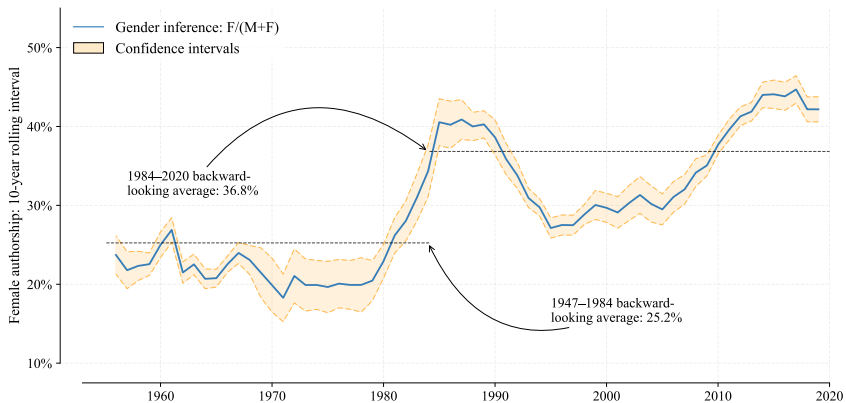


‘The GWAS Diversity Monitor Tracks Diversity in Real Time’
with Melinda C. Mills

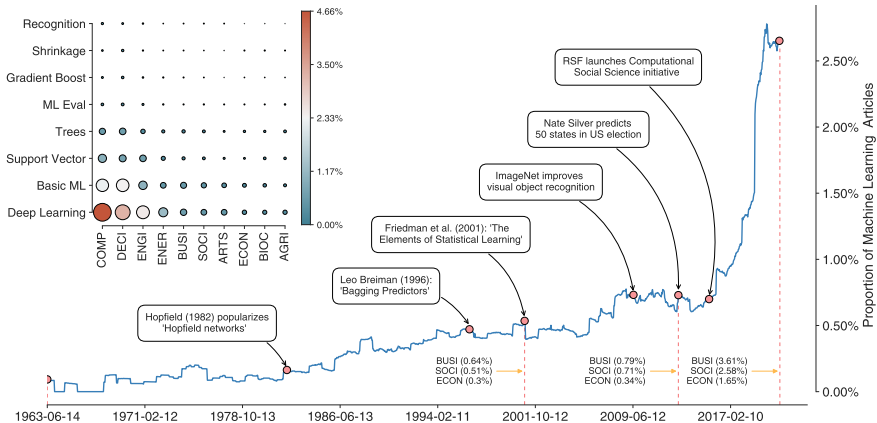
What's Next? GWAS Diversity by Funder

- The GWAS Diversity Monitor was generally a **huge** success.
 - We are always striving to generate maximal impact.
- One mechanism for impact: delineate analysis by funders.
- *Some* funders have policies on diversity of participant pools.
- Our work will generate – in real time – analysis showing differential diversity.
- Other milestones: extend by ‘Polygenic Scores’.

Other Scientometric Studies

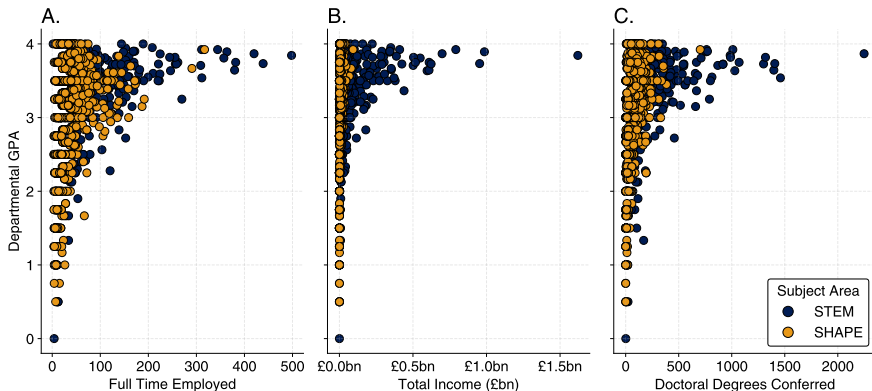


‘Population Studies at 75 Years: An empirical review’
With Melinda C. Mills



'The Rise of Machine Learning in the Academic Social Sciences'

With Dave Kirk and Mark Verhagen

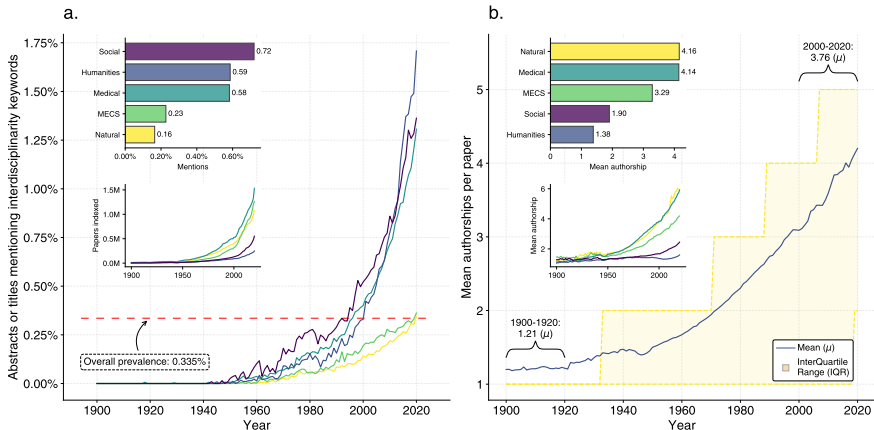


'The SHAPE of Impact'

With M.C. Mills, S. Wagner, M. Verhagen, B. Zhao, and D. Leasure, (2023)

What's Next? The Evolution of Science Lab!

- With bespoke agreement, we have access to Dimensions.
- We've used this to reconstruct the **entire scientific record***.
- This is going to power an 'Evolution of Science Lab'. RQs:
 1. How have scientific ideas evolved over the (very-) long run?
 - What new academic disciplines have gradually emerged?
 2. How are scholars internationally migrating over time?
 3. How have patterns in multi-/inter-disciplinarity changed?



'The Evolution of Interdisciplinarity'

With Kyla Chasalow and Melinda C. Mills

‘Administrative Data is an invaluable resource for Good.’

ADR UK (2018)

- UK leads the world in terms of ‘openness’ of government data.
- The value of public government data is increasingly recognised.
 - Example: the rise of ‘Government Data Science’.
- For better or worse:
 - I’m more interested in ‘hard to use’, semi-structured data.
 - I frequently work with national statistical agencies to add value.

Procurement Data

Another example of the unstructured data I work with:

- David Cameron opened data on government contracts (2010).
- This provided a totally new type of semi-structured information.
- From from local authorities, to the NHS, to central government.
- This comes in the form of spent line-items of various values.
- This sounds like it might be useful for academic research...
- The raw data is **unbelievably** messy.

Procurement Data

Another example of the unstructured data I work with:

- David Cameron opened data on government contracts (2010).
- This provided a totally new type of semi-structured information.
- From from local authorities, to the NHS, to central government.
- This comes in the form of spent line-items of various values.
- This sounds like it might be useful for academic research...
- The raw data is **unbelievably** messy.

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.
2. Every file is in a heterogeneous format in a different location.

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.
2. Every file is in a heterogeneous format in a different location.
(There are **thousands** of locations on the internet.)

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.
2. Every file is in a heterogeneous format in a different location.
(There are **thousands** of locations on the internet.)
3. Data commonly converted from a tabular file...

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.
2. Every file is in a heterogeneous format in a different location.
(There are **thousands** of locations on the internet.)
3. Data commonly converted from a tabular file... **into** a .pdf!
(This is a **direct** downgrade in the Berners-Lee 5* schema.)

Lets list a few of the problems with procurement data, which is a good example of semi-structured data more generally:

1. Some UK organizations don't believe in their legal obligation.
2. Every file is in a heterogeneous format in a different location.
(There are **thousands** of locations on the internet.)
3. Data commonly converted from a tabular file... **into** a .pdf!
(This is a **direct** downgrade in the Berners-Lee 5* schema.)
4. Naturally, of course, there are no supplier UIDs.

Just one quick example, followed by what can be made possible:

Expenditure Over Threshold

24 / 49

A3131

Expenditure Over Threshold

Department	Agency	City	Year	Category	Subcategory	Transaction number	AP Amount	VAT registration number	Postcode	Invoice number
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	COMMUNITY SERVICES	BEXLEY COUNCIL	11599453	666 000 00	6100071853
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	COMMUNITY SERVICES	BEXLEY COUNCIL	11599455	100 000 00	6100071861
Department of health	NHS	Beccy	CGO	10/04/2021	Other Gen Supplies & Inv	COMMUNITY SERVICES	BEXLEY COUNCIL	11599457	61 666 69	6100071869
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	WYLLER RESTORATION	BEXLEY COUNCIL	11599492	64 000 00	6100071887
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	UNIVERSITY COLLEGE LONDON HOSPITALS NHS FOUNDATION TRUST	11607291	68 226 00	605485857
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING - NON ACUTE	NHS SOUTH EAST SSGCP PROJECT	11608000	200 000 00	6100071894
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	BEXLEY COUNCIL	11615604	30 000 00	6100071903
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Independent Sector	ACUTE COMMISSIONING	NH-HEALTH LTD	11615704	38 525 24	787424555
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	NEWWAY NHS FOUNDATION TRUST	11615743	254 350 00	HL00055682
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Voluntary Sector	HOSPICES	GREENWICH & BEXLEY COMMUNITY HOSPICE	11622337	218 000 00	871
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	LEWISHAM & GREENWICH NHS TRUST	11636201	414 200 00	84452527
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	COMMUNITY SERVICES	NHS PROPERTY SERVICES LTD	11637214	21 724 14	144025954
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PROGRAMME PROJECTS	NHS PROPERTY SERVICES LTD	11642864	25 672 53	48413715
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	CARTFORD AND GRAVESEHAM NHS TRUST	11642923	-30 000 00	2000004184
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	CARTFORD AND GRAVESEHAM NHS TRUST	11642923	200 000 00	2000041339
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11654425	-3 662 119 34	CM00218970
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	COMMUNITY SERVICES	BEXLEY HEALTH LTD	11655539	26 332 44	BN10667
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	AMBULANCE SERVICES	LONDON AMBULANCE SERVICE NHS TRUST	11663375	587 757 29	605481192
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11666893	768 536 00	227 489
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11668507	760 530 34	233561
Department of health	NHS	Beccy	CGO	10/04/2021	Cont Care Adult 100% Fully Funded	FUNDED NURSING CARE	BURA CARE HOMES	11669592	26 512 00	2359711641
Department of health	NHS	Beccy	CGO	10/04/2021	Charges From CSU	RECHARGES	NHS SOUTH EAST CSU	11669995	85 249 00	GB054442045
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11671003	271 825 00	1206094
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	MOORFIELD EYE HOSPITAL NHS FOUNDATION TRUST	11680115	82 558 00	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	COMMUNITY SERVICES	OXLEAS NHS FOUNDATION TRUST	11671702	228 948 00	278586
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	COMMUNITY SERVICES	OXLEAS NHS FOUNDATION TRUST	11671707	141 917 00	37532
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	COMMUNITY SERVICES	OXLEAS NHS FOUNDATION TRUST	11671710	807 726 00	37592
Department of health	NHS	Beccy	CGO	10/04/2021	Income Mx <11%	BALANCE SHEET	HMC	11720565	47 733 41	604462333
Department of health	NHS	Beccy	CGO	10/04/2021	Income Mx <11% Student Loans	BALANCE SHEET	HMC	11720565	738 00	604462333
Department of health	NHS	Beccy	CGO	10/04/2021	National Insurance - 1 vNP-RMS	BALANCE SHEET	HMC	11720565	29 114 84	604462333
Department of health	NHS	Beccy	CGO	10/04/2021	National Insurance - 1 vNP-RMS	BALANCE SHEET	HMC	11720565	19 284 18	110090056570 MAR 18
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11724081	213 255 00	2020126
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11724081	268 141 00	229584
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11724081	27 309 000 00	1000041339
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11722209	266 424 00	231364
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11724114	224 793 00	2020126
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	RETAIL SALE EVENT	LEWIS & CLARK	11724081	440 310 00	2020126
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	QUEEN VICTORIA HOSPITAL NHS FOUNDATION TRUST	11735822	-20 056 34	1365
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	MOORFIELD EYE HOSPITAL NHS FOUNDATION TRUST	11735829	55 540 00	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	COMMUNITY SERVICES	OXLEAS NHS FOUNDATION TRUST	11748440	249 505 00	175451
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	QUEEN VICTORIA HOSPITAL NHS FOUNDATION TRUST	11748386	47 300 00	201000982
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	ARTFORD AND GRAVESEHAM NHS TRUST	11771723	162 132 00	2020126
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	CARTFORD AND GRAVESEHAM NHS TRUST	11771723	2 691 562 00	1000041339
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	MENTAL HEALTH CONTRACTS	SOUTH LONDON AND MAUDSLAY NHS FOUNDATION TRUST	11777727	25 202 00	80002737
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES - NON ACUTE	BEXLEY COUNCIL	11801176	100 000 00	6100071903
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	MENTAL HEALTH CONTRACTS	OXLEAS NHS FOUNDATION TRUST	11725325	378 937	378937
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	MENTAL HEALTH CONTRACTS	OXLEAS NHS FOUNDATION TRUST	11801176	36 000 00	378937
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	PROPERTIES AND FACILITIES	OXLEAS NHS FOUNDATION TRUST	11801176	100 000 00	6100071903
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	OXLEAS NHS FOUNDATION TRUST	11818020	258 781 00	378937
Department of health	NHS	Beccy	CGO	10/04/2021	Hithou Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	OXLEAS NHS FOUNDATION TRUST	11818020	55 506 00	378937
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	REPAIRS SERVICES	BEXLEY COUNCIL	11818100	20 000 00	605481192
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Voluntary Sector	DAY SERVICES	GREENWICH & BEXLEY COMMUNITY HOSPICE	11818190	276 741 00	589
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	DAY SERVICES	DAY SERVICES INTERNATIONAL	11818190	25 000 00	11818190
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11824000	1 960 511 00	882426016
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Other Public Sector	LEARNING DIFFICULTIES	BEXLEY COUNCIL	11854467	250 000 00	6100071247
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	UNIVERSITY COLLEGE LONDON HOSPITALS NHS FOUNDATION TRUST	11861126	1 167 126	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES	BELLEROVE ROAD SURGERY	11963391	65 592 39	BA-0320156
Department of health	NHS	Beccy	CGO	10/04/2021	Healthcare Srv Res CCG	NHS 111	NHS BROMLEY CCG	11963392	71 747 88	1002102149
Department of health	NHS	Beccy	CGO	10/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES	CHAP-CAO TOWN SURGERY	11963392	26 703 00	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963393	1 137 348 98	234497
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	GUY'S & ST THOMAS HOSPITAL NHS FOUNDATION TRUST	11963394	1 455 963 00	6054823417
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963395	1 133 348 98	6054823417
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	LEWISHAM & GREENWICH NHS TRUST	11963396	-129 548 00	84452527
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963397	1 167 569 00	234736
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963398	-200 561 00	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963399	229 254 29	233442
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11963400	-411 000 00	CM0225493
Department of health	NHS	Beccy	CGO	10/04/2021	Cont Care Adult 100% Fully Funded	CHC ADT FULL FUND PERKS HLTH BUD	BEXLEY COUNCIL	11963401	269 574 97	6100071885
Department of health	NHS	Beccy	CGO	10/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	UNIVERSITY COLLEGE LONDON HOSPITALS NHS FOUNDATION TRUST	11963402	68 226 00	605485857

A3131

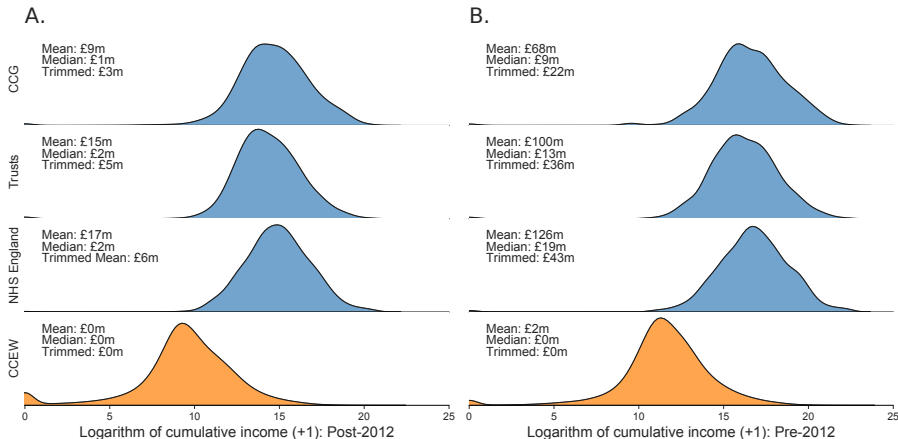
Expenditure Over Threshold

Department	Agency	City	Year	Category	Subcategory	Transaction number	AP Amount	VAT registration number	Postcode	Invoice number
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599453	666 000 00		6100017653
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599455	100 000 00		6100017661
Department of health	NHS	Bexley	CGO	10042601	Other Gen Supplies & Inv	COMMUNITY SERVICES	11599456	10 000 00		6100017669
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	WYLLER RESTORATION	11599452	64 000 00		6100017657
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	116007291	68 226 00	654485657	610009844
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING - NON ACUTE	116007290	68 226 00	654485657	610009844
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	11615604	30 000 00		6100017655
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Independent Sector	ACUTE COMMISSIONING	11615704	38 526 24	787424655	HL00056682
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11615744	254 250 00		6100017659
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Voluntary Sector	HOSPICES	11622337	218 000 00		871
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11636201	444 140 00		6100017667
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	NHS PROPERTY SERVICES LTD	11636202	218 124 14		6100017668
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PROGRAMME PROJECTS	11642864	25 672 53		48413715
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11642923	-30 000 00		2000004184
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11642924	-30 000 00		2000004185
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11654425	-3 662 116 34		CM00218970
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	AMBULANCE SERVICES	11665375	26 332 45		6100017665
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11665376	587 757 29	654481192	10011114
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11665983	768 536 00		227 489
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11665985	768 536 00		233561
Department of health	NHS	Bexley	CGO	10042601	Cont Care Adult 100% Fully Funded	FUNDED NURSING CARE	11666922	26 512 00	2359711641	ALM036320YOM036585
Department of health	NHS	Bexley	CGO	10042601	Charges From CSU	RECHARGES	11666995	85 249 00	GB55442045	731180594
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11673003	278 328 00		120606
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11680115	82 568 00		600326903
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701702	228 968 00		278586
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701707	141 917 00		37532
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701710	807 726 00		37592
Department of health	NHS	Bexley	CGO	10042601	Income Mx < 11%	BALANCE SHEET	11720565	47 733 11	654462333	100900555770 MAR 18
Department of health	NHS	Bexley	CGO	10042601	Income Mx < 11%-Student Loans	BALANCE SHEET	11720566	738 00	654462333	100900555770 MAR 18
Department of health	NHS	Bexley	CGO	10042601	National Insurance - 1 vNED-RMS	BALANCE SHEET	11720568	29 114 84	654462333	100900555770 MAR 18
Department of health	NHS	Bexley	CGO	10042601	National Insurance - 1 vNED-RMS	BALANCE SHEET	11720569	19 284 18	654462333	100900555770 MAR 18
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11724081	213 255 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11724082	288 141 00		229584
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11724083	27 126 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11722209	204 338 00		231364
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11724114	254 139 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	HEALTH CARE EQUIPMENT	11724115	200 000 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11735822	-20 056 34		1365
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11735829	30 540 00		600326908
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	COMMUNITY SERVICES	11748440	764 905 00		17541
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11748386	47 420 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11771725	167 122 00		2001026
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11771733	2 691 562 00		1000041308
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11777627	25 202 00		800027387
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Independent Sector	COMMUNITY SERVICES - NON ACUTE	11777630	56 381 00		6100017660
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11782425	1 728 242 00		37597
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11801176	36 000 00		37606
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	UTILITIES AND FACILITIES	11801175	10 000 00		48413851
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11818020	258 781 00		37605
Department of health	NHS	Bexley	CGO	10042601	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11818040	55 506 00		37604
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	HEALTHY SUPPORTIVE CARE	11818100	10 000 00		600326903
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Voluntary Sector	DAY PROGRAMMES	11818190	276 741 00		585
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Independent Sector	DAY PROGRAMMES	11818585	25 000 00	654226016	6100017657
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11824000	1 960 531 00		243431
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Clinic Public Sector	LEARNING DIFFICULTIES	11854467	250 000 00		6100017647
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11854468	250 000 00		6100017647
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	COMMUNITY SERVICES	11854469	117 122 00		6100017647
Department of health	NHS	Bexley	CGO	10042601	Healthcare Sln Res CCG	NHS 111	11854470	71 747 88		6100017647
Department of health	NHS	Bexley	CGO	10042601	Clinical/Medical-Independent Sector	COMMUNITY SERVICES	11854471	26 703 00		6100017647
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11903373	1 137 848 59		234487
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11931140	1 455 963 00	6544923417	13005559
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11933582	1 133 524 00	6544923417	1000041308
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11933581	-129 548 00		8466505
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11951210	1 187 569 00		233736
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11951211	-205 156 00		CM00218929
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11951217	229 254 29		233542
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	PLANNED CARE	11951290	-411 000 00		CM00218929
Department of health	NHS	Bexley	CGO	10042601	Cont Care Adult 100% Fully Funded	CHC ADT FULL FUND PERKS HLTH BUD	11951291	265 574 77		6100017665
Department of health	NHS	Bexley	CGO	10042601	Heath Sln Res NHS Trust-Contract Baseline	ACUTE COMMISSIONING	11959235	68 226 00	654485657	610009844

[illegible]



25 / 49

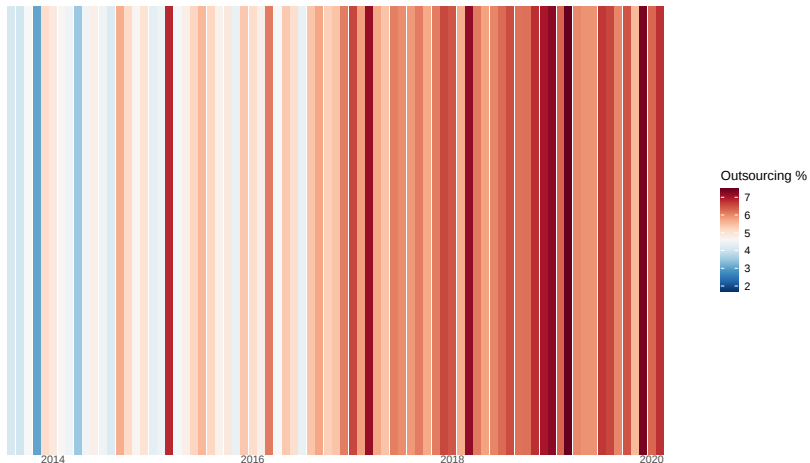


‘The Role of Non-Profits in Public Health Service Provision’

With John Mohan

What's Next? A new database of all Third Sector Organisations!

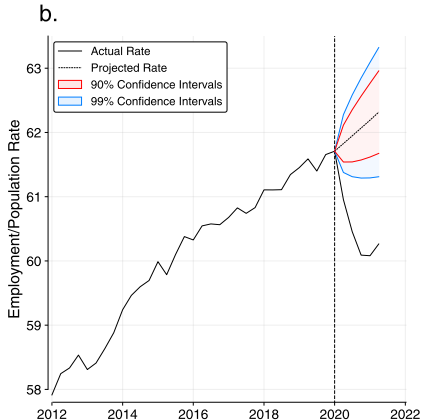
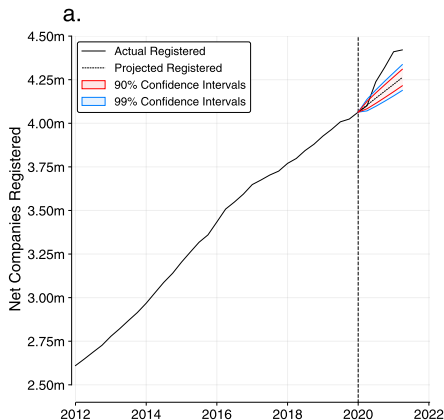
Creeping Privatisation in England's NHS



Source: Raw CCG payments approximately matched to Companies House and the Charity Commission, then manually verified

‘Creeping Commercialization in the NHS’

By Ben Goodair and Aaron Reeves



‘Modelling Labour Market Inequalities and the Covid-19 Pandemic’
With Aaron Reeves and Naomi Muggleton

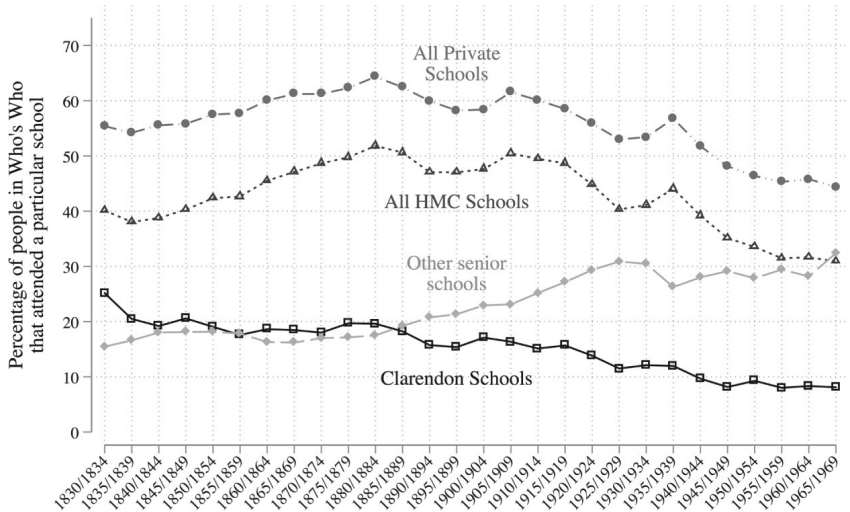
What's Next? Inequalities across the Companies House Register!

What's Next? Population Level Vehicle Data!

- We're currently working with vehicle administration data.
- The dataset contains **hundreds of millions** of observations.
- Information on emissions, characteristics, and registration.
- We are set to answer a range of questions, such as:
 1. What are the effects of unregulated exports?
 2. What are the inter-sectional characteristics of polluting?
 3. Can this data be used for estimating internal migration?

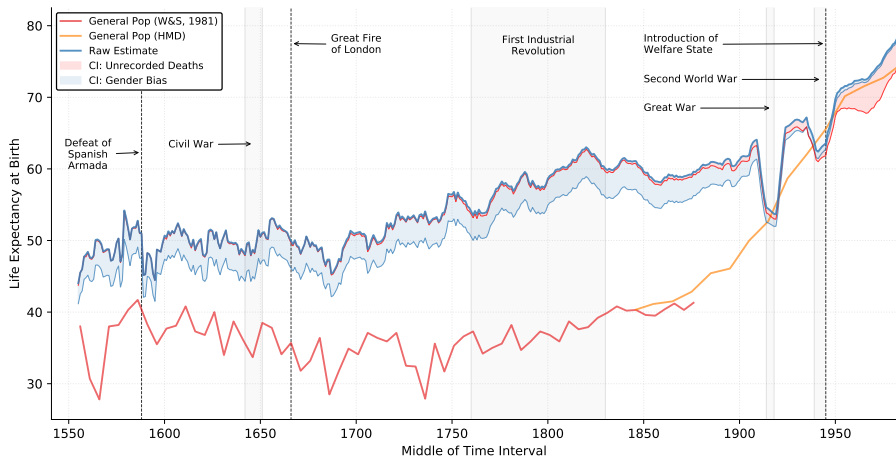
Semi-Structured Data for Inequality Research

- Data can come from a multitude of sources.
 - Sometimes, it can come from raw `.html` itself.
- One example of this is when we scraped ‘Who’s Who’ to get information on education background of professional elites.
- Another is when we scraped `thepeerage.com` to get data on dates of birth and death of British Peers.
- Exciting opportunities for digital archival research with computational methods.
- Let’s see some examples, and again consider ‘What’s Next?’



‘The decline and persistence of the old boy’

With Aaron Reeves, Sam Friedman. and Magne Flemmen



‘The Legacy of Longevity: Inequalities in UK Life Expectancy 1500-2016’ with Aaron Reeves, Felix Tropsch and Darryl Lundy

What's Next? Data on Wealth at Death

65	ABBOTT Henry Bernard of 22 India Road Strand Green London died 18 April 1908 Administration London 18 April 1908 Probate London 18 April 1908 Effects 4200 120.	ABBOTT Sarah Catherine of 24 Wykewood-road Seven Kings (widow of Benjamin Abbott) died 18 March 1907 Probate London 19 February to the 25 April 1907 Effects 4070 120.
	ABBOTT Isaac of Thompson Northamptonshire bornabout 18 October 1908 Administration Northampton 18 October 1908 Probate London 18 October 1908 Effects 4200 120.	ABBOTT Susan Georgina of Borough House (georgina) died 18 October 1908 Administration London 19 February to the 25 April 1907 Probate London 19 February to the 25 April 1907 Effects 4070 120.
	ABBOTT James of The Elms Road Green-end Banastre Somerset vicar died 10 February 1908 Probate London 11 April to Elizabeth Abbott widow. Effects 4340 120 20.	ABBOTT Thomas of Overton professor died 17 November 1907 Probate London 17 November 1907 Effects 4200 120.
	ABBOTT James of 41 Lane-end Leytonstone Essex died 20 January 1908 Administration London 21 March to John Benjamin Abbott contractor. Effects 4440 120 20.	ABBOTT Thomas of 101 Abchurch-lane Rochester Kent died 10 March 1908 Probate London 21 July to Thomas Edward Abbott and William Sackbush. Effects 4200 120.
	ABBOTT John of Bedfordshire died 10 February 1908 Administration London 21 February to Thomas Abbott farmer. Effects 4200 120.	ABBOTT Thomas of 32 Marsh-lane Middlesex died 8 June 1908 Administration London 9 July to Elizabeth Abbott widow. Effects 4200 120.
12	ABBOTT John of Mayfield House Northfield Herefordshire died 26 June 1908 Probate London 11 September to Mary Abbott widow and William Abbott railway clerk. Effects 4100 120 20.	ABBOTT Walter of Farnham near Witley Sussex died 24 November 1907 Probate London 25 July to Walter Perc Abbott and Osborne Hedley Abbott guardians. Effects 4500 120 20.
86	ABBOTT John White of 2 Westbury-gardens Hagger Lane Kensington died 21 April 1908 Probate London 21 June to Gertrude Agnes Abbott widow. Effects 4400 120 20.	ABBOTT William of Richmond-road Bexleyheath Kent (shopkeeper collector and manager) died 1 November 1907 Probate London 11 February to Elizabeth Abbott widow. Effects 4100 120 20.
	ABBOTT Joseph of 4 Progression-cottage, Gough Field Gloucestershire Lancashire engine driver died 7 November 1908 Administration London 11 December to Martha Jane Abbott widow. Effects 4200 120 20.	ABBOTT William of Haverstock-hill London died 21 January 1908 Probate London 21 March to John Elgar Chatterton clerk to the Generalist. Effects 4100 120.
3	ABBOTT Lucia Augusta of Buxton Newmarket (widow) died 28 June 1907 Administration London 28 June 1907 Probate London 28 June 1907 Effects 4200 120 20.	ABBOTT William of 28 Pool-street Putney London born 1851 died 14 October 1908 Administration London 25 November to Sarah Ann Love (widow of Edwin Love). Effects 420 30.
165	ABBOTT Mary of Cottingham London died 10 October 1908 Probate London 14 December to Charles Albert Abbott and George Abbott and George Abbott book manager. Effects 4100 120 20.	ABBOTT William of 13 Lombard-street London died 21 April 1908 Administration London 21 April 1908 Probate London 21 April 1908 Effects 420 120 20.
	ABBOTT Martha of Haverstock-hill London died 10 October 1908 Probate London 14 December to Charles Albert Abbott and George Abbott and George Abbott book manager. Effects 4100 120 20.	ABBOTT Henry of The Church Cottage near Woodhouse Devonshire died 20 May 1908 Probate London 21 June 1908 Effects 4200 120 20.
	ABBOTT Percy William of 31 Ashcroft Park Villa Bournemouth died 10 May 1908 Administration London 11 September to Arthur Abbott chemist manager. Effects 4200 120 20.	ABBOTT Robert of Rye Sussex retired farmer died 10 October 1908 Probate London 11 September to Robert James Abbott and William Abbott. Effects 4200 120.
167	ABBOTT Samuel Mitchell of Elmfield Farnham died 10 August 1908 Administration London 11 September to Elizabeth Abbott widow. Effects 4200 120 20.	ABBOTT Daniel Maria Thelma of 28 Elm of Malvern England died 10 September 1908 Administration London 11 September to George Herbert Cross commercial clerk. Effects 4200 120.
52	ABBOTT Sarah of Dunster Somersetshire spinner died 20 October 1908 Administration London 21 April to Emily Gander widow. Effects 420 30 120.	ABBOTT Anna of Upwater Sussex died 10 September 1908 Probate London 11 September to George Herbert Cross commercial clerk. Effects 4200 120.
65	ABBOTT Sarah of 41 Sandringham Road London died 18 April 1908 Probate London 19 May to Eliza Taylor (widow of Albert Taylor). Effects 4120 120.	ABBOTT Sarah of 13 Darnley-road London born 1851 died 10 September 1908 Probate London 11 September to George Herbert Cross commercial clerk. Effects 4200 120.
105	ABBOTT Sarah of The Square Bedford Northamptonshire died 10 July 1908 Probate London 11 September to Thomas Abbott engineer. Effects 4200 120 20.	

Ongoing/prospective research (with A. Reeves and N. Muggleton) plans to use probate records from 1858-1996. The idea is to:

1. Scrape 800k .pdfs containing 24m probates.
2. Pre-process images ready for OCR.
3. Run OCR on HPC in parallel.
4. Build lifetables to calculate e_0 by wealth at death.

What's Next? Estimating Inter-sectional Costs of Care

- Provision of informal care becoming an **international crisis**.
- In addition to differential demand, costs are different, too.
- However, estimating the costs of providing care is difficult.
- This motivates improvements to Synthetic Control estimators.
 - We aim to incorporate ML into the 'individual' level.
 - This will allow us to estimate different 'costs'.
 - To be used on detailed, granular data (census, biobank, etc).
 - Key idea: create individual synthetic versions (non-carers).
 - Costs can be operationalised in many ways.

Uncertainty in the Scientific Record

- There are close links between ‘Open Science’ and traditional modeling paradigms!

Uncertainty in the Scientific Record

- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening up** the black box of research, we can examine the certainty of statistical results.

Uncertainty in the Scientific Record

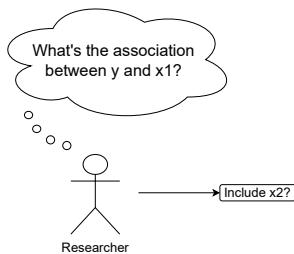
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening up** the black box of research, we can examine the certainty of statistical results.
- Does β_1 *really* equal 7.12452?

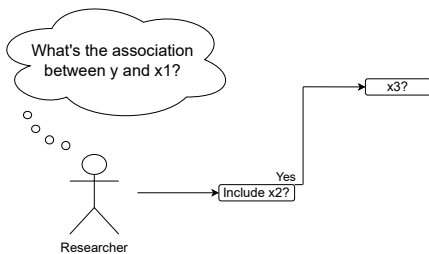
Uncertainty in the Scientific Record

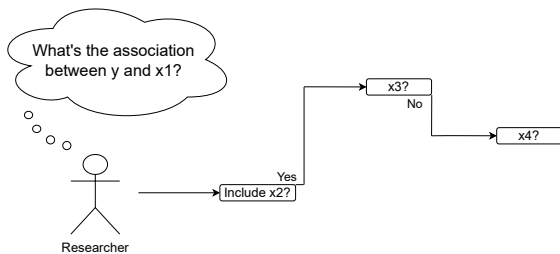
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening up** the black box of research, we can examine the certainty of statistical results.
- Does β_1 *really* equal 7.12452?
- How can we "take the ‘con’ out of econometrics", as Edward Leamer famously asked in 1983?

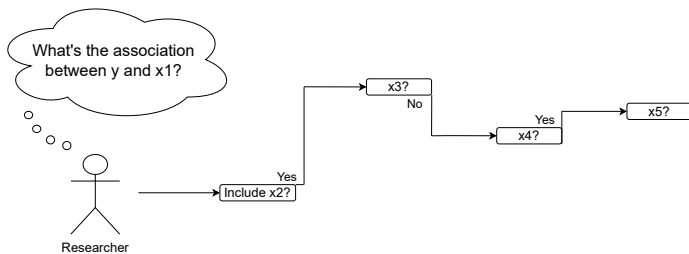
Uncertainty in the Scientific Record

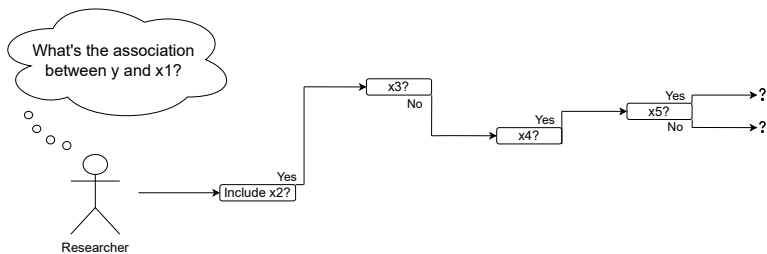
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening up** the black box of research, we can examine the certainty of statistical results.
- Does β_1 *really* equal 7.12452?
- How can we "take the ‘con’ out of econometrics", as Edward Leamer famously asked in 1983?
- **Transparency and more computational procedures!**

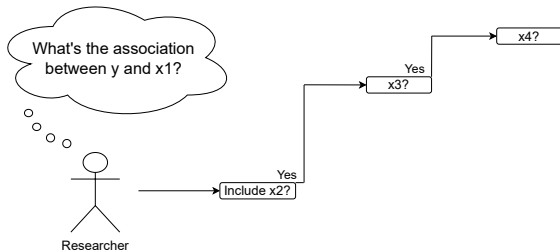


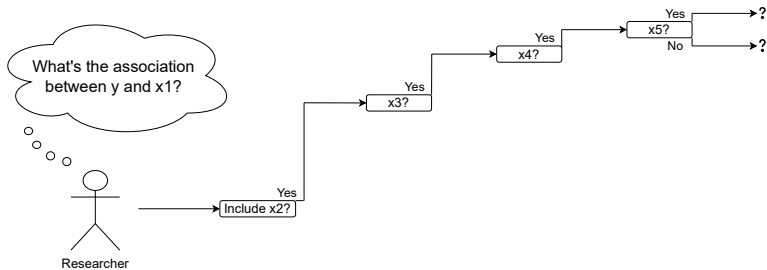


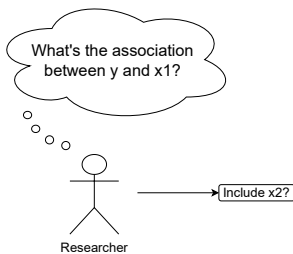


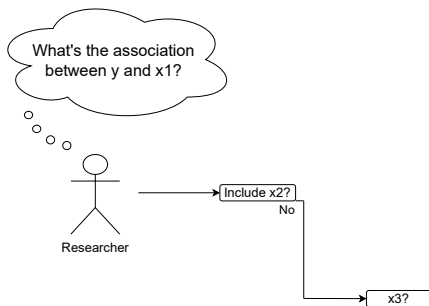


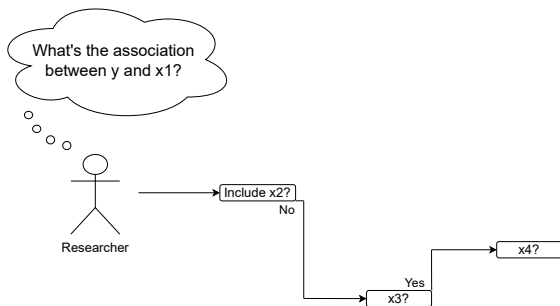


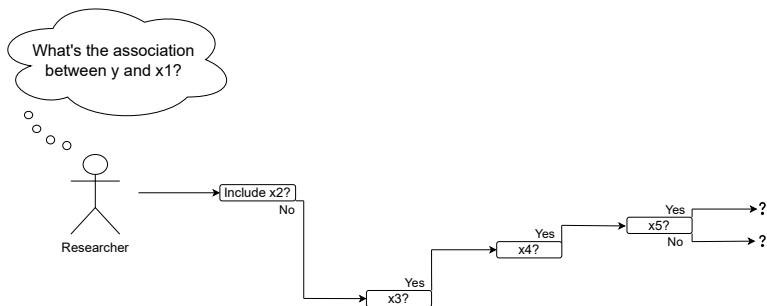


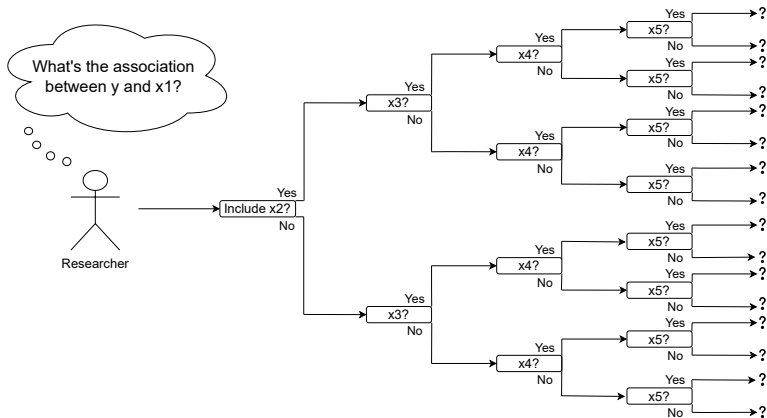


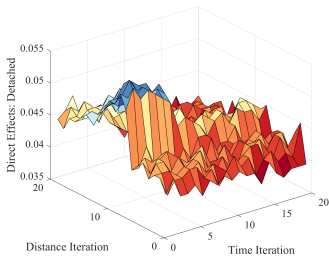
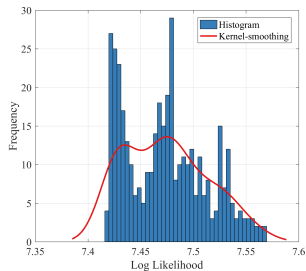
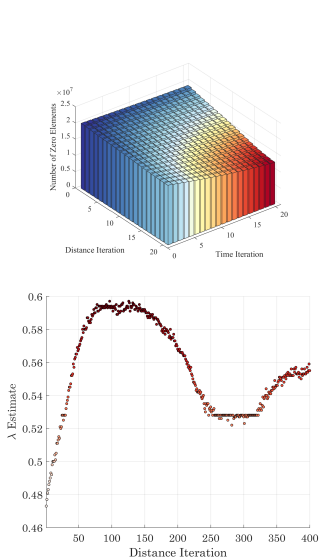












'A Grid Based Approach to Analysing Spatial Weighting Matrix Specification'

What's Next? Advancing the Spec-Curve Literature!

- Current research culture around robustness only mandates examination of a couple of choices.
 - With current computational power; why not all combinations?
 - Is specification choice the only source of variability?
- How can we visualize the *feasible* output space?
- This represents uncertainty around traditional point estimates.
- **NRobust**: The next generation of Specification Curves!
 - Bootstrap Resampling
 - Model Averaging
 - Model Selection
 - Cross-Validation

Pseudo-Random Number Generation

An invisible source of uncertainty in the scientific record: PRNGs!

Pseudo-Random Number Generation

An invisible source of uncertainty in the scientific record: PRNGs!

- **Q:** Has anyone ran a program twice, with different results?

Pseudo-Random Number Generation

An invisible source of uncertainty in the scientific record: PRNGs!

- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a 'seed'? Which seed?

Pseudo-Random Number Generation

An invisible source of uncertainty in the scientific record: PRNGs!

- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a 'seed'? Which seed?
- Maybe to eliminate variation in algorithms with PRNGs?
 - This seems to be the current 'best practice' advice.

Pseudo-Random Number Generation

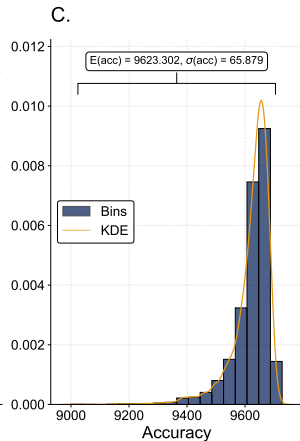
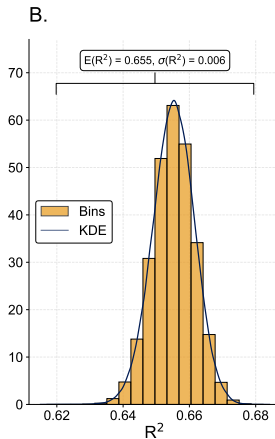
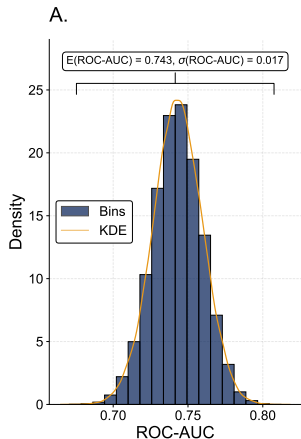
An invisible source of uncertainty in the scientific record: PRNGs!

- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a 'seed'? Which seed?
- Maybe to eliminate variation in algorithms with PRNGs?
 - This seems to be the current 'best practice' advice.
- We argue this is the **opposite of what we want!**
 - We propose you pre-specify **multiple** (complex) seeds.

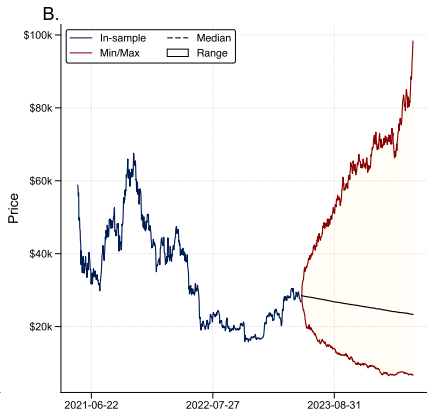
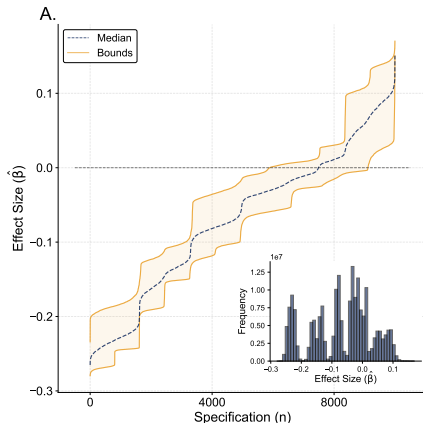
Pseudo-Random Number Generation

An invisible source of uncertainty in the scientific record: PRNGs!

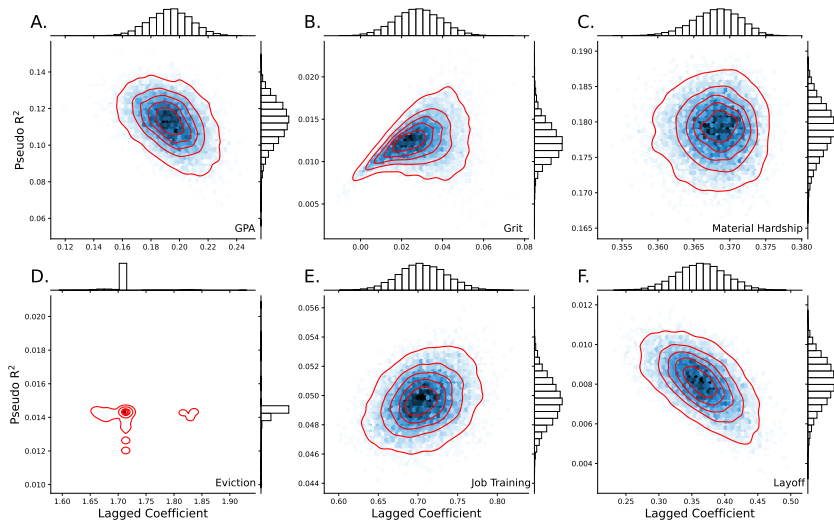
- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a 'seed'? Which seed?
- Maybe to eliminate variation in algorithms with PRNGs?
 - This seems to be the current 'best practice' advice.
- We argue this is the **opposite of what we want!**
 - We propose you pre-specify **multiple** (complex) seeds.
- We bring attention to this through multiple types of replications.
 - Simulations, machine learning, and inferential research.



‘From a Seed of Doubt Grows a Forest of Uncertainty’
With Arun Frey, Jiani Yan and Mark Verhagen



‘From a Seed of Doubt Grows a Forest of Uncertainty’
With Arun Frey, Jiani Yan and Mark Verhagen



‘From a Seed of Doubt Grows a Forest of Uncertainty’
With Arun Frey, Jiani Yan and Mark Verhagen

What's Next? Indexing Seed Variability

- The next thing to do is to formalize the replication project.
 - Does anyone have ideas for other types of seed variability?
- We are seeking to make a project website, two functions:
 1. People pre-register a project, receive a (set of) seeds.
 2. Index examples of seed variability in existent papers.
- The first improves the future of the scientific record.
- The second makes the historic scientific record more robust.

‘The future of mankind surely relies on reformulating existing challenges as predictive problems’

Avi Goldfarb at Said Business School, 2017

‘The future of mankind surely relies on reformulating existing challenges as predictive problems’

Avi Goldfarb at Said Business School, 2017

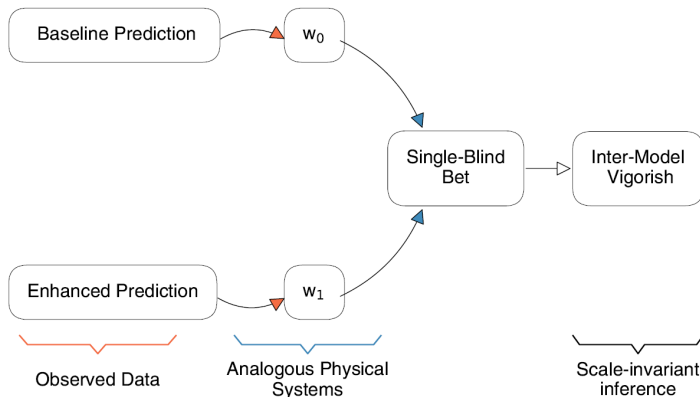
- In addition to being interested in STS...

‘The future of mankind surely relies on reformulating existing challenges as predictive problems’

Avi Goldfarb at Said Business School, 2017

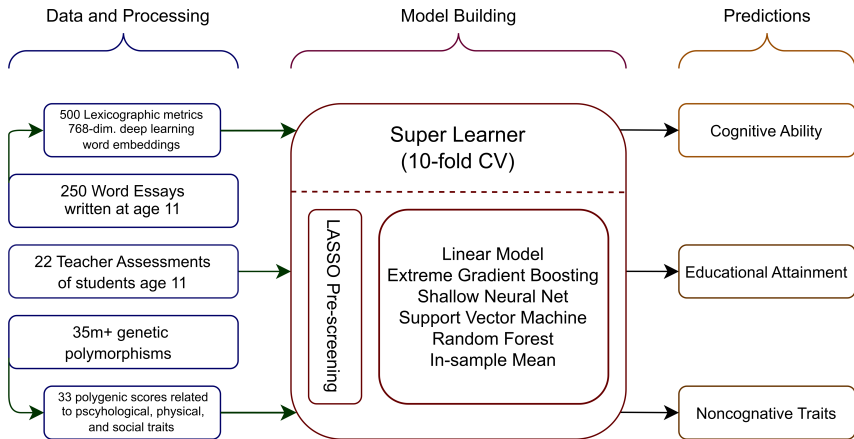
- In addition to being interested in STS...
- I am an active Machine/Deep Learning researcher!
 - I have a longstanding interest in forecasting and prediction.
- My work in this area is currently split into two aspects:
 1. The development of new methods and metrics.
 2. The application of frontier-level techniques to social problems.

Introducing the InterModel Vigorish



‘The InterModel Vigorish (IMV): A novel approach for quantifying predictive accuracy when outcomes are binary’

With Ben Domingue, Alex Rigos, Jeremy Frees *et al.*



‘Short Essays Written During Childhood Predict Cognition and Educational Attainment’

with Felix Tropsch and Tobias Wolfram

What's Next? Deep Learning for Social Care

One particular interest: predicting urgent care need in real-time.

- This is made possible by 'Technology Enabled Care'.
- We are in the process of mapping TEC ecosystems.
- The core idea: Use sensor data to predict emergencies.
- We are currently scoping, and undertaking ethical reviews.
- Data available through bespoke agreements with councils.
- We are planning to develop deep survival based models.

What's Next? Indexing Social Prediction

- Society is increasingly interested in 'how predictable' things are.
- This generates multiple interesting research questions:
 1. What are the **predictive limits** of health/social systems?
 2. With more HPCs/data, how much **irreducible error** remains?
 3. How can we incorporate a large number of **small signals**?
 - Are '**SWAS**' **ethically** or practically viable?
 4. How can we ensure social prediction is undertaken **ethically**?

Conclusion: **Thank you for attending!**

I hope I've been able confer my enthusiasm for:

- Robust and 'assumption free' social science.
- 'Semi-structured' data which is found 'in the wild'.
- Open, Responsible, and Reproducible Research.

Please don't hesitate to get in touch at:

charles.rahal@demography.ox.ac.uk

and/or

github.com/crahal!