

Adventures in Open and Assumption-Free Research

Charlie Rahal

LCDS and Nuffield College, University of Oxford

A talk at the Oxford Institute for Population Ageing, 27th February, 2024



In the spirit of ‘Open Science’:

crahal.github.io/files/oipa_talk.pdf

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

github.com/crahal

Lots of examples: hopefully something for everyone!

Everything here is made with free, and ‘Open’ tools:

- Linux (operating system)
- Python (data wrangling and statistical software)
- $\text{\LaTeX}$ (typesetting)

Everything here is made with free, and ‘Open’ tools:

- Linux (operating system)
- Python (data wrangling and statistical software)
- $\text{\LaTeX}$ (typesetting)

Why mention this? We’re talking about ‘Open’ Science!

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

What is the ‘Scientific Record’?

‘The scientific record is crucial for the advancement of knowledge and serves as a foundation for scientific progress. It provides a means for scientists to communicate their research findings to the broader scientific community, allowing for scrutiny, validation, and further exploration.’

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.

*"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 anecdotes from my time as a doctoral student:

What motivates a talk about ‘Open Science’?

- Four anecdotes from my time as a doctoral student:
 1. Make sure your coauthors have code.

What motivates a talk about ‘Open Science’?

- Four anecdotes from my time as a doctoral student:
 1. Make sure your coauthors have code.
 2. Be prepared to dismantle bad tests of theory.

What motivates a talk about ‘Open Science’?

- Four anecdotes from my time as a doctoral student:
 1. Make sure your coauthors have code.
 2. Be prepared to dismantle bad tests of theory.
 3. Critically analyze everyone else's data.

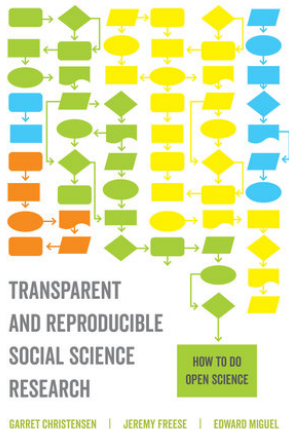
What motivates a talk about ‘Open Science’?

- Four anecdotes from my time as a doctoral student:
 1. Make sure your coauthors have code.
 2. Be prepared to dismantle bad tests of theory.
 3. Critically analyze everyone else’s data.
 4. Consider using open-source software.

What motivates a talk about ‘Open Science’?

- Four anecdotes from my time as a doctoral student:
 1. Make sure your coauthors have code.
 2. Be prepared to dismantle bad tests of theory.
 3. Critically analyze everyone else’s data.
 4. Consider using open-source software.
- Disclaimer: presently the LNL Co-ordinator for UKRN.

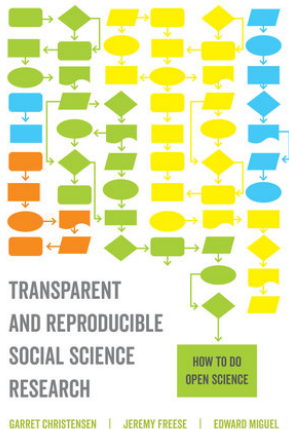
More traditional motivation, just in case you aren't sold...:



This book...

Is most definitely worth
reading!*

More traditional motivation, just in case you aren't sold...:



This book...

Is most definitely worth
reading!*

*Not just because one of
the authors is a coauthor of
a paper included later!

It does an **excellent** job of
motivating replications.

One of the most important parts of the book is this table:

Compared to Original	Original Procedures	Key Differences
Same Data	Verification	Reanalysis
Different Data	Direct Replication	Extension

- Maybe you've heard 'Verification' called 'Exact Replication'.
- Increasing push towards teaching concepts of 'Replication'.
- One area where Sociology has really progressed forwards.

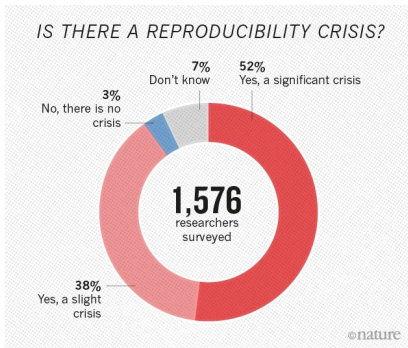
How it started:

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:



What does ‘Open Science’ look like?

1. Preprints and ‘Open Access’.
2. Publicly available data.
3. Open source code



What does ‘Open Science’ look like?

1. Preprints and ‘Open Access’.
2. Publicly available data.
3. Open source code

 **HARVARD**
Dataverse

[Add Data](#) ▾ [Search](#) ▾ [About](#) [User Guide](#) [Support](#) [Sign Up](#) [Log In](#)

Deposit and share your data. Get academic credit.

Harvard Dataverse is a repository for research data. Deposit data and code here.

[Add a dataset](#) +

Organize datasets and gather metrics in your own repository.

A dataverse is a container for all your datasets, files, and metadata.

[Add a dataverse](#) +

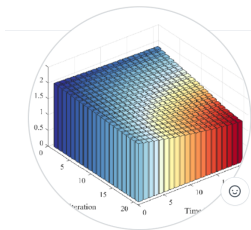
Publishing your data is easy on Harvard Dataverse!

Learn about getting started creating your own dataverse repository here.

[Getting started](#) ↗

What does 'Open Science' look like in practice?

1. Preprints and 'Open Access'.
2. Publicly available data.
3. Open source code



Pinned

Customize your pins



[GWASReview](#)

Public ::

A Scientometric Review of Genome Wide Association Studies

● HTML ☆ 15 📄 7



[NHSSpend](#)

Public ::

A home for the NHSSpend Library

● HTML ☆ 1 📄 1



[InterModel_Vigorish](#)

Public ::

Introducing the InterModel Vigorish (IMV)

● Jupyter Notebook ☆ 1



[OxfordDemSci/gwasdiversitymonitor](#)

Public ::

Source code for the GWAS Diversity Monitor

● CSS ☆ 12 📄 1

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}
```

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!

Why show you this? Well, Python is great at lots of useful things:

1. Scientific Computing (SciPy, numpy)
2. Data visualisation (Matplotlib, seaborn)
3. Web scraping (requests)
4. Data wrangling (Pandas)

A natural choice for un/semi-structured data. Python also excels at NLP/ML, but maybe a little less with classical statistics.

Python is **great** at building new datasets. But why bother?

'We have been using the same data sets over and over again, and since we wanted new answers, we have been developing new econometric techniques to try to transform the data, and get more meaningful information out of them. Just having a fresh, new data set brings a whole new perspective, and I think people are starting to realize that, and gradually people are becoming more interested in data collection itself'

A. Cavallo, 15 November 2012 in Taylor et al. (2014)

Python is **great** at building new datasets. But why bother?

'We have been using the same data sets over and over again, and since we wanted new answers, we have been developing new econometric techniques to try to transform the data, and get more meaningful information out of them. Just having a fresh, new data set brings a whole new perspective, and I think people are starting to realize that, and gradually people are becoming more interested in data collection itself'

A. Cavallo, 15 November 2012 in Taylor et al. (2014)

New datasets allow fresh **insight** and are typically **open**!

But what is **semi-structured** data?

“A form of structured data that does not obey the tabular structure of data models associated with relational databases or other forms of data tables, but nonetheless contains tags or other markers to separate semantic elements”

Different to **tabular** surveys classical social scientists are used to.

But, often, our goal is to parse it and make it tabular all the same.

A lot of my semi-structured work is ‘**scientometric**’ and ‘**bibliometric**’ in nature:

- **Scientometrics** is the field of study which concerns itself with measuring and analysing science.
- **Bibliometrics** is the use of statistical methods to analyze quantitative information pertaining to books, articles and other publications.

These fields are naturally related to ‘**Open Science**’.

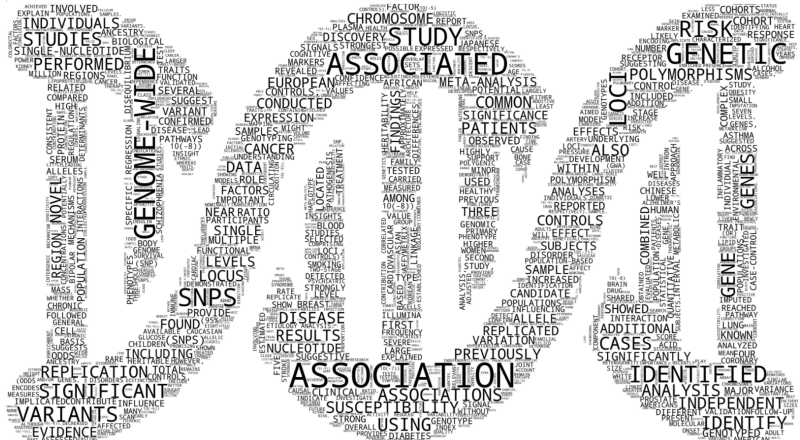
Data comes from various providers (usually APIs).

Lets use a 'Traffic Light' system to rank by openness:

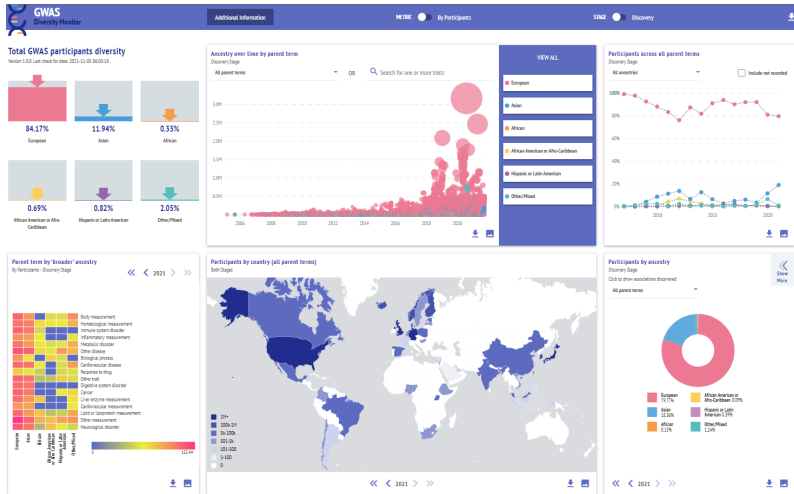
- Web of Science
- Google Scholar
- Scopus
- PubMed
- JSTOR/Constellate
- CrossRef
- The General Index
- Dimensions

This list gets longer every day...

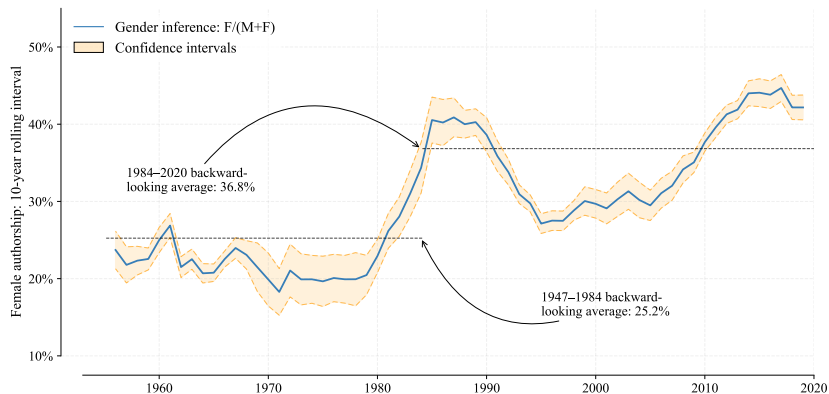
- My early scientometric work analyzed GWAS.
- It took curated lists of catalogued information from the European Bioinformatics Institute. It then called **semi-structured** data from PubMed.
- We published two papers which:
 1. Conduct a thorough scientometric review.
 2. Launch a substantial interactive dashboard.



'The Scientometrics of Genome Wide Association Studies'
With Melinda C. Mills

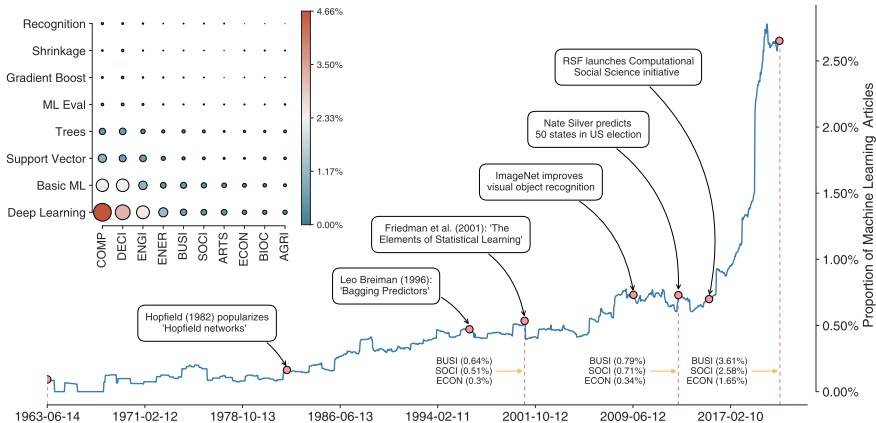


‘The GWAS Diversity Monitor Tracks Diversity in Real Time’
with Melinda C. Mills



‘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

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.
 - Increasingly recognised by the ESRC (e.g. ADRNs).
- For better or worse:
 - I’m more interested in ‘hard to use’, semi-structured data.

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.
 - Increasingly recognised by the ESRC (e.g. ADRNs).
- For better or worse:
 - I’m more interested in ‘hard to use’, semi-structured data.
 - Using curated data from the ONS SRS would be too easy.

TRANSPARENCY INTERNATIONAL UK
Open Access Home Other versions **DONATE** About

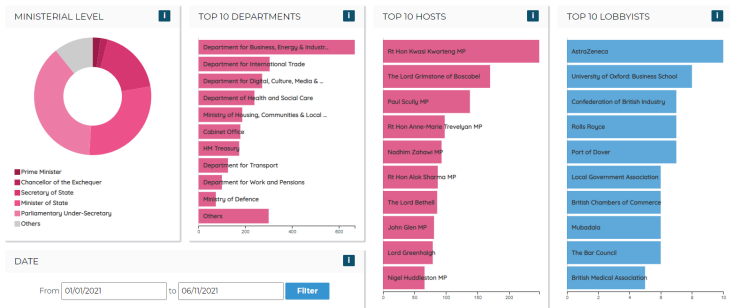
2,623 out of 74,435 meetings

SEARCH BY LOBBYIST, HOST, SUBJECT

Reset

UK GOVERNMENT

LOBBYIST



openaccess.transparency.org.uk

With Ian Knowles, Paulo Serôdio, David Stuckler and TI UK

Procurement Data

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

Procurement Data

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

If only there was a **Government Data Service** to curate it...

Procurement Data

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

If only there was a **Government Data Service** to curate it...

However, I am currently working with the ONS to improve collection!

Lets list a few of the problems with procurement data. Some are common to semi-structured data, some are not:

1. Some organisations 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. Some are common to semi-structured data, some are not:

1. Some organisations 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 UUIDs.

Just one quick example, followed by what we can do...

A3131

Expenditure Over Threshold

Department/Agency	City	Year	Category	Subcategory	Transaction number	AP Amount	VAT registration number	Purchase Invoice number
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599853	606 000.00	6100071863
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599852	100 000.00	6100071861
Department of health	Benby	CGG	30/04/2021	Other Gen Supplies & Inv	COMMUNITY SERVICES	11599851	15 669.99	6100071860
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	WYLLIE RESTAURANT	11599850	64 000.00	6100071859
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11600729	68 226.00	605483887
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING - NON ACUTE	11600730	68 226.00	6100071858
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	ACUTE COMMISSIONING	11615024	30 000.00	6100071857
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Independent Sector	ACUTE COMMISSIONING	11615025	38 526.24	787424855
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11615026	58 137.51	HL020056882
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	HOSPICES	11626337	218 000.00	871
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	LEWISHAM & GREENWICH NHS TRUST	11634201	440 444.00	8442627
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11634202	218 162.11	14620506
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642864	25 672.53	48413715
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642865	-30 000.00	2000004184
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642866	200 000.00	2000004185
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11654422	-3 662 119.34	CM00218970
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	AMBUANCE SERVICES	11665375	587 727.29	605481192
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11665983	786 836.00	227 489
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11665985	760 530.34	233561
Department of health	Benby	CGG	30/04/2021	Cont Care Adult 100% Fully Funded	FUNDED NURSING CARE	11666922	26 512.00	2387311641
Department of health	Benby	CGG	30/04/2021	Charges From CSU	RECHARGES	11669995	85 249.00	GB054442045
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11671003	278 826.00	1206094
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11680115	82 568.00	CM00254903
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701702	228 869.00	278586
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701703	141 917.00	278587
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701710	807 726.00	278622
Department of health	Benby	CGG	30/04/2021	Income Mx < 11%	BALANCE SHEET	11720068	47 733.11	6044623331
Department of health	Benby	CGG	30/04/2021	Income Mx < 11% Student Loans	BALANCE SHEET	11720069	738.00	6044623331
Department of health	Benby	CGG	30/04/2021	National Insurance - 1 Yr NP-RMS	BALANCE SHEET	11720065	29 114.84	6044623331
Department of health	Benby	CGG	30/04/2021	National Insurance - 1 Yr NP-RMS	BALANCE SHEET	11720066	19 284.18	6044623331
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724281	213 245.00	2029126
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11724282	288 141.00	2295884
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724283	27 100.00	2295885
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11722209	266 424.00	231364
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724114	224 749.00	231365
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	LEWISHAM & GREENWICH NHS TRUST	11734201	440 444.00	8442627
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	11738262	-30 056.34	1365
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	11738269	55 840.00	CM0026848
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11748400	249 505.00	175451
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11748386	47 300.00	201000982
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11771723	10 100.00	263322
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11771723	2 691 512.00	100004308
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11777727	25 202.00	80002737
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES - NON ACUTE	11781016	64 381.00	CM00218980
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11782492	1 728 242.00	37897
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11801176	36 000.00	37898
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ESTATES AND FACILITIES	11801410	51 209.79	484138351
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11818020	258 781.00	37895
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11818021	85 506.00	37894
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	LEARNING DIFFICULTIES	11818022	4 416.81	6800019263
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	HOSPICES	11818190	276 741.00	589
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Independent Sector	WYLLIE RESTAURANT	11818191	64 000.00	6100071857
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11824800	1 869 511.00	231431
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	LEARNING DIFFICULTIES	11854467	250 000.00	6100071247
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11854468	250 000.00	6100071247
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES	11963191	65 592.39	8442627
Department of health	Benby	CGG	30/04/2021	Healthcare Srv Res CCG	NHS 111	13263323	71 747.88	BA-0320156
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Independent Sector	COMMUNITY SERVICES	13263324	26 700.00	1062021149
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11903373	1 137 848.99	234497
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11931401	1 455 963.00	604623417
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11931402	1 133 848.99	604623417
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11933581	-129 548.00	8442625
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951210	1 187 569.00	234376
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951211	1 187 569.00	234377
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951212	229 254.29	234340
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951213	229 254.29	234341
Department of health	Benby	CGG	30/04/2021	Cont Care Adult 100% Fully Funded	CHC ADT FULL FUND PERKS HLTH PRG	11961590	205 574.97	CM00218971
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11962539	68 226.00	605483887

A3131

Expenditure Over Threshold

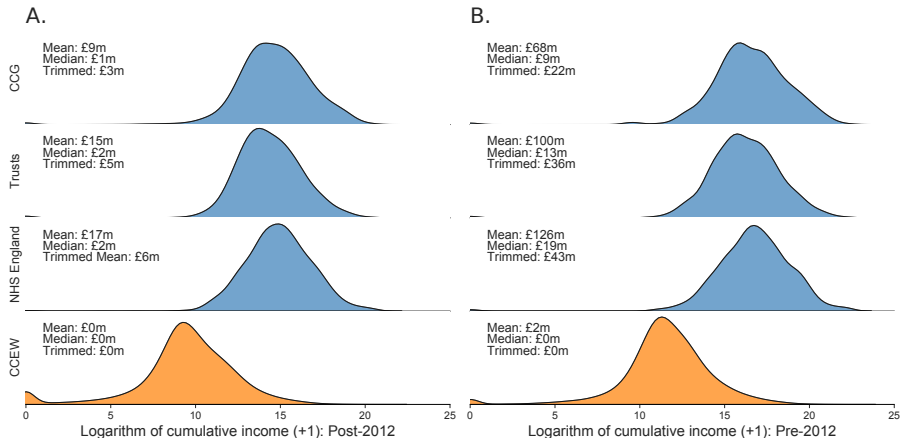
Department/Agency	City	Year	Category	Subcategory	Transaction number	AP Amount	VAT registration number	Invoice number
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599653	606 000.00	6100071861
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599652	1440 000.00	6100071861
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	COMMUNITY SERVICES	11599651	165699.00	6100071861
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	WYLLER RESTORATION	11599642	64 000.00	6100071861
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11600729	68 226.00	601008644
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11600728	165699.00	601008644
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	11615604	30 000.00	6100071861
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	ACUTE COMMISSIONING	11615603	38 526.24	787424655
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	ACUTE COMMISSIONING	11615602	58 147.4	787424655
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	HOSPICES	11623337	218 000.00	871
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11634001	444 000.00	60402567
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11634002	218 146.1	60402567
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642864	25 672.53	48413715
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642863	-30 000.00	2000004184
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11642862	2000004184	2000004184
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11654423	-3 626 116.34	CM00218970
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	AMBUCLANCE SERVICES	11665375	587 757.29	6040191392
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11665983	768 636.00	227 489
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11665982	760 530.34	233561
Department of health	Benby	CGG	30/04/2021	Cont Care Adult 100% Fully Funded	FUNDED NURSING CARE	11666989	26 912.00	2387911641
Department of health	Benby	CGG	30/04/2021	Charges From CSU	RECHARGES	11669999	85 249.00	GB054442045
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11671003	278 826.00	1206094
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11680115	82 568.00	600129403
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701702	228 969.00	278586
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701701	141 917.00	278586
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11701710	807 726.00	278922
Department of health	Benby	CGG	30/04/2021	Income Mx < 11%	BALANCE SHEET	11720068	47 733.1	604462331
Department of health	Benby	CGG	30/04/2021	Income Mx < 11%-Student Loans	BALANCE SHEET	11720068	738.00	604462331
Department of health	Benby	CGG	30/04/2021	National Insurance < 1 vNE-RMS	BALANCE SHEET	11720068	29 114.84	604462331
Department of health	Benby	CGG	30/04/2021	National Insurance < 1 vNE-RMS	BALANCE SHEET	11720068	19 284.18	604462331
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724081	213 245.00	2020126
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11724082	288 141.00	2020126
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724083	27 126.00	2020126
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11722029	204 833.00	231364
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11724114	242 139.00	2020126
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	ACUTE COMMISSIONING	11734640	200 000.00	600129403
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	11738262	-20 056.34	1365
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Over Under Performance	ACUTE COMMISSIONING	11738263	30 540.00	600129403
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	COMMUNITY SERVICES	11748440	178 400.00	17541
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11748438	47 400.00	202010982
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11771723	262 132.00	202010982
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11771723	2 691 512.00	1000041308
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11777827	25 202.00	800027387
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	11777828	56 381.00	600129403
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11782402	1 728 240.00	37897
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	MENTAL HEALTH CONTRACTS	11801176	30 000.00	37898
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	CHILD AND ADOLESCENT MENTAL HEALTH	11801175	180 000.00	484138351
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11801820	258 781.00	37895
Department of health	Benby	CGG	30/04/2021	Hilltop Foundation Trusts	CHILD AND ADOLESCENT MENTAL HEALTH	11813400	58 506.00	37894
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	WYLLER SUPPORTIVE SERVICES	11813401	118 100.00	600129403
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	DAY SERVICES	11815190	276 741.00	589
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	DAY SERVICES	11815191	25 000.00	600129403
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11824000	1 860 531.00	604226016
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Clinic Public Sector	LEARNING DIFFICULTIES	11854467	250 000.00	6100071247
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11854468	117 126.00	6100071247
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	11903191	66 592.39	BA-0320156
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	NHS 111	13262323	71 147.88	7062021149
Department of health	Benby	CGG	30/04/2021	Clinical/Medical-Voluntary Sector	COMMUNITY SERVICES	11903192	51 209.19	600129403
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11903237	1 137 848.98	234497
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11931140	1 455 963.00	60402567
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11931141	1 133 848.98	60402567
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11933581	-129 548.00	8486605
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951210	1 187 569.00	232736
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951211	200 156.00	CM00218970
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951212	229 254.29	232542
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	PLANNED CARE	11951290	-411 000.00	CM00218970
Department of health	Benby	CGG	30/04/2021	Cont Care Adult 100% Fully Funded	CHC ADT FULL FUND PERKS HLTH PRG	11965329	269 532.97	6100071861
Department of health	Benby	CGG	30/04/2021	Health Srv Res P8th Trust-Contract Baseline	ACUTE COMMISSIONING	11965330	68 226.00	60402567

A3131

Expenditure Over Threshold

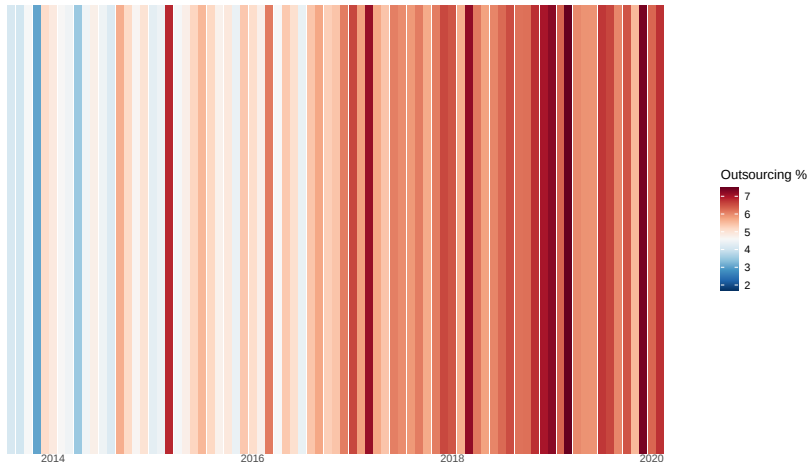
Department Name	Fund	Category	Date	Expense Type	Expense Area	Vendor	Transaction Number	AP Amount	YTD	YTD Description Number	Purchase Invoice Number
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Off Public Sector	COMMUNITY SERVICES	BEXLEY COUNCIL	11059633	658.00	000	6100071658
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Off Public Sector	COMMUNITY SERVICES	BEXLEY COUNCIL	11059632	10.00	000	6100071661
Department of Health	NHS	Beasey	COG	10/04/2021	Other Gen Supplies & Srv	COMMUNITY SERVICES	BEXLEY COUNCIL	11059637	97.00	000	6100071674
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Off Public Sector	WYTH RESEARCH	11059662	84.00	000	6100071687	
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	UNIVERSITY COLLEGE LONDON HOSPITALS NHS FOUNDATION TRUST	11060237	116.00	000	6100080444
Department of Health	NHS	Beasey	COG	10/04/2021	Cont Care - Adult 100% Fully Funded	CHIC ADULT FULLY FUNDED	CRAGWOOD KIDS GROUP PRIORITY	11060225	174.00	000	6100080300
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING - NON ACUTE	BEXLEY COUNCIL	11059636	18.00	000	6100080392
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Independent Sector	ACUTE COMMISSIONING	INPHATH LTD	11051870	35.00	200	6100080582
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	MEOWAY NHS FOUNDATION TRUST	11059774	35.00	000	6100082383
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Independent Sector	ACUTE COMMISSIONING	WYTH RESEARCH & BEXLEY COMMUNITY HOSPICE	11059774	35.00	000	6100082383
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	LEIGHAM & GREENWOOD NHS TRUST	11058401	4,440.14	000	6100082657
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PROGRAMME PROJECTS	NHS PROPERTY SERVICES LTD	11061697	216.72	140	6100082696
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PROGRAMME PROJECTS	NHS PROPERTY SERVICES LTD	11062084	181.17	000	6100082718
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	DARTFORD AND GRAVESEND NHS TRUST	11062423	-30.30	000	6100083044
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	DARTFORD AND GRAVESEND NHS TRUST	11062427	10,000.00	000	6100083049
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11062509	-3,000.00	000	6100083092
Department of Health	NHS	Beasey	COG	10/04/2021	Clinical/Medical-Commercial Sector	COMMUNITY SERVICES	BEXLEY HEALTH CARE	11065539	28.93	420	6100083196
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	AMBULANCE SERVICES	LONDON AMBULANCE SERVICE NHS TRUST	11063378	587.73	000	6100083192
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11065989	30.00	000	6100083449
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11065985	760.50	000	6100083541
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11065985	18.00	000	6100083541
Department of Health	NHS	Beasey	COG	10/04/2021	Changes from CGU	RECHARGES	NHS SOUTH EAST CGU	11065996	45.25	000	6100083541
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	PLANNED CARE	KING'S COLLEGE HOSPITAL NHS FOUNDATION TRUST	11067000	219.25	000	6100083541
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	ACUTE COMMISSIONING	MOORFLOOS EYE HOSPITAL NHS FOUNDATION TRUST	11068116	62.00	000	6100083593
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	COMMUNITY SERVICES	OCELEX NHS FOUNDATION TRUST	11070102	201.95	000	6100083596
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	COMMUNITY SERVICES	OCELEX NHS FOUNDATION TRUST	11070107	141.97	000	6100083596
Department of Health	NHS	Beasey	COG	10/04/2021	Health Srv Nhs Pbn Trust-Contract Baseline	COMMUNITY SERVICES	OCELEX NHS FOUNDATION TRUST	11070110	807.20	000	6100083596
Department of Health	NHS	Beasey	COG	10/04/2021	Income tax v17	BALANCE SHEET	HMRC	11072005	471.83	000	6100083596
Department of Health	NHS	Beasey	COG	10/04/2021	Income tax v17	BALANCE SHEET	HMRC	11072006	30.00	000	6100083596
Department of Health											





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

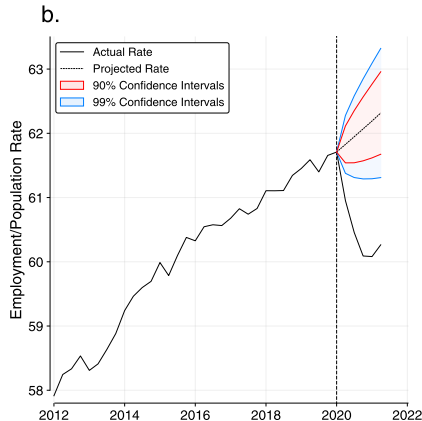
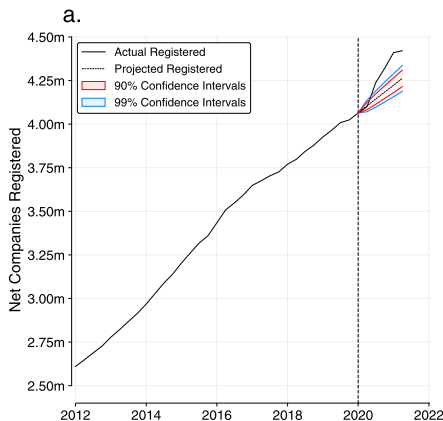
Creeping Privatisation in England's NHS



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

‘Creeping Commercialization in the NHS’

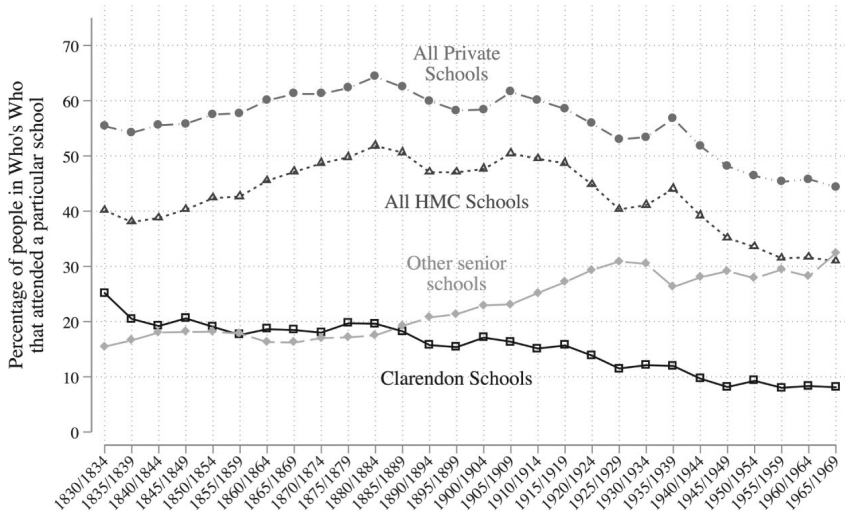
By Ben Goodair and Aaron Reeves



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

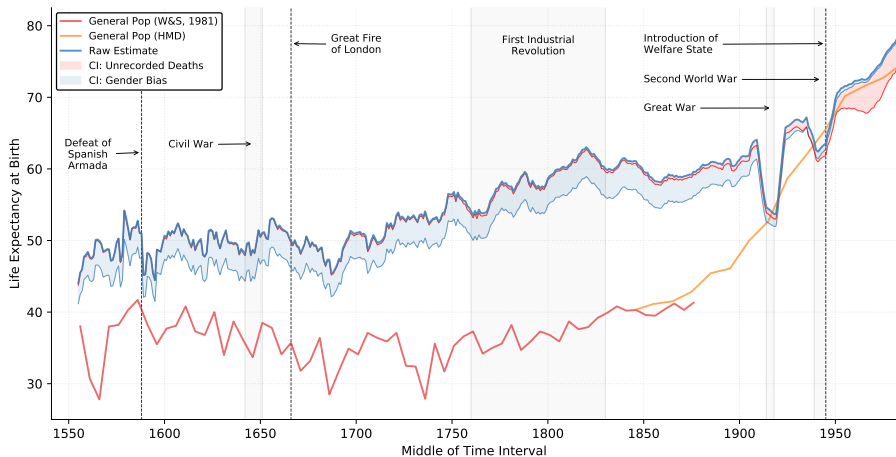
Semi-Structured Data for Inequality Research

- Data doesn't always come from APIs.
- 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 'elites'.
- Another is when we scraped `thepeerage.com` to get data on dates of birth and death of British Peers.
- We might want to (there are competing views):
 - Check ToS and `robots.txt`.
 - Contact the webmaster first.



‘The decline and persistence of the old boy’

With Aaron Reeves, Sam Friedman. and Magne Flemmen

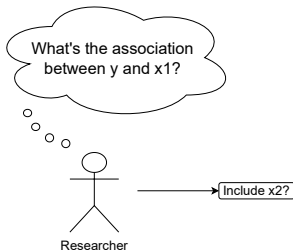


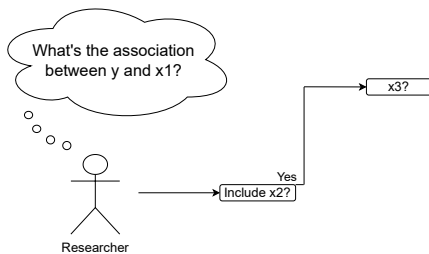
‘The Legacy of Longevity: Inequalities in UK Life Expectancy 1500-2016’
with Aaron Reeves, Felix Tropsch and Darryl Lundy

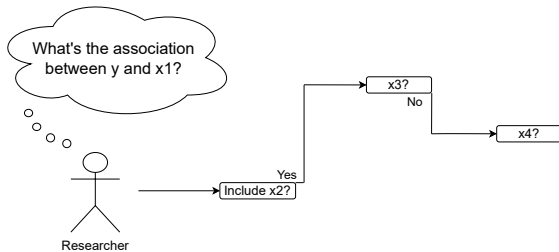
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening** research, we can examine the certainty of statistical results.
- Does β_1 really equal 7.12452?

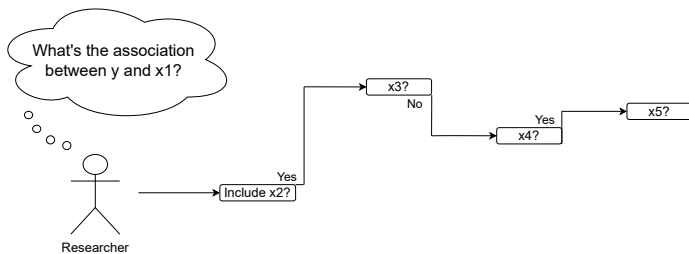
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening** research, we can examine the certainty of statistical results.
- Does β_1 really equal 7.12452?
- How can we "take the ‘con’ out of econometrics"?

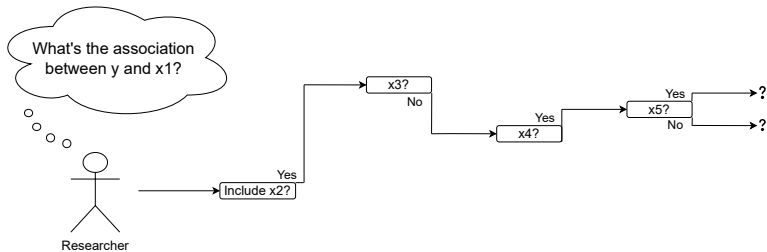
- There are close links between ‘Open Science’ and traditional modeling paradigms!
- By **opening** research, we can examine the certainty of statistical results.
- Does β_1 really equal 7.12452?
- How can we "take the ‘con’ out of econometrics"?
- Transparency and more computational procedures!

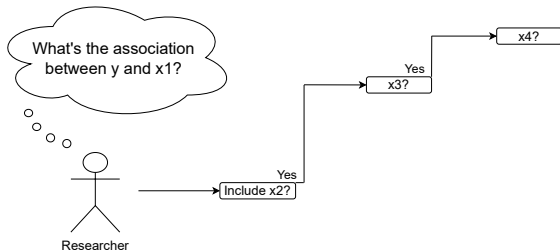


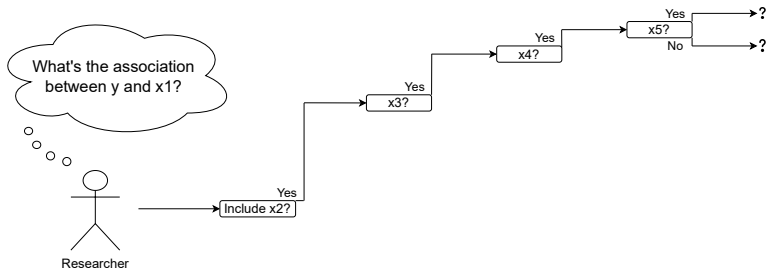


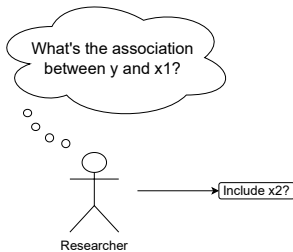


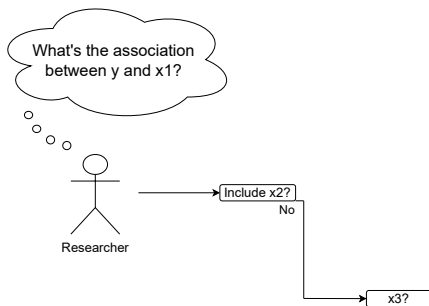


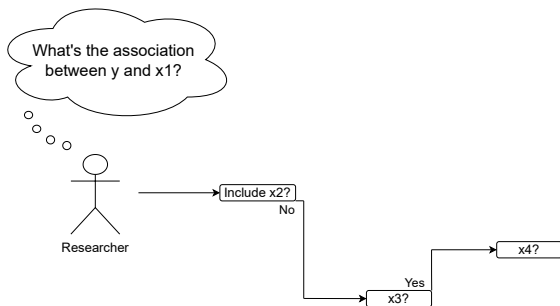


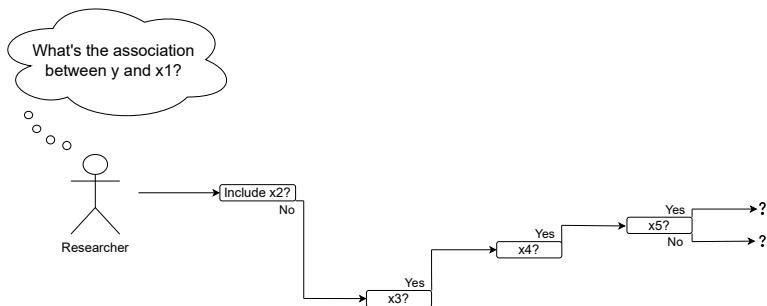


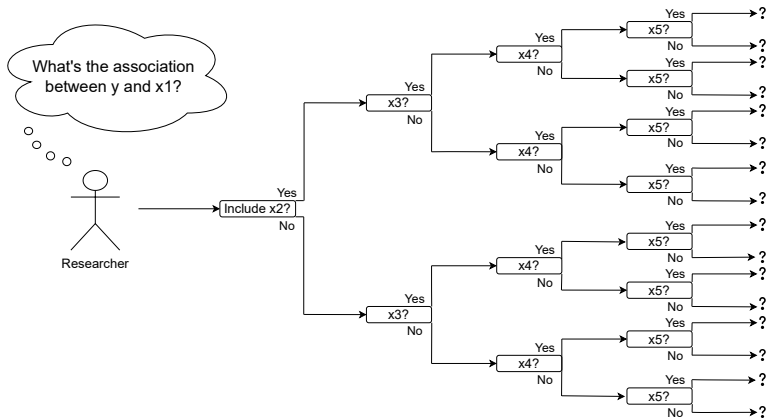




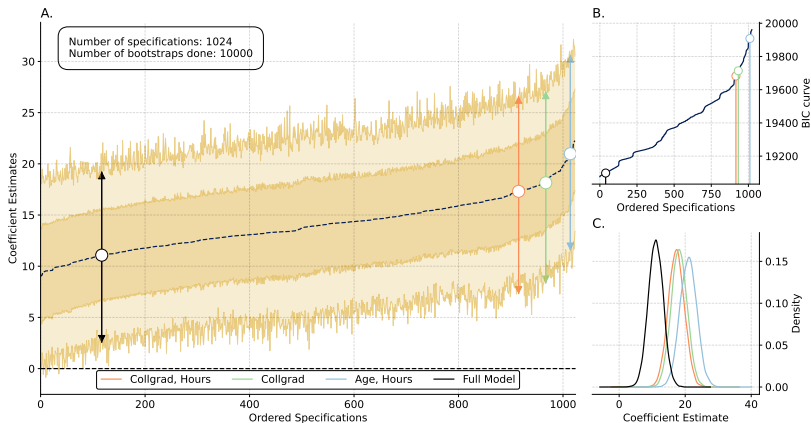




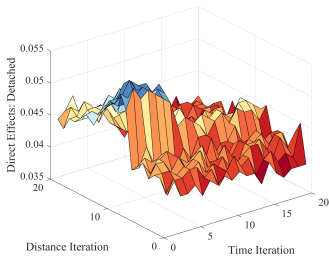
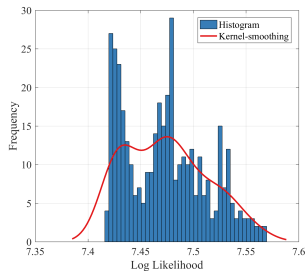
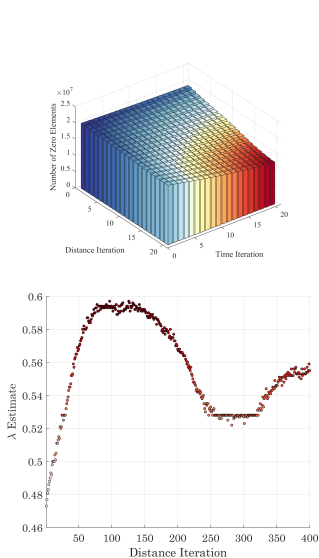




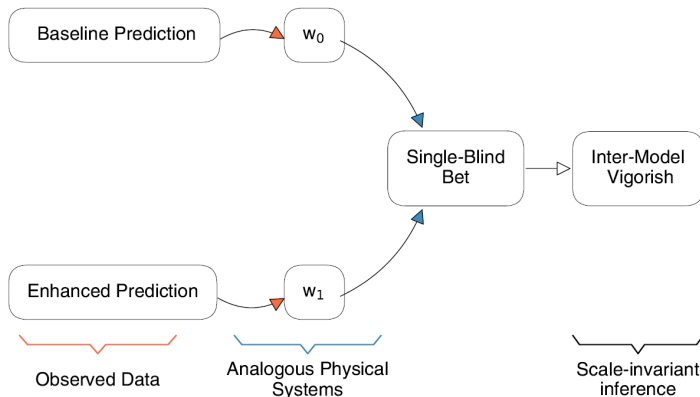
- Our work tackling this type of uncertainty is still in progress!



- The idea is to advance the **specification curve literature**.
- In particular: selection, **averaging**, and **influence analysis**.
- See github.com/centre-for-care/nrobust for updates!



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



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

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

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.
- **Q:** Has anyone ran a program twice, with different results?

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.
- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a '**seed**'?

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.
- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a '**seed**'?
- **Q:** How did you decide on what seed to use?

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.
- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a '**seed**'?
- **Q:** How did you decide on what seed to use?
- Maybe to eliminate variation in algorithms with PRNGs?
 - This seems to be the current 'best practice' advice.

Pseudo-Random Number Generation

- I'd like to end with a few slides on what I'm working on **right now**.
- **Q:** Has anyone ran a program twice, with different results?
- **Q:** Has anyone here ever set a '**seed**'?
- **Q:** How did you decide on what seed to use?
- 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 (Cont.)

- We want to consider possible variation independent of arbitrary variation in seed choice.

Pseudo-Random Number Generation (Cont.)

- We want to consider possible variation independent of arbitrary variation in seed choice.
- This is an extremely important and scarcely research problem.

Pseudo-Random Number Generation (Cont.)

- We want to consider possible variation independent of arbitrary variation in seed choice.
- This is an extremely important and scarcely research problem.
- Pseudo-Random Number Generators occur **everywhere!**

Pseudo-Random Number Generation (Cont.)

- We want to consider possible variation independent of arbitrary variation in seed choice.
- This is an extremely important and scarcely research problem.
- Pseudo-Random Number Generators occur **everywhere!**
- The variation in estimand can be **huge**.

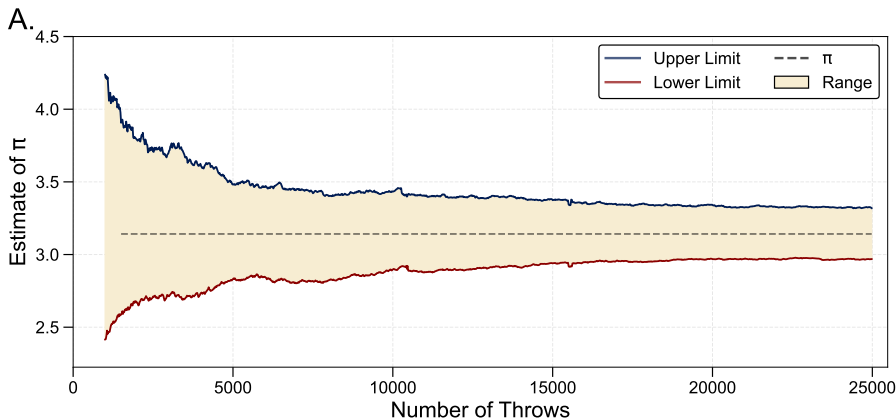
Pseudo-Random Number Generation (Cont.)

- We want to consider possible variation independent of arbitrary variation in seed choice.
- This is an extremely important and scarcely research problem.
- Pseudo-Random Number Generators occur **everywhere!**
- The variation in estimand can be **huge**.
- We bring attention to this through multiple types of replications.
 - Simulations, machine learning, topic modelling, and inferential research.

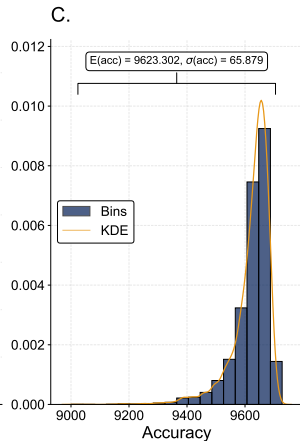
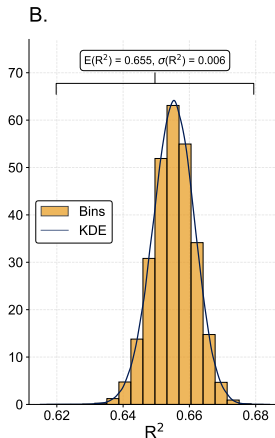
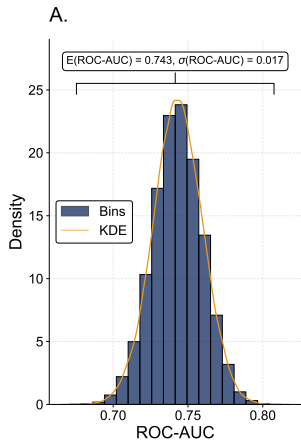
- **Problem Statement:** By setting one seed, we have no idea about the potential key variation of our estimand as a function of how random numbers were generated. This is computationally unintensive, but scientifically dour.

- **Problem Statement:** By setting one seed, we have no idea about the potential key variation of our estimand as a function of how random numbers were generated. This is computationally unintensive, but scientifically dour.
- **Solution:** Visualize the outcome space of a of large number (10k? 100k?) seeds simultaneously. This is computationally intensive, but scientifically illuminating.

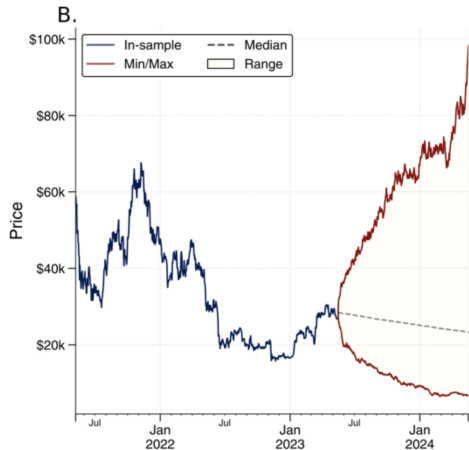
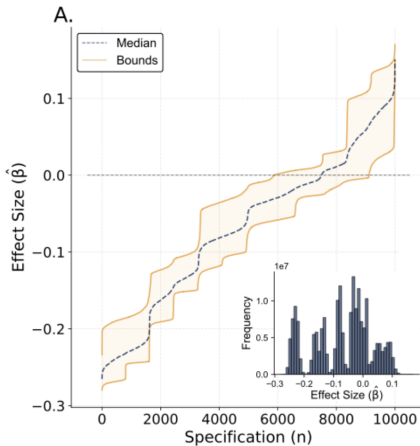
- **Problem Statement:** By setting one seed, we have no idea about the potential key variation of our estimand as a function of how random numbers were generated. This is computationally unintensive, but scientifically dour.
- **Solution:** Visualize the outcome space of a of large number (10k? 100k?) seeds simultaneously. This is computationally intensive, but scientifically illuminating.
- **Replication:** We can consider whether an original result is in the tail of our distribution (or IQR?) or not.



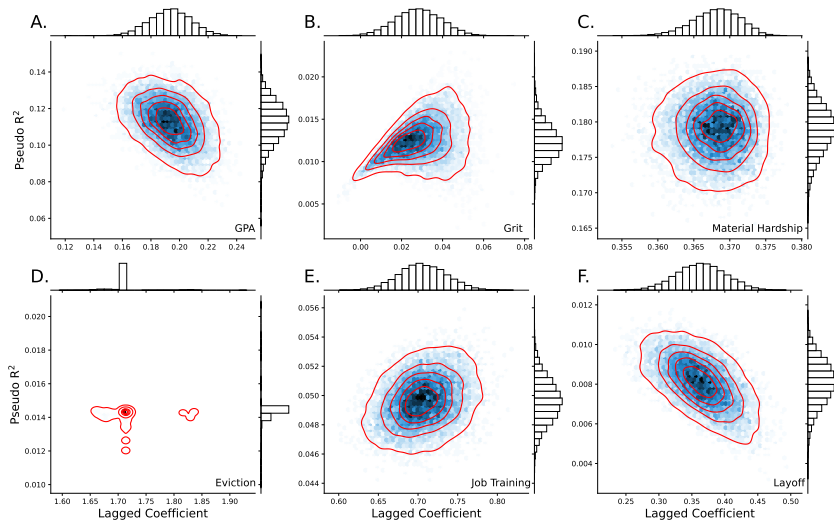
‘From a Seed of Doubt Grows a Forest of Uncertainty’
With Arun Frey, Jiani Yan and Mark Verhagen



‘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



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

I hope I've been able confer my excitement, enthusiasm and inspiration for:

- 'Assumption free' social science.
- 'Semi-structured' data found 'in the wild'.
- Openness and Transparency in research.

If just one paper, program, or better estimators enters scientific record, everything worth while it 😊.

Thank you for attending!

Thank you to my collaborators!

Don't hesitate to get in touch at:

charles.rahall@demography.ox.ac.uk

and

github.com/crahall!