



# PreFer

## Predicting Fertility data challenge

**How predictable are fertility outcomes?**

Introducing the PreFer data challenge and its potential for fertility research



Elizaveta Sivak



Understanding  
fertility outcomes  
by quantifying the  
(un)predictable



university of  
groningen



Eyra

the dreamteam  
Tom Emery  
Javier Garcia-Bernardo  
Seyit Höcük  
Kasia Karpinska  
Angelica Maineri  
Adriënne Mendrik  
Joris Mulder  
Malvina Nissim  
Paulina Pankowska

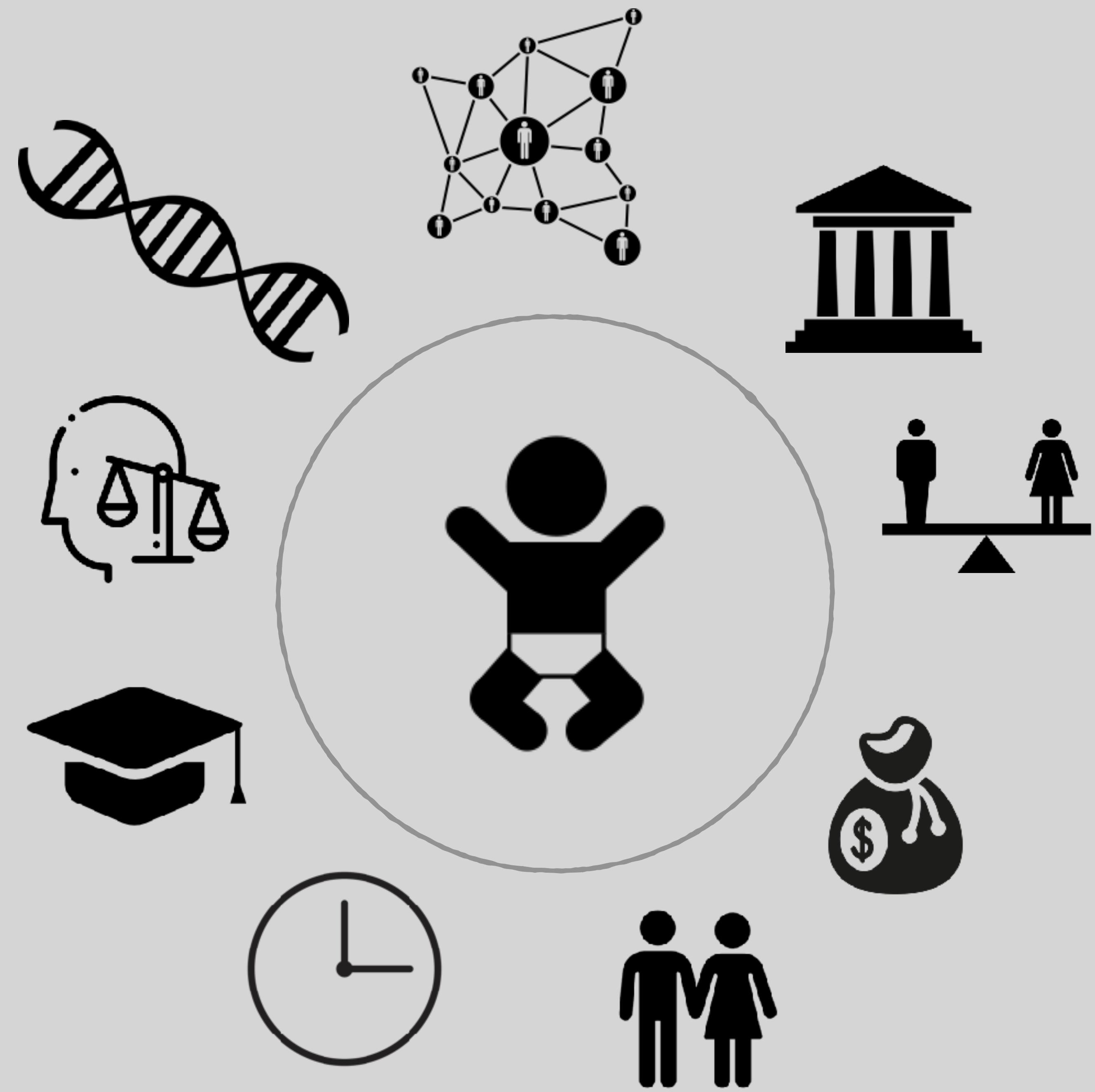


DATA SCIENCE CHALLENGE

# PreFer

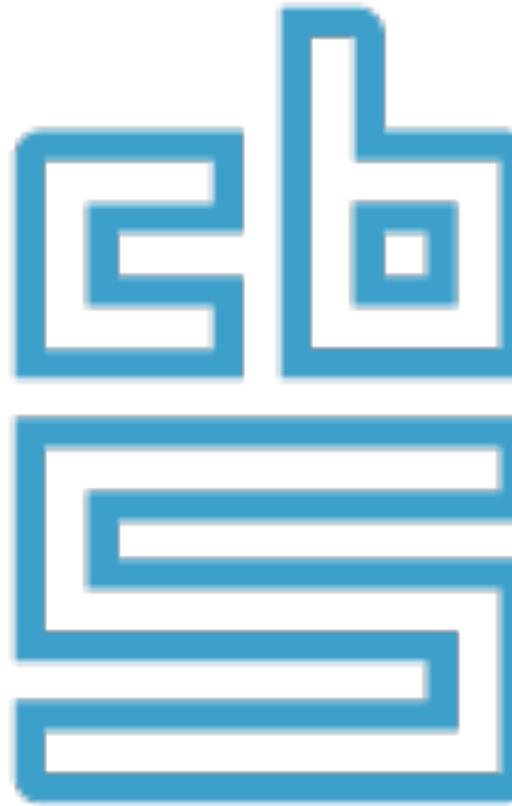


## Predicting Fertility data challenge



variables  
explain  
little

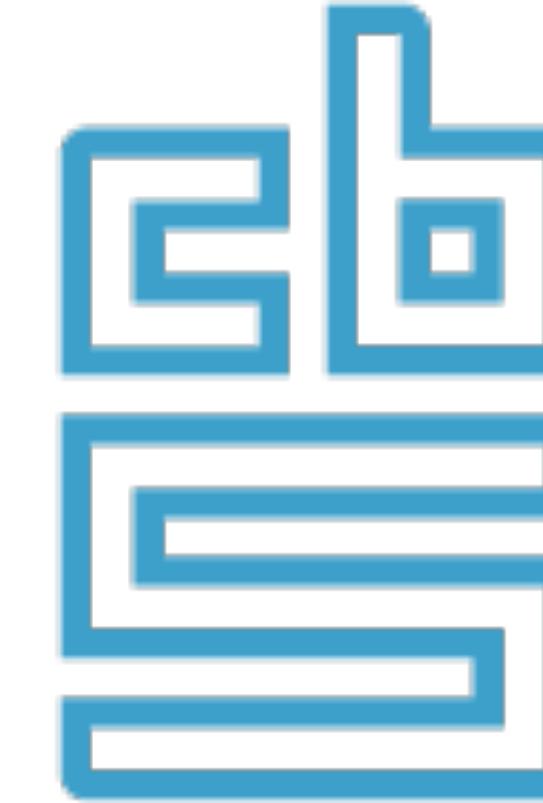
**Fewer  
births  
because of  
study and  
flexwork?**



“total effect on fertility ...  
rather small

variables  
explain  
little

Fewer  
births  
because of  
study and  
flexwork?



“total effect on fertility ...  
rather small

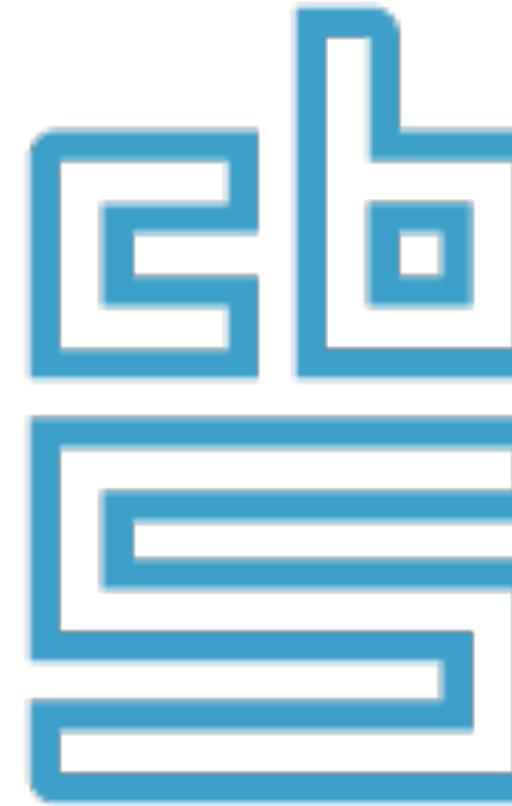


surprising  
patterns

variables  
explain  
little

# Fewer births because of study and flexwork?

“total effect on fertility ...  
rather small



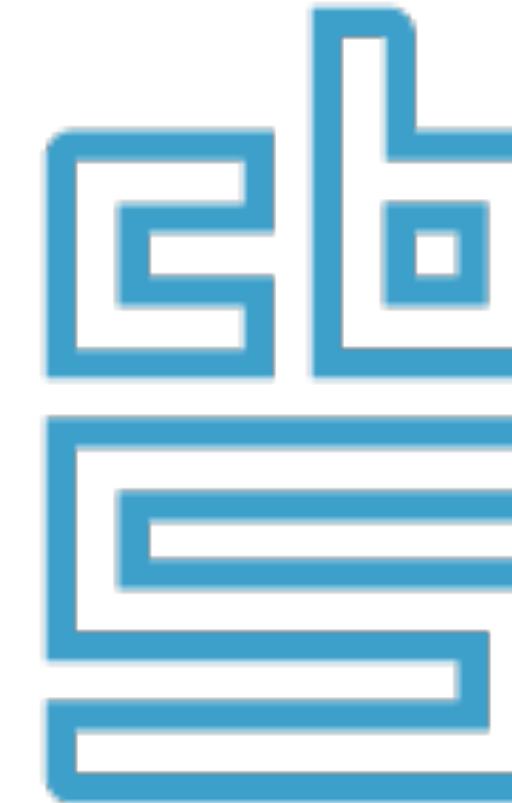
# incomparable results



# surprising patterns

variables  
explain  
little

# Fewer births because of study and flexwork?



“total effect on fertility ...  
rather small

# incomparable results

Downloaded from <http://rstb.royalsocietypublishing.org> on March 28, 2014

PHILOSOPHICAL  
TRANSACTIONS B

[rstb.royalsocietypublishing.org](http://rstb.royalsocietypublishing.org)

Review

See this article: July G, Booth L. 2010. Wealth, fertility and adaptive behaviour in industrial populations. *R. Soc. B* 375: 20100150. doi:10.1098/rstb.2010.0150

Accepted: 25 December 2010

The contribution of H to a theme issue:  
*'Understanding variation in human fertility: what can we learn from evolutionary demography?'*

**Subject areas:**  
Inference, evolution, ecology

**Keywords:**  
income, income, human behavioural ecology,  
inference, ecology, research

**Author for correspondence:**  
See links  
e-mail: [julian@york.ac.uk](mailto:julian@york.ac.uk)

Electronic supplementary material is available at [http://rstb.royalsocietypublishing.org](http://rstb.royalsocietypublishing.org/10.1098/rstb.2010.0150).

THE ROYAL SOCIETY  
PUBLISHING

## Wealth, fertility and adaptive behaviour in industrial populations

Get Stulp<sup>1</sup> and Louise Barrett<sup>2</sup>

<sup>1</sup>Department of Sociology, University of Groningen / Inter-University Center for Social Science Theory and Methodology (CSTS), Grote Rozenstraat 30, 9712 TG Groningen, The Netherlands

<sup>2</sup>Department of Psychology, University of Leiden/Lifebridge, Leiden, NL 2300 RA Leiden, The Netherlands

✉ E-mail: [rstulp@rug.nl](mailto:rstulp@rug.nl)

The link of association between wealth and fertility in contemporary industrial populations has often been used to question the value of an evolutionary perspective on human behaviour. Here, we first present the history of this debate, and the evolutionary explanations for why wealth and fertility—the number of children—are decoupled in modern industrial settings. We suggest that the nature of the relationship between wealth and fertility remains an open question because of the multi-faceted nature of wealth, and because existing cross-sectional studies are ambiguous with respect to how maternal wealth and fertility are linked. A literature review of longitudinal studies on wealth and fertility shows that the majority of these report positive effects of wealth, although levels of fertility seem to fall below those that would maximize fitness. We emphasize that reproductive decision-making reflects a complex interplay between individual and societal factors that needs simple evolutionary interpretation, and highlight the role of economic inequality in fertility decisions. We conclude by discussing whether the wealth–fertility relationship can inform us about the adaptiveness of modern fertility behaviour, and argue against simplistic claims regarding maladaptive behaviour in humans.

### 1. Introduction

In an update to Jane Akers's famous pronouncement of 'a wealth environment' acknowledging that a single man in possession of a good fortune must be in want of a wife [1] (p. 14), Vilfredo Pareto suggested that, in contemporary society, 'there is a regular relationship between wealth and fertility—the married of all classes are less close to a "normal ergonomy" [i.e. living disease]' [2] argued similarly that wealth and fertility were decoupled in industrial societies, given that 'middle-class men did not have more offspring despite higher starting income'. These papers have been used to characterize the 'initial theoretical problem of sociobiology': it is evolutionary theory asserts, individuals are attempting to maximize their fitness; then more resources should translate into a larger number of offspring, as seen in a range of pre-industrial populations [3–9]. The lack of a positive relationship between income and reproductive success also fits with the long-term pattern of fertility decline in many nations; whereby fewer children are born in more prosperous economies [10], whatever people are doing with the resources they acquire is, statistically, they are not, apparently, investing them in having more children.

Now, we revisit Vilfredo Pareto [2] and Pearce [11] using them as a spring board for a review of the literature on wealth and fertility among industrial populations (see also [12]). We then present a new review focused exclusively on longitudinal studies that enable stronger inferences to be made about the links between wealth and reproduction. Finally, we discuss the extent to which the association between wealth and fertility speaks to the issue of maladaptive behaviour, and argue for a more biomedial approach to human fertility.

© RSTB Royal Society Publishing or its licensor. All rights reserved.

## Societies, Part I

rstb.royalsocietypublishing.org

ciated in modern industrial nations debate, beginning with evolutionary and social traits (e.g., fertility) offers a modern industrial critique that our behavior can no longer be explained by traditional measurements in society; in particular, such research between ancestral and modern environments provide insight into the evolution of human behavior. Having made this argument, we turn to the analysis of large-scale databases, even made about which may exist in such data, to represent an excellent tool for investigating the evolutionary

data—Secondary database—

Journal of Social Medicine, Kappel Street,

## Societies, Part II

rstb.royalsocietypublishing.org

utility in industrial populations have been used to argue that economic inequality has led to a decrease in human variability, increasing fertility (and possibly birth rate) with wider availability of large-scale databases, even made about which may exist in such data, to represent an excellent tool for investigating the evolutionary

(doi:10.1098/rstb.2011.0150)

Journal of Social Medicine, Kappel Street,

Inter-University Center for Social Science, Groningen, The Netherlands

<sup>1</sup> Department of Sociology and Nuffield College, University of Oxford, Manor Road, Oxford OX1 3QD, UK.

<sup>2</sup> Department of Psychology, University of Leiden/Lifebridge, Leiden, NL 2300 RA Leiden, The Netherlands

Springer

Springer



# surprising patterns

# non-replicable results

# Replication Crisis

PSYCHOLOGY

## Estimating the reproducibility of psychological science

Open Science Collaboration\*

ROYAL SOCIETY  
OPEN SCIENCE

[rsos.royalsocietypublishing.org](http://rsos.royalsocietypublishing.org)

Research



click for updates

Cite this article: Smaldino PE, McElreath R.  
2016 The natural selection of bad science.

## The natural selection of bad science

Paul E. Smaldino<sup>1</sup> and Richard McElreath<sup>2</sup>

<sup>1</sup>Cognitive and Information Sciences, University of California, Merced, CA 95343, USA

<sup>2</sup>Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

PES, 0000-0002-7133-5620; RME, 0000



General Article

## False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons<sup>1</sup>, Leif D. Nelson<sup>2</sup>, and Uri Simonsohn<sup>1</sup>

<sup>1</sup>The Wharton School, University of Pennsylvania, and <sup>2</sup>Haas School of Business, University of California, Berkeley

aps  
ASSOCIATION FOR  
PSYCHOLOGICAL SCIENCE

Psychological Science  
22(11) 1359–1366  
© The Author(s) 2011  
Reprints and permission:  
[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)  
DOI: 10.1177/0956797611417632  
<http://pss.sagepub.com>  
SAGE

# Replication (crisis) in Family Sociology / Demography?



## Reasons unlikely

- ✓ *Strong methods*
- ✓ *Strong focus on representative data*
- ✓ *Less measurement error*
- ✓ *Open data*
- ✓ *Large N*
- ✓ *Often descriptive*



## Reasons not unlikely

# Replication (crisis) in Family Sociology / Demography?



## Reasons unlikely

- ✓ *Strong methods*
- ✓ *Strong focus on representative data*
- ✓ *Less measurement error*
- ✓ *Open data*
- ✓ *Large N*
- ✓ *Often descriptive*



## Reasons not unlikely

- ✗ *Non-experimental*
- ✗ *Correlational, but little causal inference*
- ✗ *Large N, yet star gazing*
- ✗ *Controlling at will*
- ✗ *Long reign linearity*

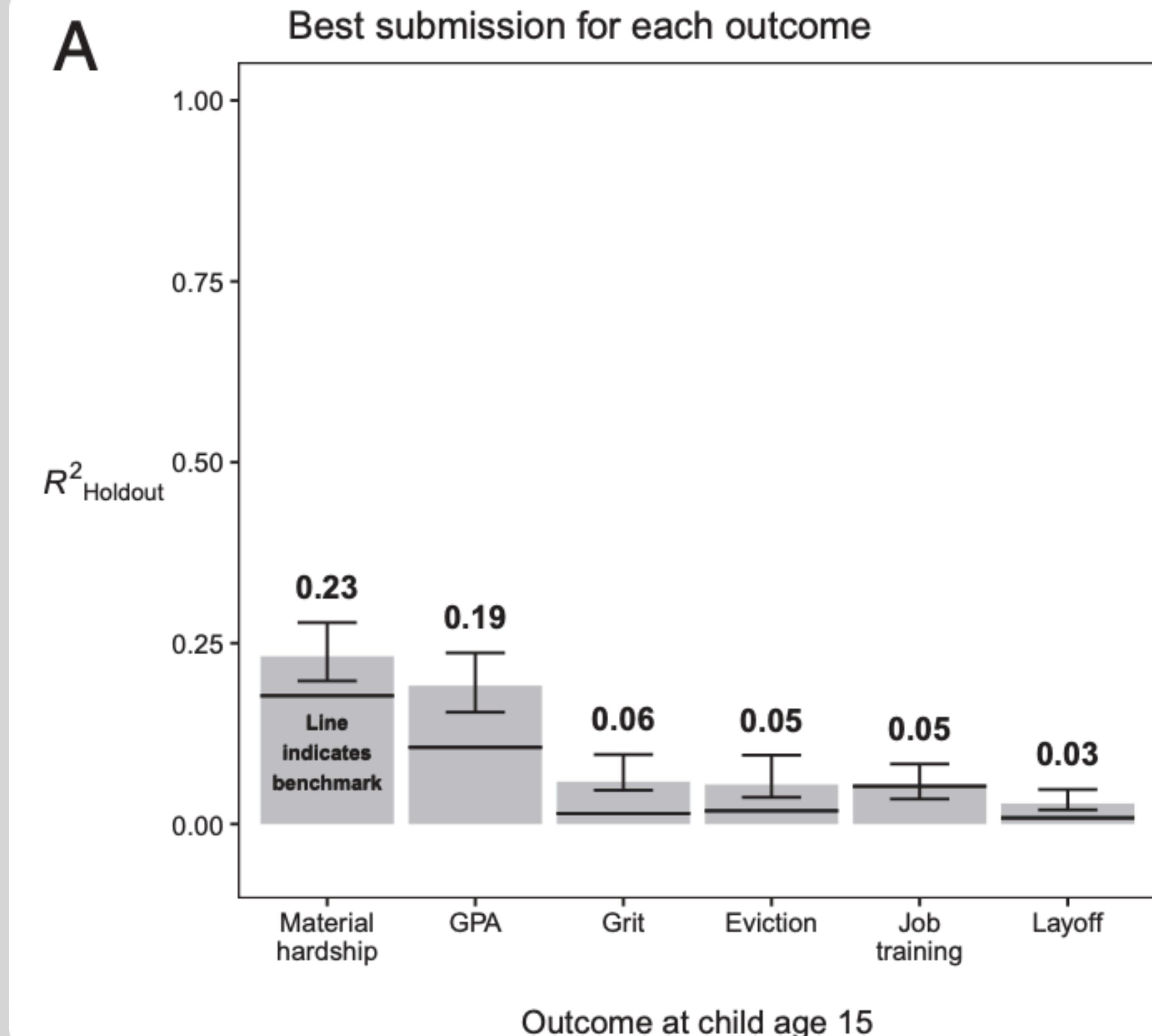
# Predictability Crisis?

Check for updates

## Measuring the predictability of life outcomes with a scientific mass collaboration

Matthew J. Salganik<sup>a,1</sup>, Ian Lundberg<sup>a</sup>, Alexander T. Kindel<sup>a</sup>, Caitlin E. Ahearn<sup>b</sup>, Khaled Al-Ghoneim<sup>c</sup>, Abdullah Almaatouq<sup>d,e</sup>, Drew M. Altschul<sup>f</sup>, Jennie E. Brand<sup>b,g</sup>, Nicole Bohme Carnegie<sup>h</sup>, Ryan James Compton<sup>i</sup>, Debanjan Datta<sup>j</sup>, Thomas Davidson<sup>k</sup>, Anna Filippova<sup>l</sup>, Connor Gilroy<sup>m</sup>, Brian J. Goode<sup>n</sup>, Eaman Jahani<sup>o</sup>, Ridhi Kashyap<sup>p,q,r</sup>, Antje Kirchner<sup>s</sup>, Stephen McKay<sup>t</sup>, Allison C. Morgan<sup>u</sup>, Alex Pentland<sup>v</sup>, Kivan Polimis<sup>w</sup>, Louis Raes<sup>x</sup>, Daniel E. Rigobon<sup>x</sup>, Claudia V. Roberts<sup>y</sup>, Diana M. Stanescu<sup>z</sup>, Yoshihiko Suhara<sup>t</sup>, Adaner Usmani<sup>aa</sup>, Erik H. Wang<sup>x</sup>, Muna Adem<sup>bb</sup>, Abdulla Alhajri<sup>cc</sup>, Bedoor AlShebli<sup>dd</sup>, Redwane Amin<sup>ee</sup>, Ryan B. Amos<sup>y</sup>, Lisa P. Argyle<sup>ff</sup>, Livia Baer-Bositis<sup>gg</sup>, Moritz Büchi<sup>hh</sup>, Bo-Ryehn Chung<sup>ii</sup>, William Eggert<sup>ll</sup>, Gregory Faletto<sup>kk</sup>, Zhilin Fan<sup>ll</sup>, Jeremy Freese<sup>gg</sup>, Tejomay Gadgil<sup>mm</sup>, Josh Gagné<sup>gg</sup>, Yue Gao<sup>nn</sup>, Andrew Halpern-Manners<sup>bb</sup>, Sonia P. Hashim<sup>y</sup>, Sonia Hausen<sup>gg</sup>, Guanhua He<sup>oo</sup>, Kimberly Higuera<sup>gg</sup>, Bernie Hogan<sup>pp</sup>, Ilana M. Horwitz<sup>qq</sup>, Lisa M. Hummel<sup>gg</sup>, Naman Jain<sup>x</sup>, Kun Jin<sup>rr</sup>, David Jurgens<sup>ss</sup>, Patrick Kaminski<sup>bb,tt</sup>, Areg Karapetyan<sup>uu,ww</sup>, E. H. Kim<sup>gg</sup>, Ben Leizman<sup>y</sup>, Naijia Liu<sup>z</sup>, Malte Möser<sup>y</sup>, Andrew E. Mack<sup>x</sup>, Mayank Mahajan<sup>y</sup>, Noah Mandell<sup>ww</sup>, Helge Marahrens<sup>bb</sup>, Diana Mercado-Garcia<sup>aa</sup>, Viola Mocz<sup>xx</sup>, Katarina Mueller-Gastell<sup>gg</sup>, Ahmed Musse<sup>yy</sup>, Qiankun Niu<sup>ee</sup>, William Nowak<sup>zz</sup>, Hamidreza Omidvar<sup>aa</sup>, Andrew Or<sup>y</sup>, Karen Ouyang<sup>y</sup>, Katy M. Pinto<sup>bb</sup>, Ethan Porter<sup>cc</sup>, Kristin E. Porter<sup>dd</sup>, Crystal Qian<sup>y</sup>, Tamkinat Rauf<sup>gg</sup>, Anahit Sargsyan<sup>ee</sup>, Thomas Schaffner<sup>y</sup>, Landon Schnabel<sup>gg</sup>, Bryan Schonfeld<sup>z</sup>, Ben Sender<sup>ff</sup>, Jonathan D. Tang<sup>y</sup>, Emma Tsurkov<sup>gg</sup>, Austin van Loon<sup>gg</sup>, Onur Varol<sup>gg,hh,ii</sup>, Xiafei Wang<sup>ll</sup>, Zhi Wang<sup>hh,ii</sup>, Julia Wang<sup>y</sup>, Flora Wang<sup>ff</sup>, Samantha Weissman<sup>y</sup>, Kirstie Whitaker<sup>kk,ii</sup>, Maria K. Wolters<sup>mm</sup>, Wei Lee Woon<sup>nn</sup>, James Wu<sup>ooo</sup>, Catherine Wu<sup>y</sup>, Kengran Yang<sup>aa</sup>, Jingwen Yin<sup>ll</sup>, Bingyu Zhao<sup>ppp</sup>, Chenyun Zhu<sup>ll</sup>, Jeanne Brooks-Gunn<sup>qqq,rrr</sup>, Barbara E. Engelhardt<sup>yy,ii</sup>, Moritz Hardt<sup>ss</sup>, Dean Knox<sup>x</sup>, Karen Levy<sup>tt</sup>, Arvind Narayanan<sup>y</sup>, Brandon M. Stewart<sup>a</sup>, Duncan J. Watts<sup>uu,ww,vv,ww</sup>, and Sara McLanahan<sup>a,1</sup>

*data challenge:*  
predicting life outcomes  
based on ~6000 variables by  
160 teams  
both theory- & data-driven



# Predictability Crisis?

“

Social scientists studying the life course must find a way to reconcile a widespread belief that understanding has been generated by these data—as demonstrated by more than 750 published journal articles using the Fragile Families data with the fact that the very same data could not yield accurate predictions of these important outcomes.

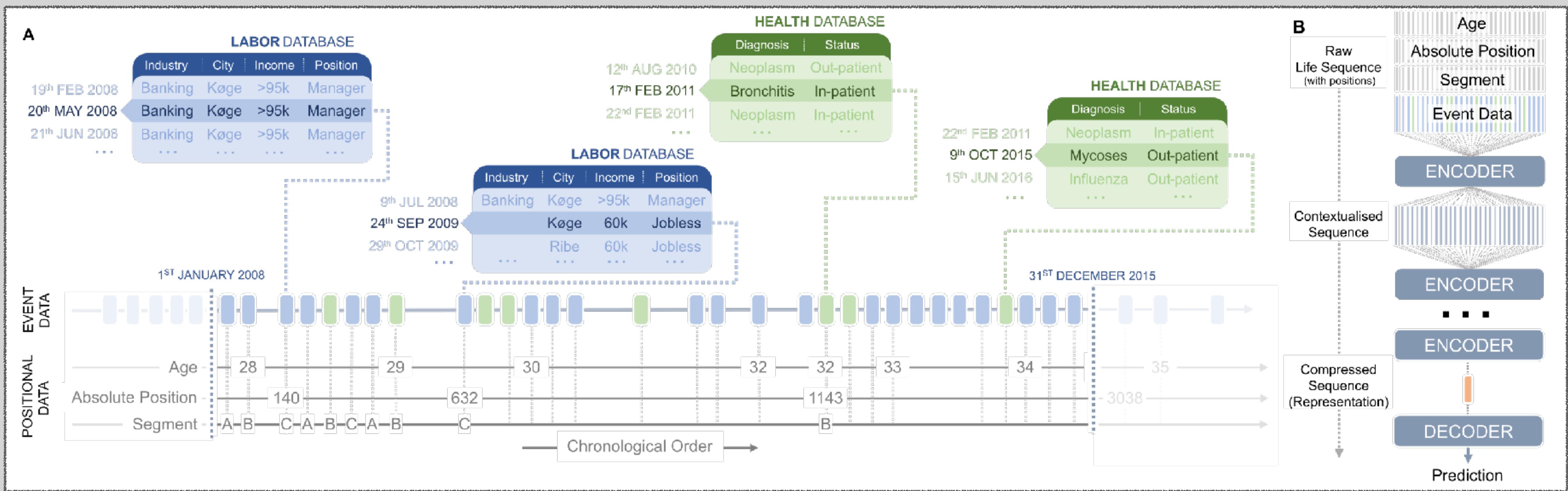
# Using Sequences of Life-events to Predict Human Lives

Germans Savcisen, Tina Eliassi-Rad, Lars Kai Hansen, Laust Hvas Mortensen,  
Lau Lilleholt, Anna Rogers, Ingo Zettler, and Sune Lehmann

June 6, 2023

“

we show that accurate individual predictions are indeed possible



# Prediction

a shift towards **prediction**  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



# out-of-sample predictive ability

- ✓ is easy(ier) to understand
- ✓ can be compared across analytical techniques
- ✓ can be compared across models
- ✓ is less gameable

European Sociological Review, VOLUME 26 | NUMBER 1 | 2010 63–82  
DOI: 10.1111/j.1465-3806.2009.00626.x  
Online publication 9 March 2009

## Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It

Carina Mood

Logistic regression estimates do not behave like linear regression estimates in an important respect. They are affected by omitted variables unrelated to the independent variables in the model. This is something that has gone largely unnoticed by sociologists. Important: Interpret log odds ratios or odds ratios as effect measures, the degree of unobserved heterogeneity in the model. Interpret log-odds ratios or odds ratios for similar models across groups or across models with different independent variables in these problems and possible ways of overcoming them.

### Introduction

The use of logistic regression is routine in the social sciences when studying outcomes that are naturally or necessarily represented by binary variables. Examples are many in stratification research (education/transitions, promotion), demographic research (divorce, childbirth, race-leaving), social medicine (diagnosis, mortality), research into social exclusion (unemployment, benefit take up), and research about political behavior (voting, participation in collective action). When fitting a dichotomous dependent variable, sociologists almost automatically turn to logistic regression, and this practice is generally recommended in textbooks in quantitative methodology. However, our common ways of interpreting results from logistic regression have some important problems.<sup>1</sup>

The problems stem from unobservables, or the fact that we can seldom include in a model all variables that affect an outcome. Unobserved heterogeneity is

© The Author 2009. Published by Oxford University Press, All rights reserved.  
For permission, please e-mail: journals.permissions@oxfordjournals.org



Annual Review of Sociology

Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models

Richard Breen,<sup>1</sup> Kristian Bernt Karlson,<sup>2</sup> and Anders Holm<sup>3</sup>

<sup>1</sup>Nuffield College and Department of Sociology, University of Oxford, OX1 4EN, United Kingdom; email: richard.breen@nuffield.ox.ac.uk

<sup>2</sup>Department of Sociology, University of Copenhagen, DK-1333 Copenhagen, Denmark

<sup>3</sup>Department of Sociology, University of Western Ontario, London, Ontario N6A 3G2, Canada



www.annualreviews.org

- Download Squires
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annual Rev. Sociol. 2010. 39: 63–82

Published online in Advance of Print on May 11, 2010  
The Annual Review of Sociology is reprinted or reproduced by permission of Annual Reviews.

http://www.annualreviews.org/10.1111/j.1465-3806.2009.00626.x

Copyright © 2010 by Annual Reviews.

All rights reserved.

### Keywords

logit, probit, KLIM method, F-standardization, marginal effects, linear probability model, mediation

### Abstract

Methods textbooks in sociology and other social sciences routinely recommend the use of the logit/probit model when an outcome variable is binary, an ordered logit or ordered probit when it is ordinal, and a multinomial logit when it has more than two categories. But these methodological guidelines take little or no account of a body of work that, over the past 10 years, has pointed to problematic aspects of these nonlinear probability models and, particularly, to difficulties in interpreting their parameters. In this review, we draw on that literature to explain the problems, show how they manifest themselves in research, discuss the strengths and weaknesses of alternatives that have been suggested, and point to lines of further analysis.



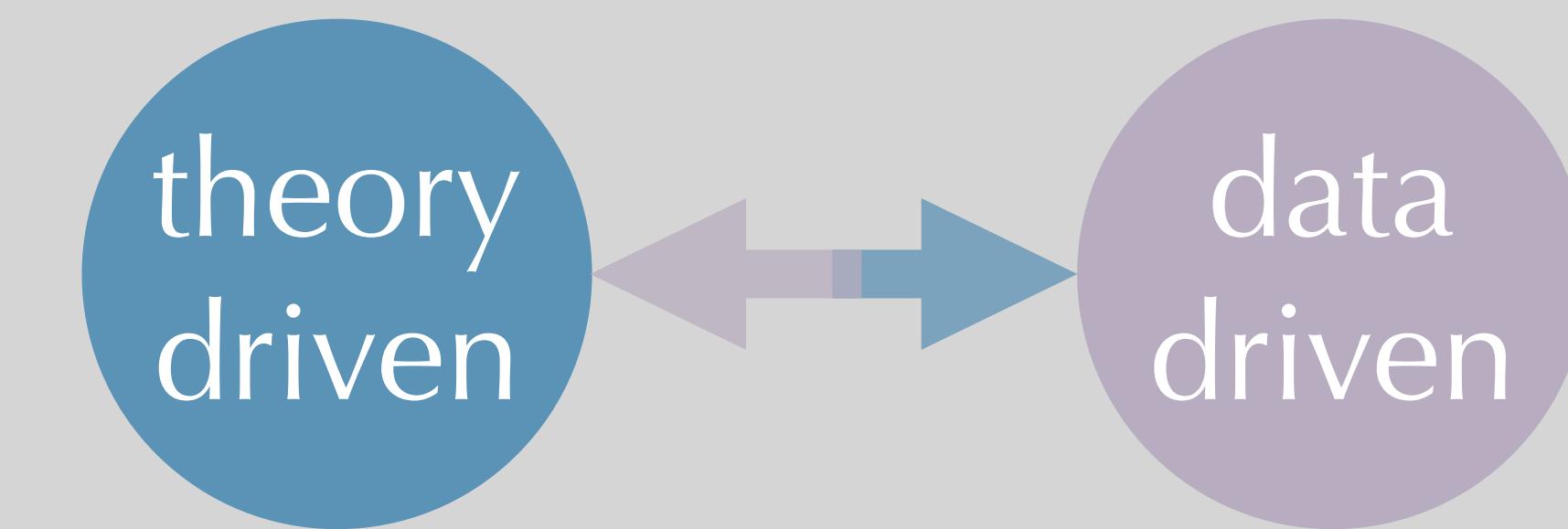
# Prediction

a shift towards **prediction**  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



facilitates dialogue  
theory- and data-  
driven models

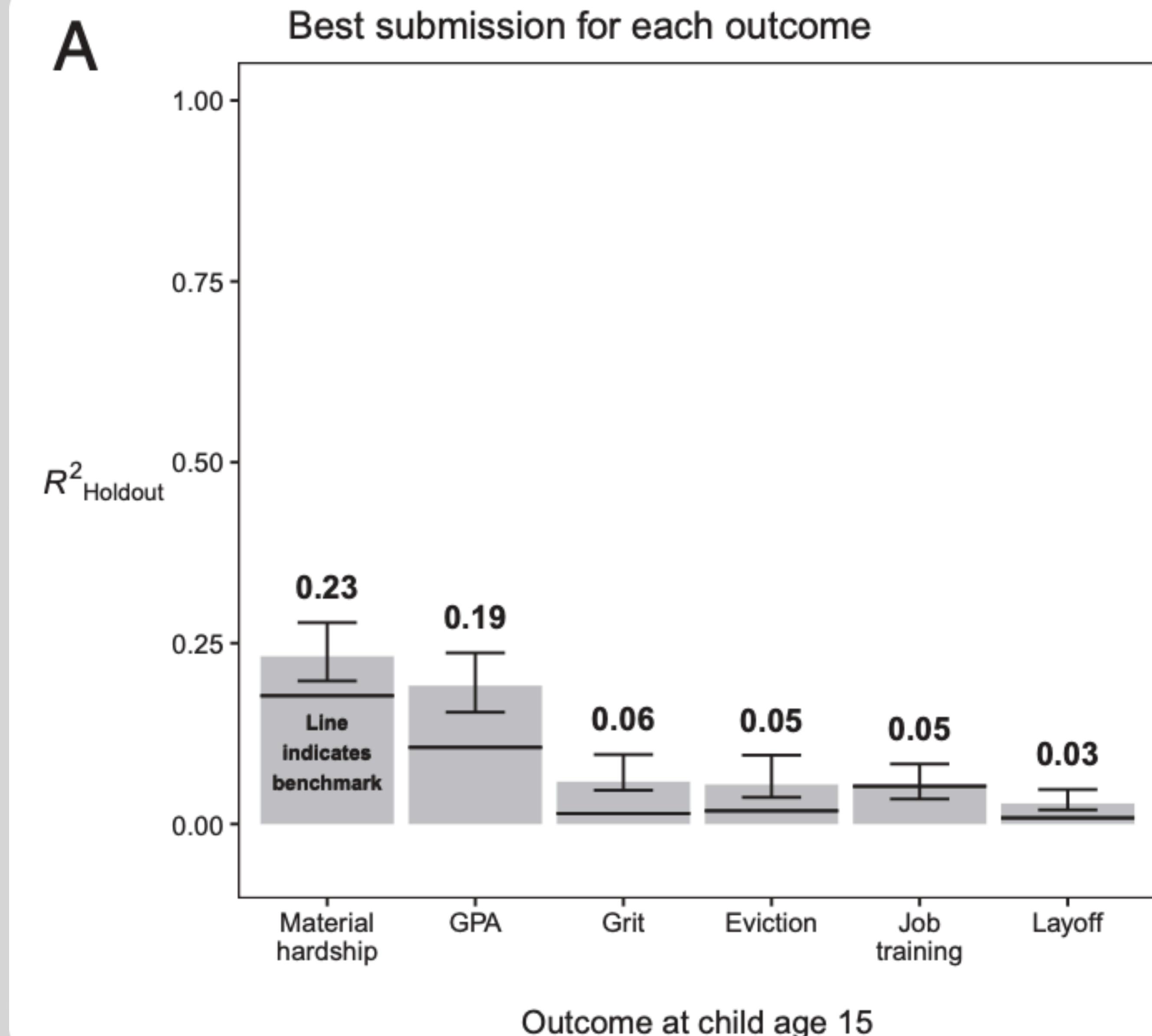
# Predictability Crisis?



## Measuring the predictability of life outcomes with a scientific mass collaboration

Matthew J. Salganik<sup>a,1</sup>, Ian Lundberg<sup>a</sup>, Alexander T. Kindel<sup>a</sup>, Caitlin E. Ahearn<sup>b</sup>, Khaled Al-Ghoneim<sup>c</sup>, Abdullah Almaatouq<sup>d,e</sup>, Drew M. Altschul<sup>f</sup>, Jennie E. Brand<sup>b,g</sup>, Nicole Bohme Carnegie<sup>h</sup>, Ryan James Compton<sup>i</sup>, Debanjan Datta<sup>j</sup>, Thomas Davidson<sup>k</sup>, Anna Filippova<sup>l</sup>, Connor Gilroy<sup>m</sup>, Brian J. Goode<sup>n</sup>, Eaman Jahani<sup>o</sup>, Ridhi Kashyap<sup>p,q,r</sup>, Antje Kirchner<sup>s</sup>, Stephen McKay<sup>t</sup>, Allison C. Morgan<sup>u</sup>, Alex Pentland<sup>v</sup>, Kivan Polimis<sup>w</sup>, Louis Raes<sup>x</sup>, Daniel E. Rigobon<sup>x</sup>, Claudia V. Roberts<sup>y</sup>, Diana M. Stanescu<sup>z</sup>, Yoshihiko Suhara<sup>t</sup>, Adaner Usmani<sup>aa</sup>, Erik H. Wang<sup>x</sup>, Muna Adem<sup>bb</sup>, Abdulla Alhajri<sup>cc</sup>, Bedoor AlShebli<sup>dd</sup>, Redwane Amin<sup>ee</sup>, Ryan B. Amos<sup>y</sup>, Lisa P. Argyle<sup>ff</sup>, Livia Baer-Bositis<sup>gg</sup>, Moritz Büchi<sup>hh</sup>, Bo-Ryehn Chung<sup>ii</sup>, William Eggert<sup>ll</sup>, Gregory Faletto<sup>kk</sup>, Zhilin Fan<sup>ll</sup>, Jeremy Freese<sup>gg</sup>, Tejomay Gadgil<sup>mm</sup>, Josh Gagné<sup>gg</sup>, Yue Gao<sup>nn</sup>, Andrew Halpern-Manners<sup>bb</sup>, Sonia P. Hashim<sup>y</sup>, Sonia Hausen<sup>gg</sup>, Guanhua He<sup>oo</sup>, Kimberly Higuera<sup>gg</sup>, Bernie Hogan<sup>pp</sup>, Ilana M. Horwitz<sup>qq</sup>, Lisa M. Hummel<sup>gg</sup>, Naman Jain<sup>x</sup>, Kun Jin<sup>rr</sup>, David Jurgens<sup>ss</sup>, Patrick Kaminski<sup>bb,tt</sup>, Areg Karapetyan<sup>uu,ww</sup>, E. H. Kim<sup>gg</sup>, Ben Leizman<sup>y</sup>, Naijia Liu<sup>z</sup>, Malte Möser<sup>y</sup>, Andrew E. Mack<sup>x</sup>, Mayank Mahajan<sup>y</sup>, Noah Mandell<sup>ww</sup>, Helge Marahrens<sup>bb</sup>, Diana Mercado-Garcia<sup>aa</sup>, Viola Mocz<sup>xx</sup>, Katarina Mueller-Gastell<sup>gg</sup>, Ahmed Musse<sup>yy</sup>, Qiankun Niu<sup>ee</sup>, William Nowak<sup>zz</sup>, Hamidreza Omidvar<sup>aaa</sup>, Andrew Or<sup>y</sup>, Karen Ouyang<sup>y</sup>, Katy M. Pinto<sup>bbb</sup>, Ethan Porter<sup>cc</sup>, Kristin E. Porter<sup>dd</sup>, Crystal Qian<sup>y</sup>, Tamkinat Rauf<sup>gg</sup>, Anahit Sargsyan<sup>eee</sup>, Thomas Schaffner<sup>y</sup>, Landon Schnabel<sup>gg</sup>, Bryan Schonfeld<sup>z</sup>, Ben Sender<sup>fff</sup>, Jonathan D. Tang<sup>y</sup>, Emma Tsurkov<sup>gg</sup>, Austin van Loon<sup>gg</sup>, Onur Varol<sup>ggghhhjjj</sup>, Xiafei Wang<sup>ll</sup>, Zhi Wang<sup>hhhjjj</sup>, Julia Wang<sup>y</sup>, Flora Wang<sup>ff</sup>, Samantha Weissman<sup>y</sup>, Kirstie Whitaker<sup>kkkll</sup>, Maria K. Wolters<sup>mmm</sup>, Wei Lee Woon<sup>nnn</sup>, James Wu<sup>ooo</sup>, Catherine Wu<sup>y</sup>, Kengran Yang<sup>aaa</sup>, Jingwen Yin<sup>ll</sup>, Bingyu Zhao<sup>ppp</sup>, Chenyun Zhu<sup>ll</sup>, Jeanne Brooks-Gunn<sup>qqqrrr</sup>, Barbara E. Engelhardt<sup>yy</sup>, Moritz Hardt<sup>sss</sup>, Dean Knox<sup>x</sup>, Karen Levy<sup>ttt</sup>, Arvind Narayanan<sup>y</sup>, Brandon M. Stewart<sup>a</sup>, Duncan J. Watts<sup>uuu,vvv,wwwww</sup>, and Sara McLanahan<sup>a,1</sup>

*data challenge:*  
predicting life outcomes  
based on ~6000 variables by  
160 teams  
both theory- & data-driven



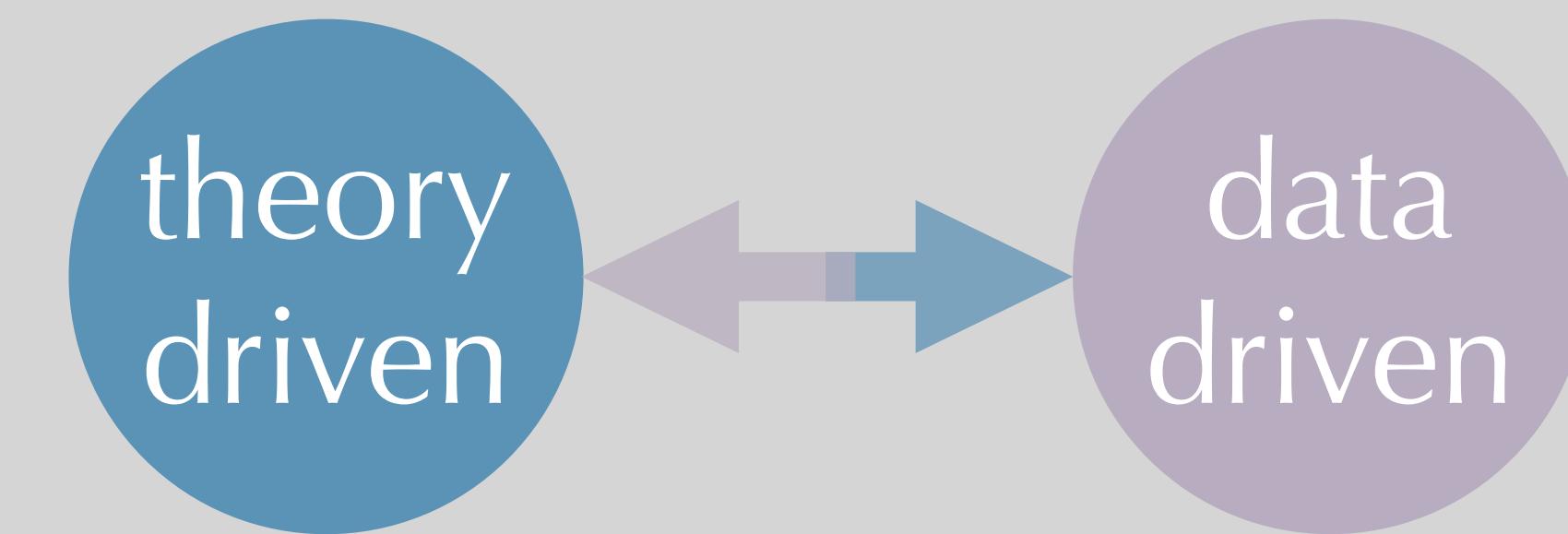
# Prediction

a shift towards **prediction**  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



facilitates dialogue  
theory- and data-  
driven models



measure of distance  
theory and practice



out-of-sample predictive ability  
is a measure of how useful  
our theory is in the real world

Articles

## The perils of policy by p-value: Predicting civil conflicts

Michael D Ward

*Department of Political Science, Duke University*

Brian D Greenhill

*Department of Political Science, University of Washington*

Kristin M Bakke

*Department of Political Science, University College London*

Journal of  
*Peace*  
Research

Journal of Peace Research  
47(4) 363–375  
© The Author(s) 2010  
Reprints and permission:  
[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)  
DOI: 10.1177/0022343309356491  
[jpr.sagepub.com](http://jpr.sagepub.com)

SAGE

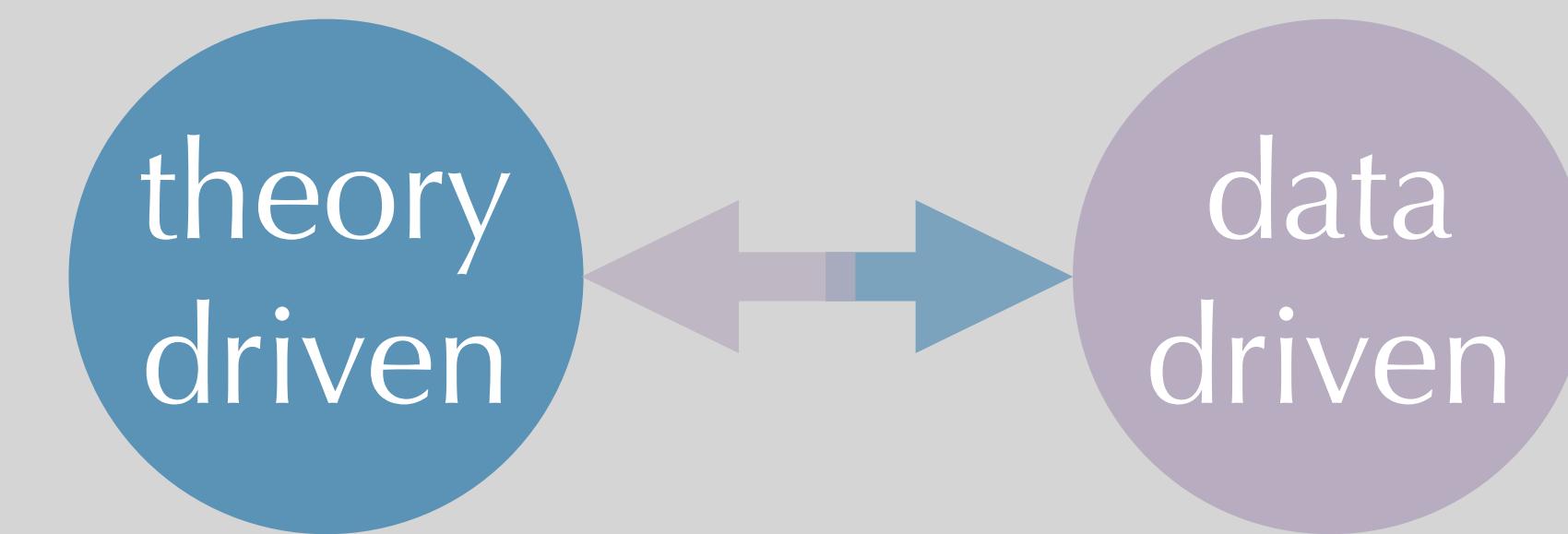
# Prediction

a shift towards **prediction**  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



facilitates dialogue  
theory- and data-  
driven models



measure of distance  
theory and practice

# data challenge



theory  
driven

data  
driven

theory- and data-driven teams  
engage in common task  
using common data  
and common metric

## ⌚ Active Competitions

Hotness ▾



### Google AI4Code – Understand Code in...

Predict the relationship between co...

Featured

Code Competition · 166 Teams

**\$150,000**

3 months to go



### JPX Tokyo Stock Exchange Prediction

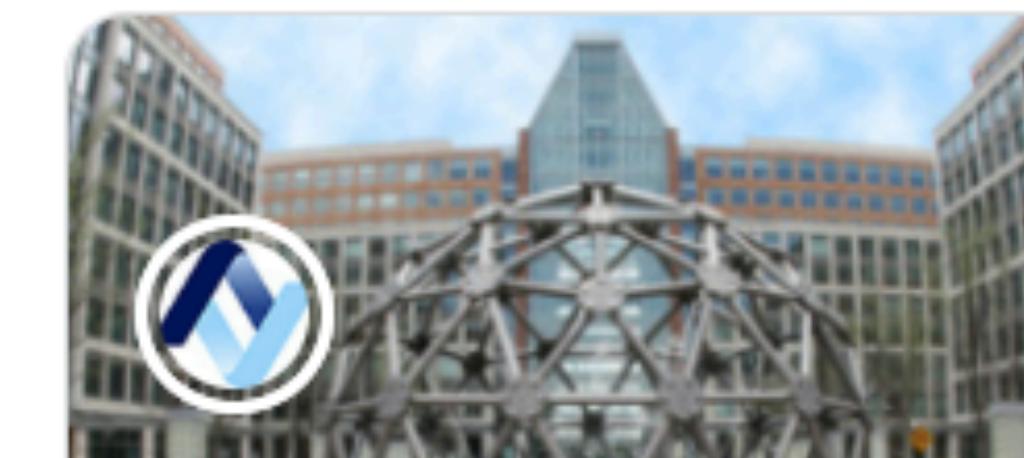
Explore the Tokyo market with your ...

Featured

Code Competition · 983 Teams

**\$63,000**

2 months to go



### U.S. Patent Phrase to Phrase Matching

Help Identify Similar Phrases in U.S. ...

Featured

Code Competition · 1258 Teams

**\$25,000**

a month to go



### Foursquare - Location Matching

Match point of interest data across ...

Featured

Code Competition · 489 Teams

**\$25,000**

2 months to go

“

secret sauce of data science  
Donoho, 2015

# Prediction Benchmarks

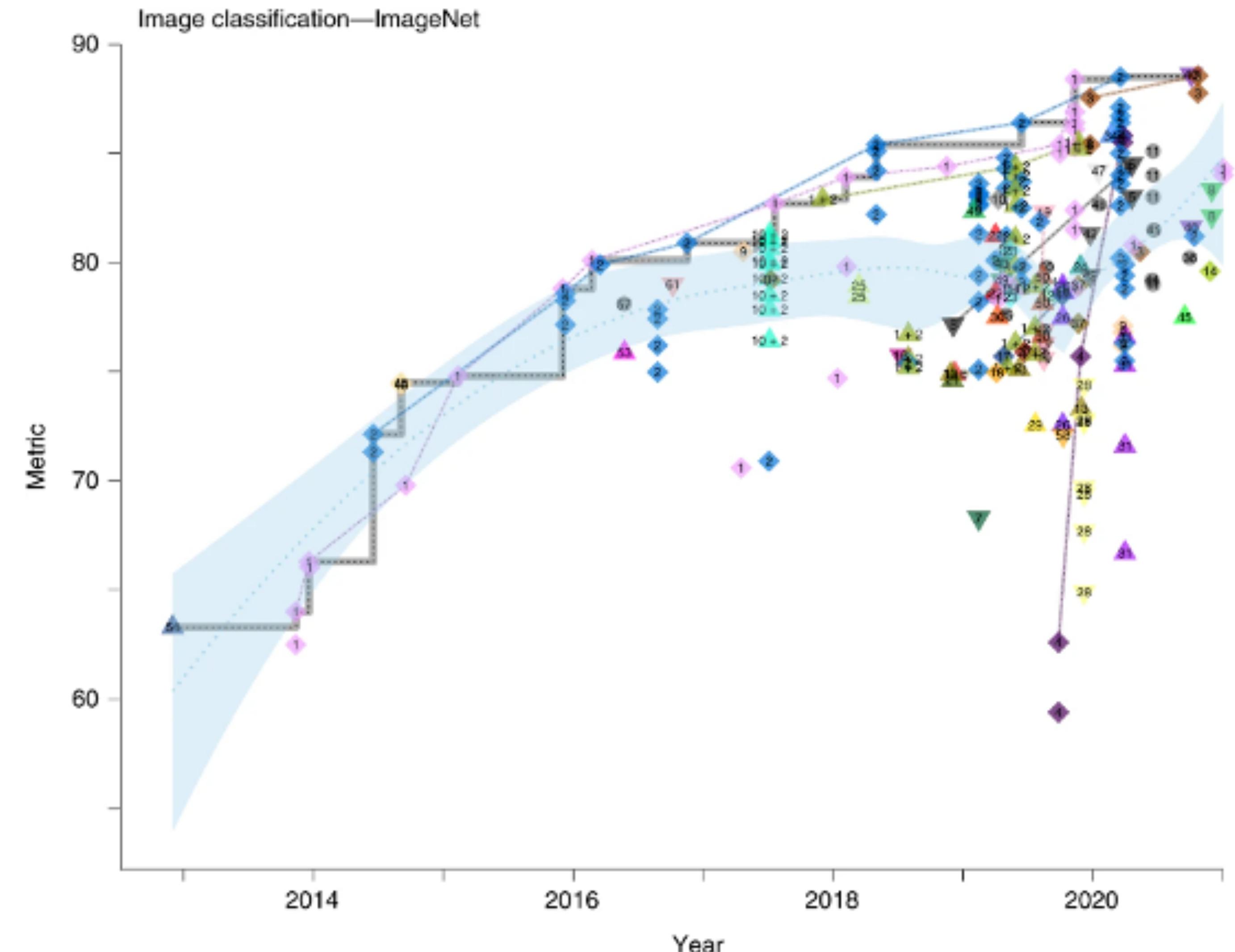
“

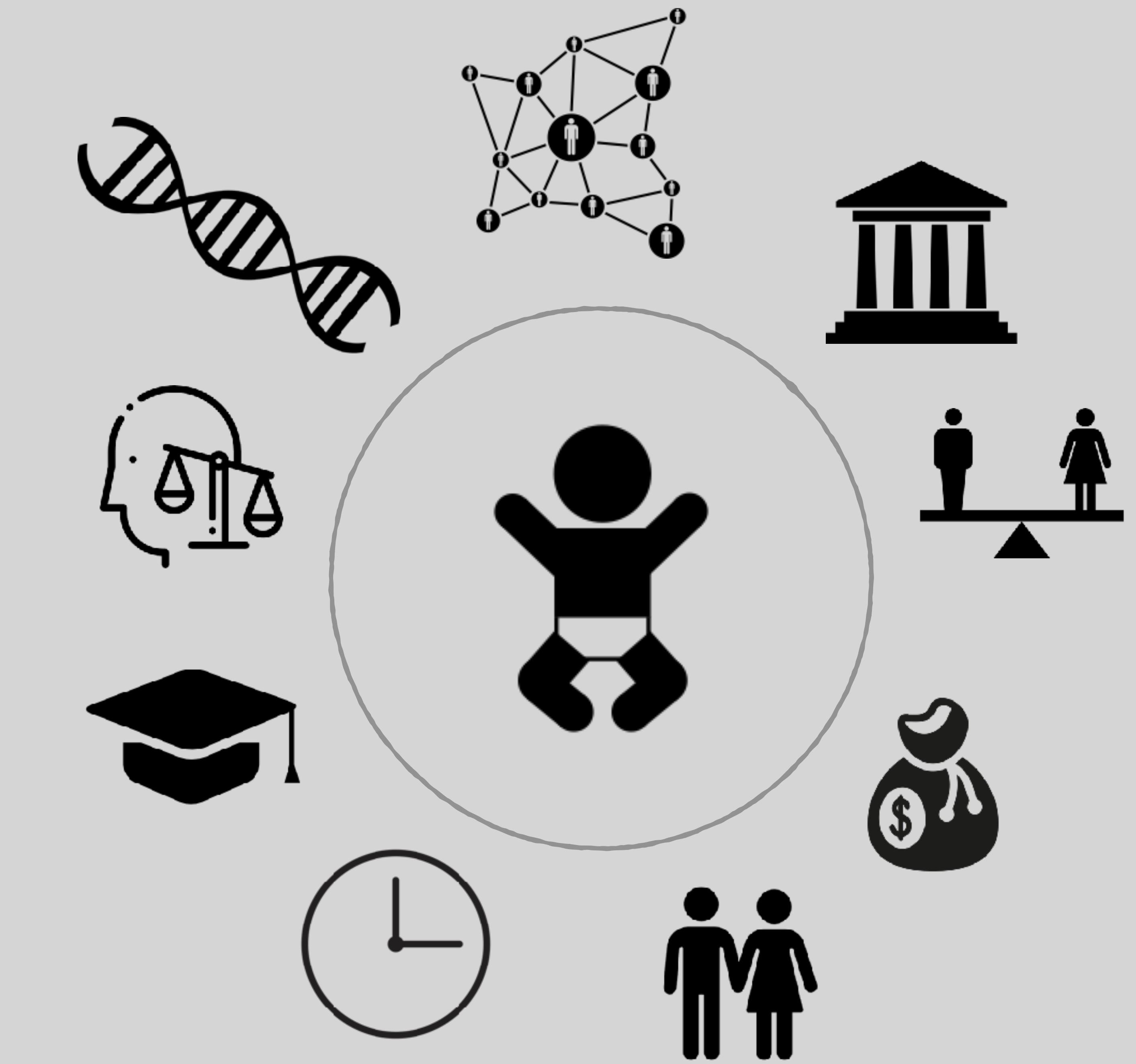
Progress usually comes from many small improvements; a change of 1% can be a reason to break out the champagne

Liberman, 2012

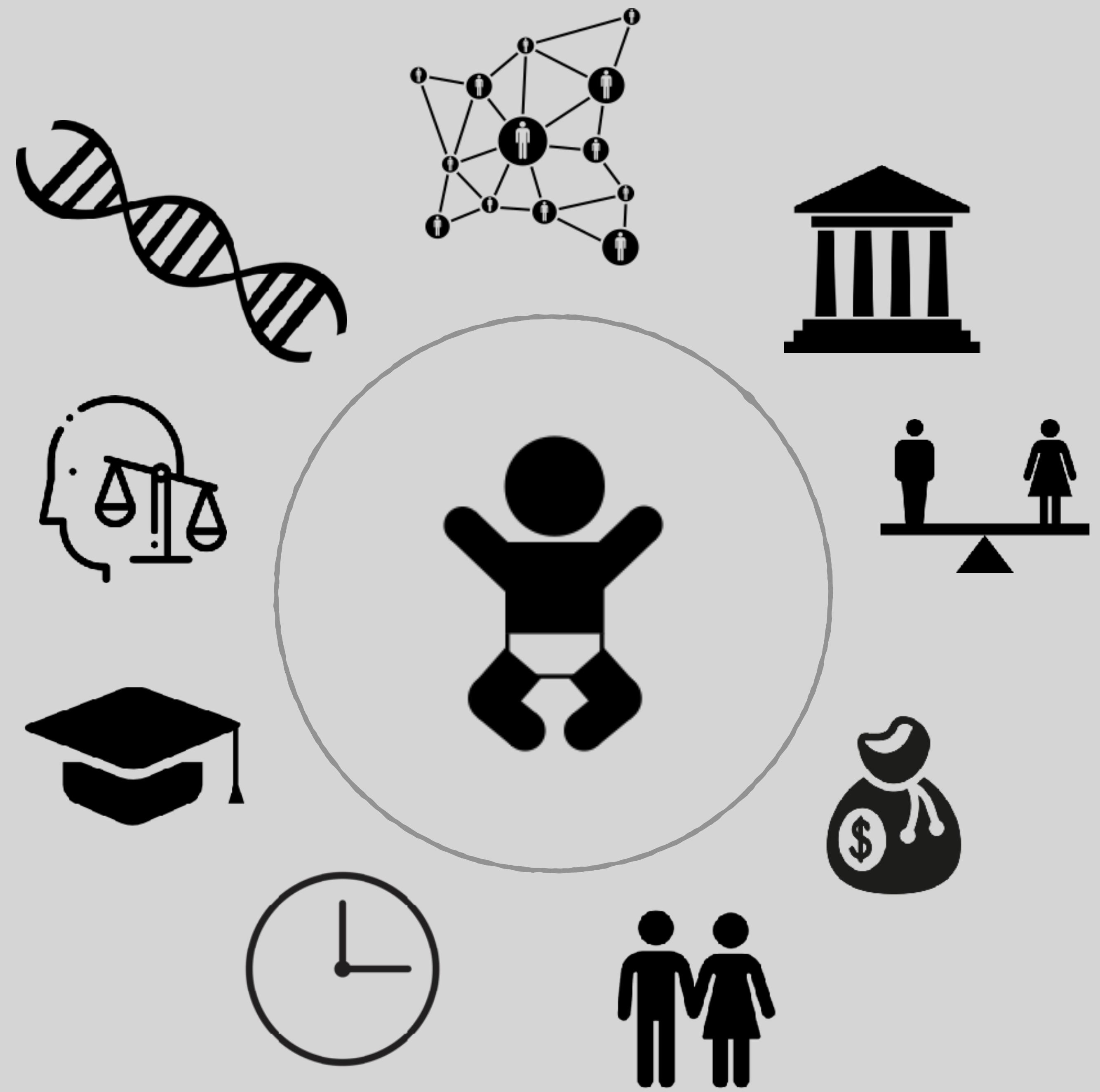
## Progress in accuracy over time for ImageNet.

From: [Research community dynamics behind popular AI benchmarks](#)

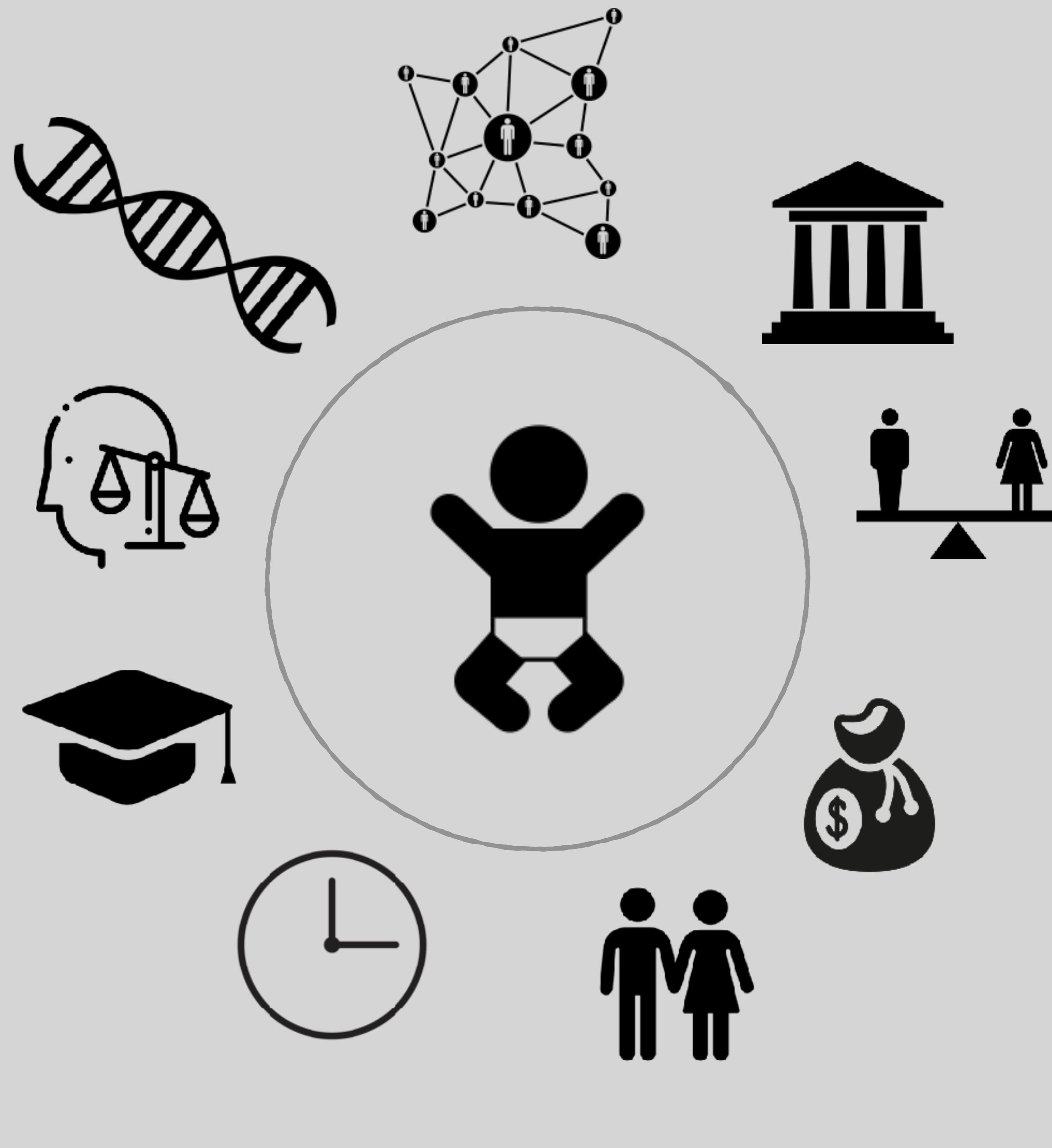




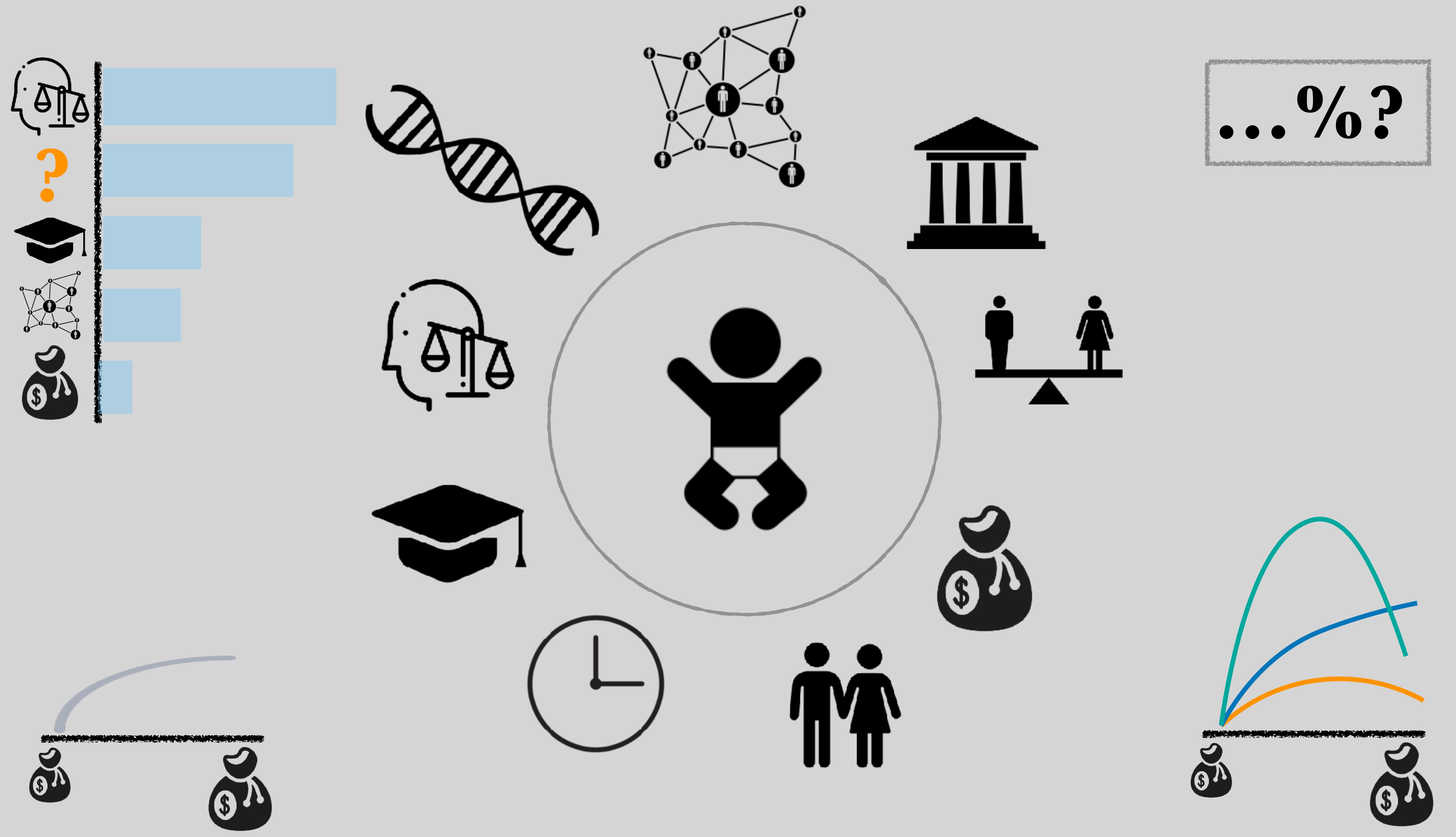
... %?



... %?



... %?



# Pre Fer 9/10

## Predicting Fertility data challenge



# Predicting Fertility data challenge

theory- and data-driven teams  
engage in common task  
using common data  
and common metric

# Common Task



To predict who will have a child  
in the next three years

**Outcome  
['21-'23]**

**Background data**  
[data from 1995/2007  
up to 2020]  
[ages 18-45]

70%

**TRAIN**

30%

**HOLD-  
OUT**

# Common Task

## Rationale

 **Difficult!**

[minimal policy-relevant test]

 **Parity-specific**

 **Data availability**  
[longer prediction timespan means fewer background data]



To predict who will have a child in the next three years

**Outcome  
['21-'23]**

**Background data**  
[data from 1995/2007 up to 2020]  
[ages 18-45]

70%

**TRAIN**

30%

**HOLD-OUT**

# Common Data

1



LISS panel

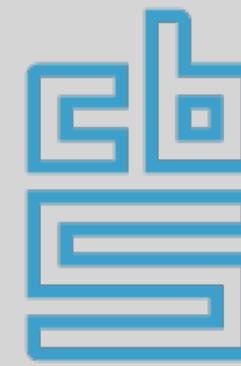
panel survey [2007 - current]

~1200 cases

‘objective’ and ‘subjective’ measures

15 waves x 10 core surveys  
1000s variables

2



Social Statistics Netherlands

register data [1995 - current]

6 million cases

‘objective’ measures

100s variables  
(10000s variables?)

# Common Data

1



LISS panel

Background variables

Health

Religion and Ethnicity

Social Integration and Leisure

Family and Household

Work and Schooling

Personality

Politics and Values

Economic Situation:

Assets, Income, Housing

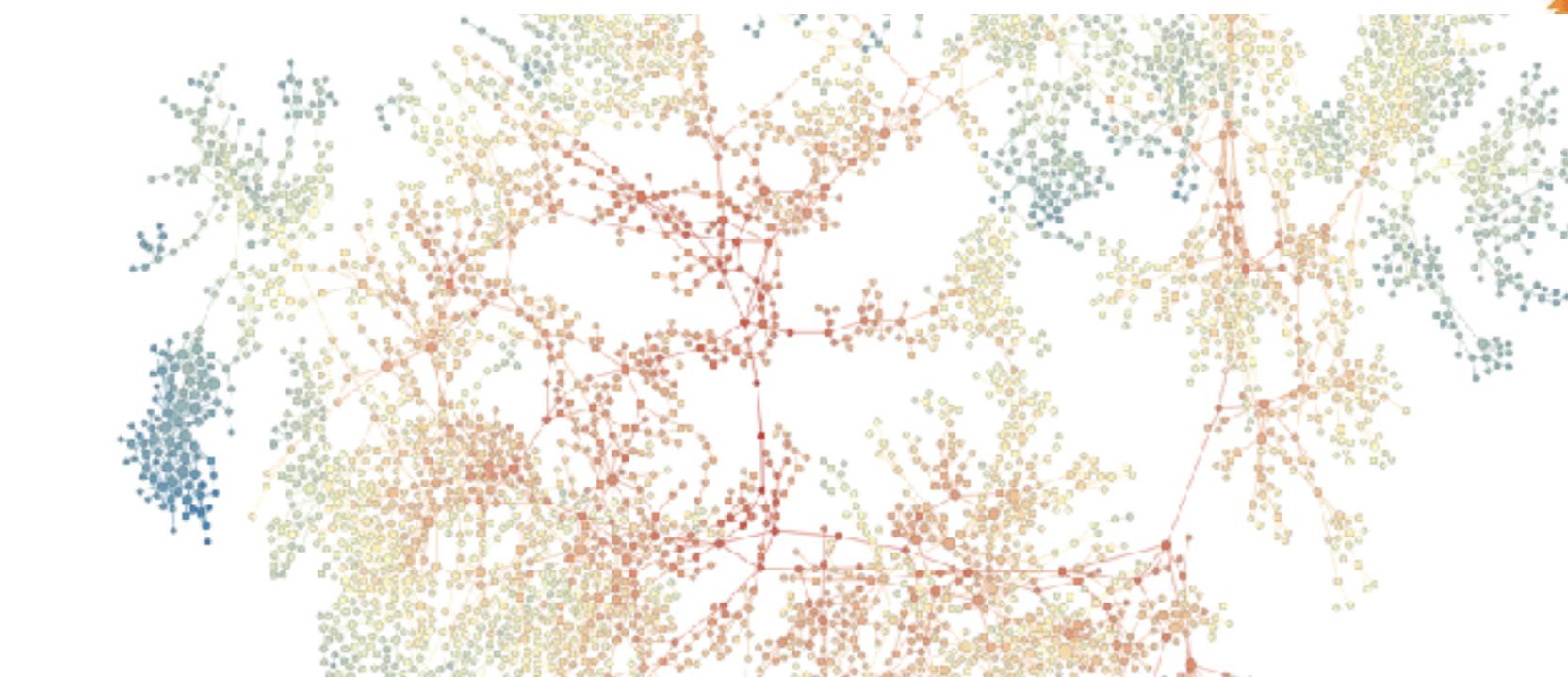
2



Social Statistics Netherlands

**Population-Scale  
Network Analysis**

*A new research data infrastructure in  
computational social science*



# Common Metric: F1

$$\text{precision} = \frac{\text{green}}{\text{green} + \text{red}}$$

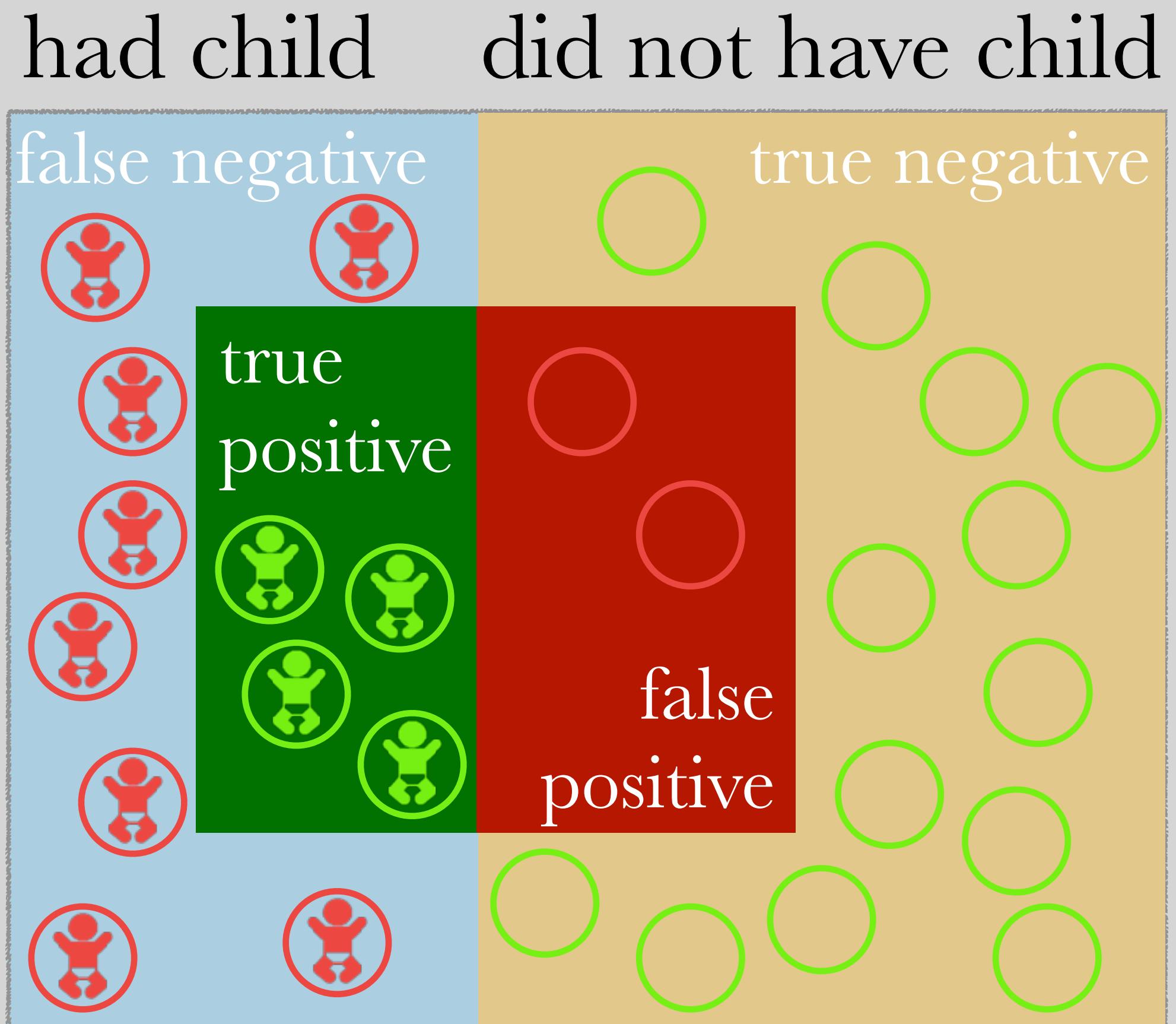
among those predicted  
to have child,  
% who had child

$$\text{recall} = \frac{\text{green}}{\text{blue} + \text{green}}$$

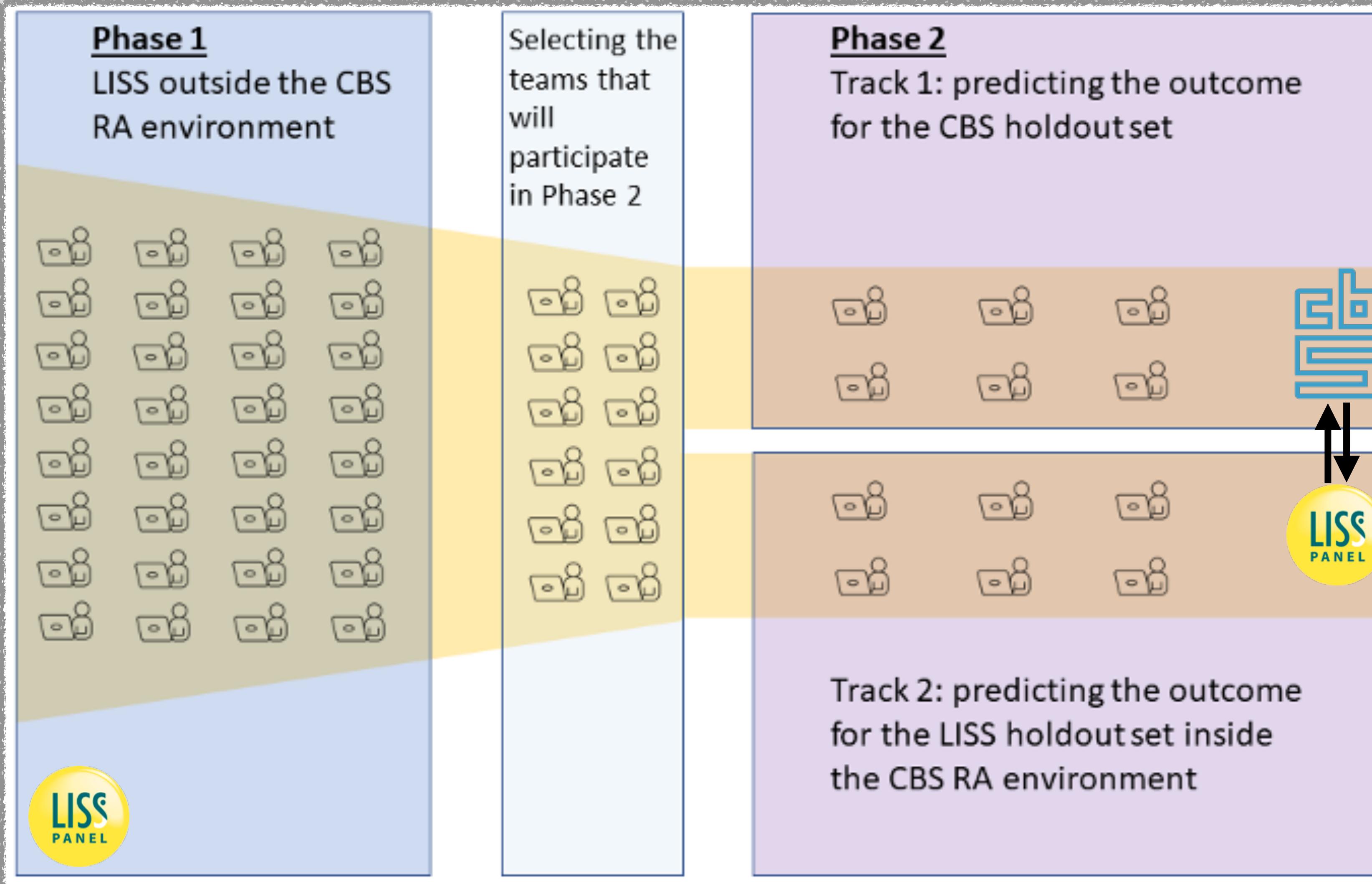
among those who  
had child,  
% predicted to have child

harmonic mean of precision and recall:

$$\text{F1} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$



# Overview



## Evaluation criteria:

- ✓ F1 score [*3 winners*]
- ✓ Qualitative criteria [*2 winners*]
  - innovativeness: novel approach from social- or data science
  - improving understanding: what have we learned about fertility



**Eyra**

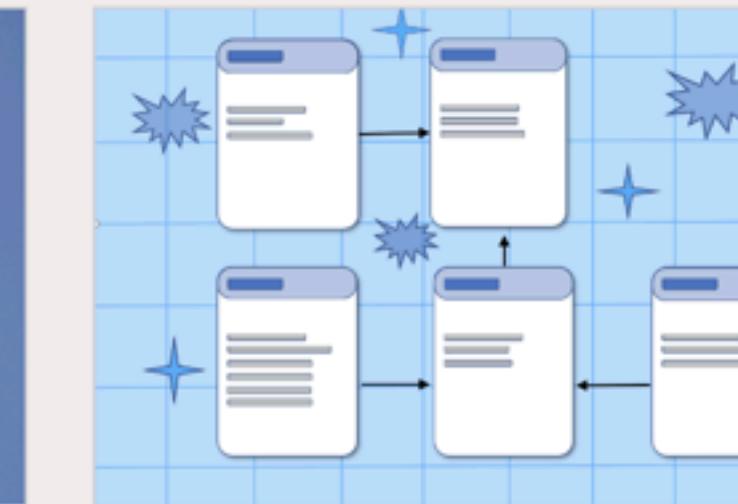


## DETAILS ABOUT THE CHALLENGE



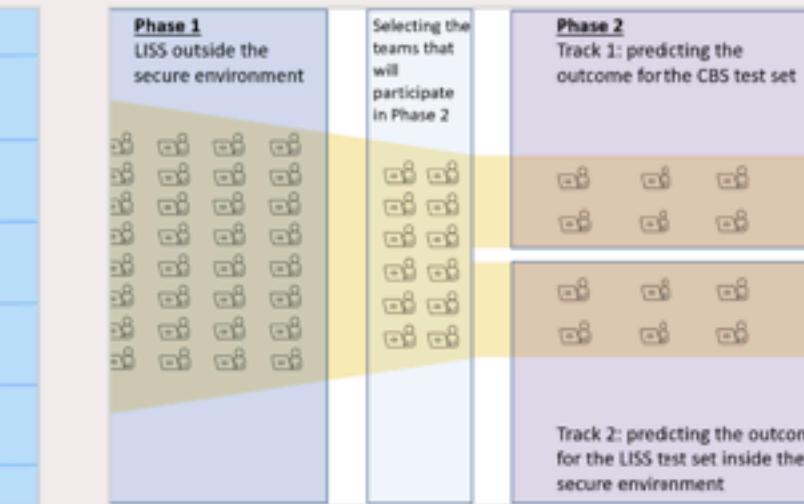
### The goal and research questions

The goal of the data challenge is to assess the current predictability of individual-level fertility and improve our understanding of fertility behaviour.



### Data

PreFer uses two datasets: the LISS panel and Dutch population registries data.



### Phases of the challenge

The challenge includes two phases.



### Evaluation and winners

Evaluation criteria and determining the winners.



### Submission

Description of the submission process.



### Special issue and community paper

Results will be published in a community paper and in a special issue of a journal.

# Timeline

January-March 2024	April-May 2024	June-September 2024	October 2024	2024 - 2025
<b>Application</b>  Sign up to get notified when the application opens	<b>Phase 1</b>  Predict the outcome for the LISS holdout set	<b>Phase 2</b>  Predict the outcome for the LISS holdout set or/and CBS holdout set inside CBS RA	<b>Evaluation</b>  Evaluating the submissions and announcing the winners	<b>Analysis</b>  Analyzing submitted methods, preparing publications

↑  
select winner phase 1 & teams for phase 2

# Who Can Participate?

1



LISS panel

**No restrictions, but:**

Access after data agreement

Data on your own computer

2



Social Statistics Netherlands

**Restricted access**

Only after vetting procedure

Remote secure environment

No uploads/downloads possible

Only available from within  
European Union + selected  
countries

# Why Participate?



- eternal glory
- talk at and paid-for-trip to conference in exotic Netherlands
- publish paper [special issue]
- contribute to fertility research / computational social science
- work with amazing data
- test your ML skills and favourite algorithms
- students: learn new skills
- use in/for teaching



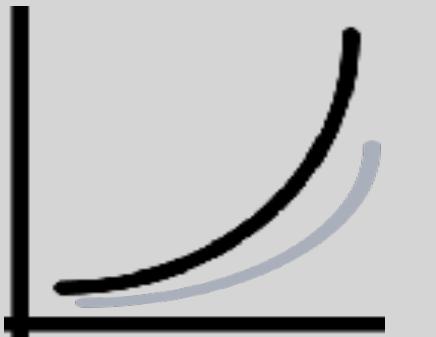
# What Can We Learn?

## Science

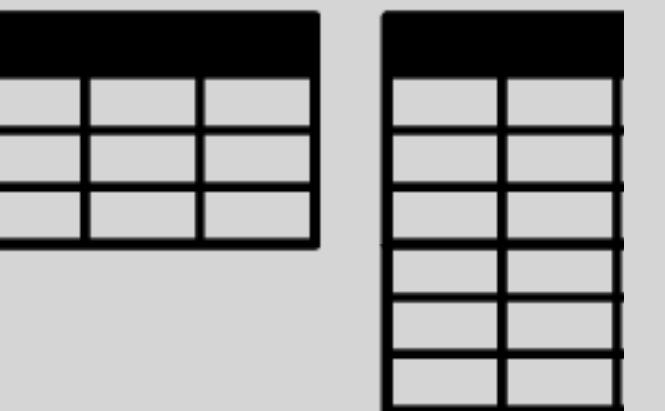
Establish predictive ability  
and set benchmarks



Novel understanding through  
e.g. non-linearity, interactions



Scale versus scope,  
long versus wide



“subjective” versus  
“objective” measures



transfer learning

$$1+1=3$$

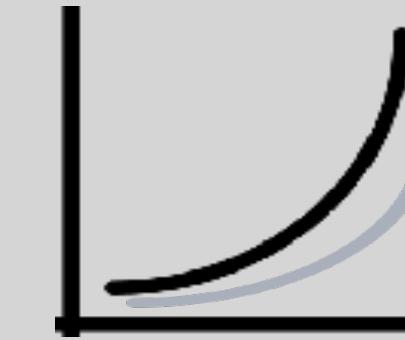
success further in the future

2025?

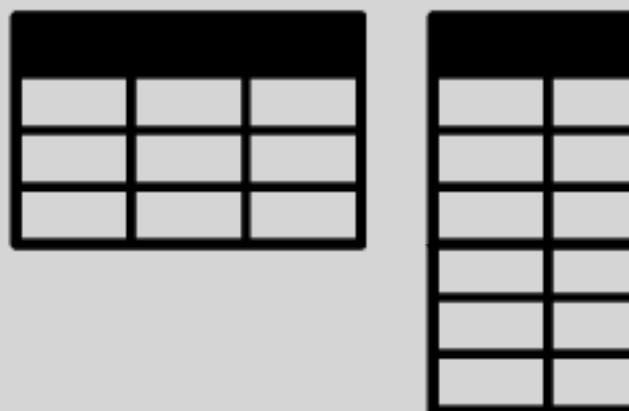
# What Can We Learn?

## Science

Establish predictive ability  
and set benchmarks



Scale versus scope,  
long versus wide



“subjective” versus  
“objective” measures



transfer learning

$$1+1=3$$

success further in the future

2025?

## Policy

Debate on using intentions  
in forecasting



Quantifying unmet needs



## Predicting Fertility data challenge

- Be a part of a unique data challenge
- Contribute to fertility research & computational social sciences
- Publish research
- Work with amazing data:
  - LISS panel
  - Dutch population registries

**SIGN UP HERE!**



[preferdatachallenge.nl](http://preferdatachallenge.nl)  
[g.stulp@rug.nl](mailto:g.stulp@rug.nl)