# How should Sociologists build models in an age of Big Data, Machine Learning, and High Performance Computing?

12th October, 2025

Daniel Valdenegro, Duiyi Dai, Jiani Yan and **Charlie Rahal**

Invited Talk at Tsinghua University

October 12, 2025



# The plan for today!

We're going to:

- Give a short talk about model uncertainty
- Introduce RobustiPy with some simple:
  - Simulated examples
  - Empirical examples
- Time permitting, host a short discussion about the future of model uncertainty work.

RESEARCH ARTICLE | SOCIAL SCIENCES |

# Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty

Nate Breznau, Eike Mark Rinke, Alexander Wuttke, +162, and Tomasz Żółtak Authors Info & Affiliations

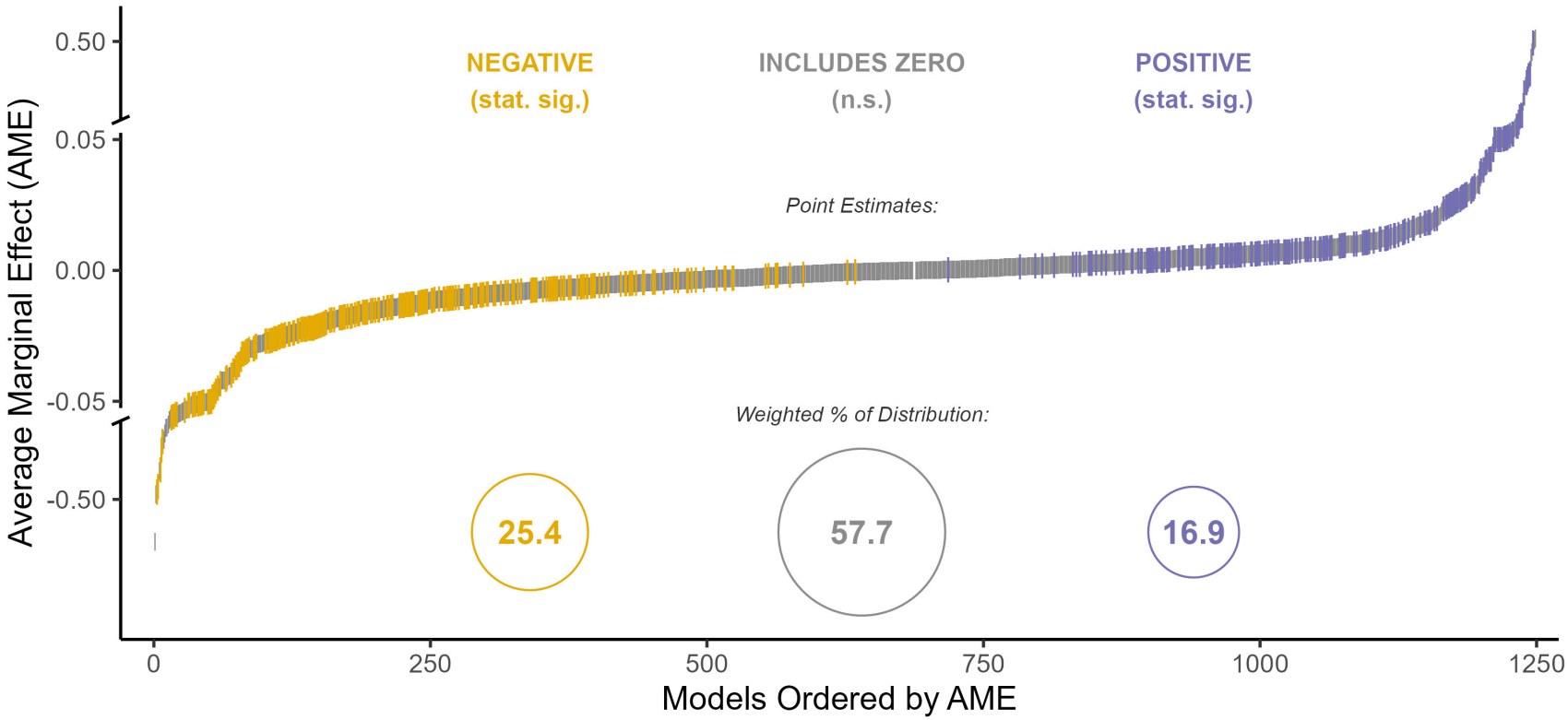
Edited by Douglas Massey, Princeton University, Princeton, NJ; received March 6, 2022; accepted August 22, 2022

**October 28, 2022** | 119 (44) e2203150119 | https://doi.org/10.1073/pnas.2203150119

THIS ARTICLE HAS BEEN UPDATED

THIS ARTICLE HAS BEEN CORRECTED + VIEW RELATED CONTENT +

65,932 | 58 PDF/EPUB


NEGATIVE
(stat. sig.)
INCLUDES ZERO
(n.s.)
POSITIVE
(stat. sig.)
Point Estimates:
Weighted % of Distribution:
25.4
57.7
16.9
Average Marginal Effect (AME)
0.50
0.05
0.00
-0.05
-0.50
0
250
500
750
1000
1250
Models Ordered by AME

# The problem: Types of Model Uncertainty

- **Parameter Uncertainty**:
  - Variability due to estimation from limited data.
- **Structural Uncertainty**:
  - Inadequate model to capture system complexities fully.
  - Simplifications that may oversimplify or miss critical aspects.
- **Data Uncertainty**:
  - Errors in data collection or measurement.
  - Outliers affecting model stability or accuracy.
- **Model Formulation Uncertainty**:
  - Choice of mathematical model structure.
  - Assumptions about relationships between variables.

Some types of uncertainty can be handled by researchers, some can't.

# Let's consider a simple example

- Imagine we want to know the effect of one variable $x_1$, on an output variable $y$.
- Example: $x_1$ is educational attainment, and $y$ is income.
- That is to say: our 'estimand' is $\beta_1$ in the following regression:

$$y_i = \alpha + \beta_1 x_{i,1} + \varepsilon_i \tag{1}$$

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | R | R-squared: | 0.033 |
| Model: | OLS | Adj. R-squared: | -0.011 |
| Method: | Least Squares | F-statistic: | 0.9856 |
| Date: | Sun, 12 Oct 2025 | Prob (F-statistic): | 0.381 |
| Time: | 05:05:32 | Log-Likelihood: | -237.19 |
| No. Observations: | 47 | AIC: | 480.4 |
| Df Residuals: | 44 | BIC: | 485.9 |
| Df Model: | 2 | | |
| Covariance Type: | HC1 | | |

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 111.2518 | 46.834 | 2.375 | 0.018 | 19.459 | 203.044 |
| Inequality | -0.1997 | 0.172 | -1.160 | 0.246 | -0.537 | 0.138 |
| Age | 0.1299 | 0.449 | 0.289 | 0.773 | -0.751 | 1.011 |

| | | | |
|---|---|---|---|
| Omnibus: | 8.799 | Durbin-Watson: | 2.457 |
| Prob(Omnibus): | 0.012 | Jarque-Bera (JB): | 8.002 |
| Skew: | 0.956 | Prob(JB): | 0.0183 |
| Kurtosis: | 3.655 | Cond. No. | 2.79e+03 |

But don't we also need to correct for things like age?

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | R | R-squared: | 0.033 |
| Model: | OLS | Adj. R-squared: | -0.011 |
| Method: | Least Squares | F-statistic: | 0.9856 |
| Date: | Sun, 12 Oct 2025 | Prob (F-statistic): | 0.381 |
| Time: | 05:05:32 | Log-Likelihood: | -237.19 |
| No. Observations: | 47 | AIC: | 480.4 |
| Df Residuals: | 44 | BIC: | 485.9 |
| Df Model: | 2 | | |
| Covariance Type: | HC1 | | |

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 111.2518 | 46.834 | 2.375 | 0.018 | 19.459 | 203.044 |
| Inequality | -0.1997 | 0.172 | -1.160 | 0.246 | -0.537 | 0.138 |
| Age | 0.1299 | 0.449 | 0.289 | 0.773 | -0.751 | 1.011 |

| | | | |
|---|---|---|---|
| Omnibus: | 8.799 | Durbin-Watson: | 2.457 |
| Prob(Omnibus): | 0.012 | Jarque-Bera (JB): | 8.002 |
| Skew: | 0.956 | Prob(JB): | 0.0183 |
| Kurtosis: | 3.655 | Cond. No. | 2.79e+03 |

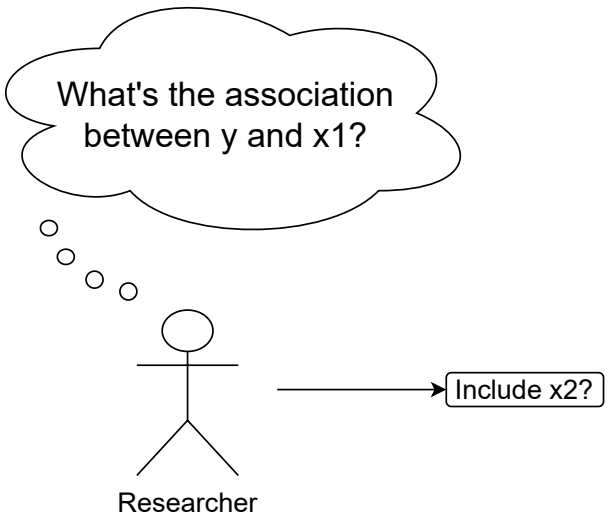
What about if we build out a fuller model, in order to help this estimation satisfy the Gauss-Markov properties?

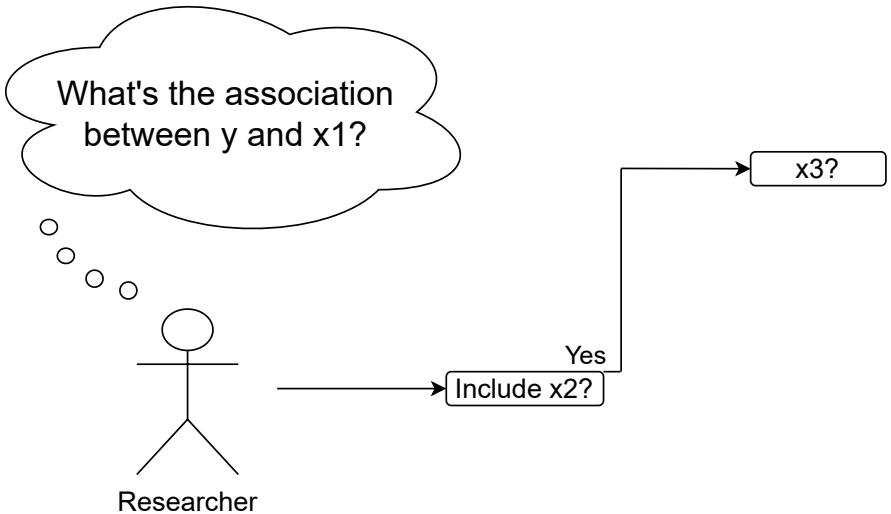
OLS Regression Results

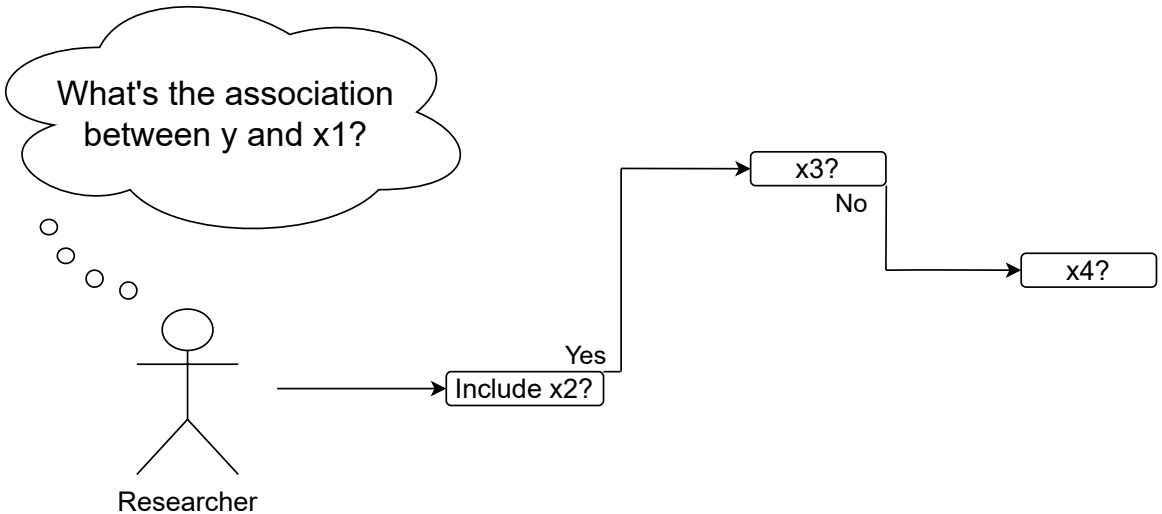
| | | | |
|---|---|---|---|
| Dep. Variable: | R | R-squared: | 0.522 |
| Model: | OLS | Adj. R-squared: | 0.436 |
| Method: | Least Squares | F-statistic: | 7.276 |
| Date: | Sun, 12 Oct 2025 | Prob (F-statistic): | 1.39e-05 |
| Time: | 05:06:00 | Log-Likelihood: | -220.64 |
| No. Observations: | 47 | AIC: | 457.3 |
| Df Residuals: | 39 | BIC: | 472.1 |
| Df Model: | 7 | | |
| Covariance Type: | HC1 | | |

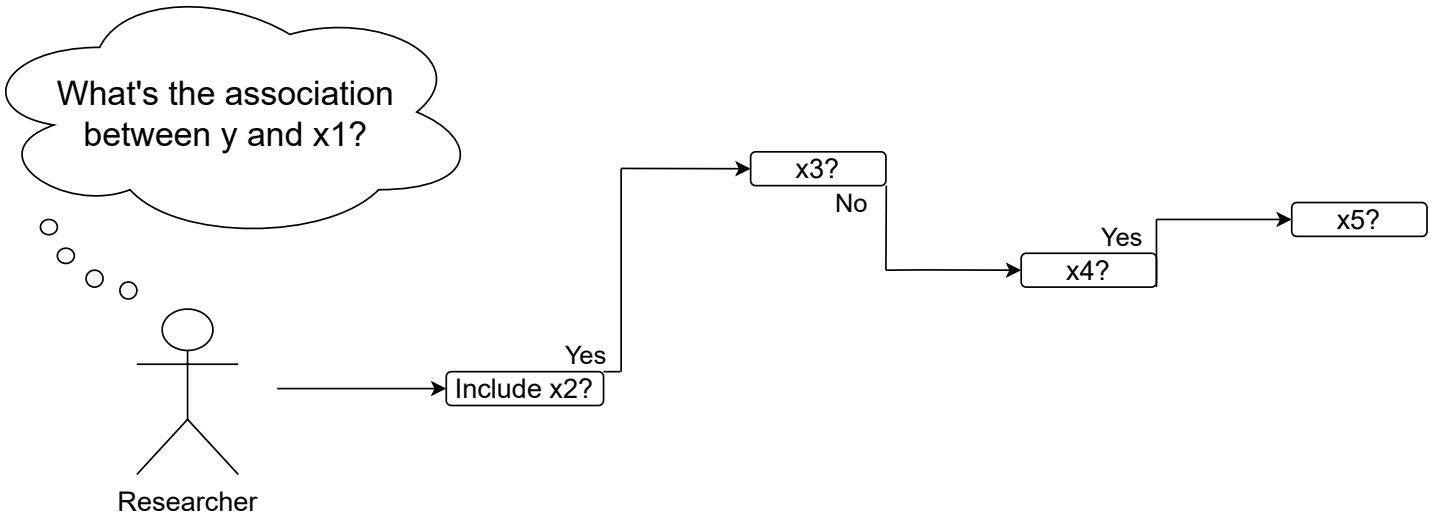
| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | -764.2379 | 223.962 | -3.412 | 0.001 | -1203.196 | -325.280 |
| Inequality | 0.8445 | 0.264 | 3.196 | 0.001 | 0.327 | 1.363 |
| Age | 0.9750 | 0.578 | 1.688 | 0.091 | -0.157 | 2.107 |
| Ed | 0.9510 | 0.793 | 1.199 | 0.231 | -0.604 | 2.506 |
| Unemployment | -0.0593 | 0.302 | -0.196 | 0.844 | -0.652 | 0.533 |
| Males | 0.2211 | 0.241 | 0.918 | 0.359 | -0.251 | 0.693 |
| N | 0.2714 | 0.117 | 2.326 | 0.020 | 0.043 | 0.500 |
| Wealth | 0.4449 | 0.125 | 3.559 | 0.000 | 0.200 | 0.690 |

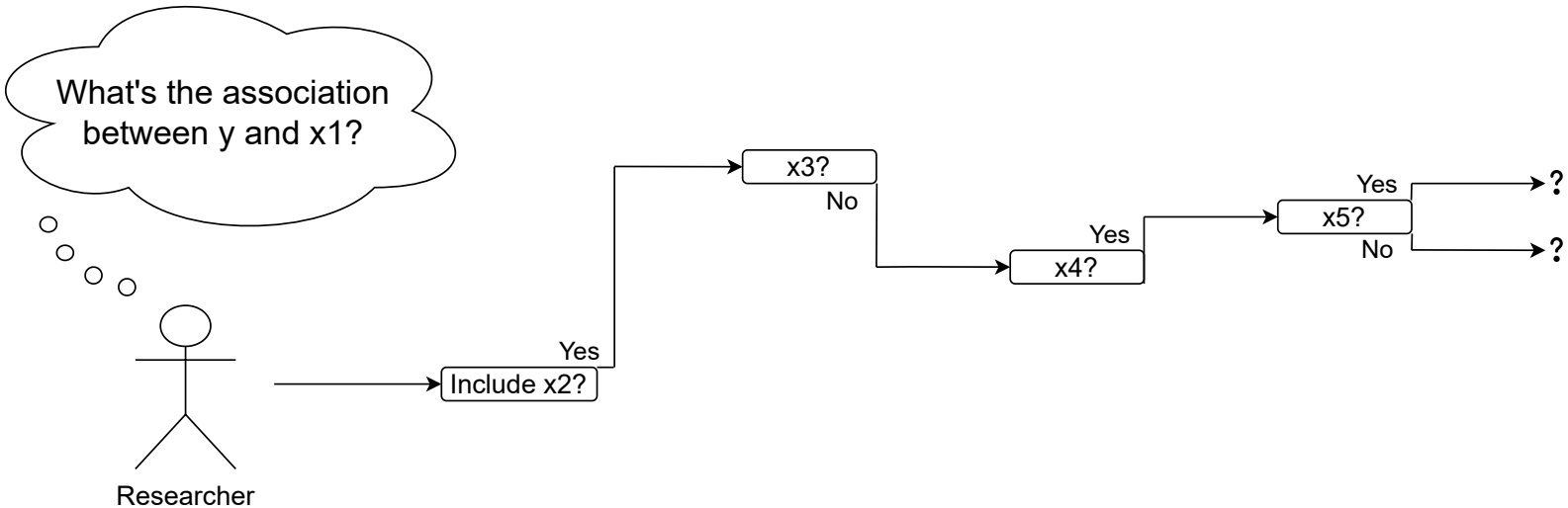
| | | | |
|---|---|---|---|
| Omnibus: | 11.181 | Durbin-Watson: | 1.794 |
| Prob(Omnibus): | 0.004 | Jarque-Bera (JB): | 10.897 |
| Skew: | 1.061 | Prob(JB): | 0.00430 |
| Kurtosis: | 4.031 | Cond. No. | 5.08e+04 |

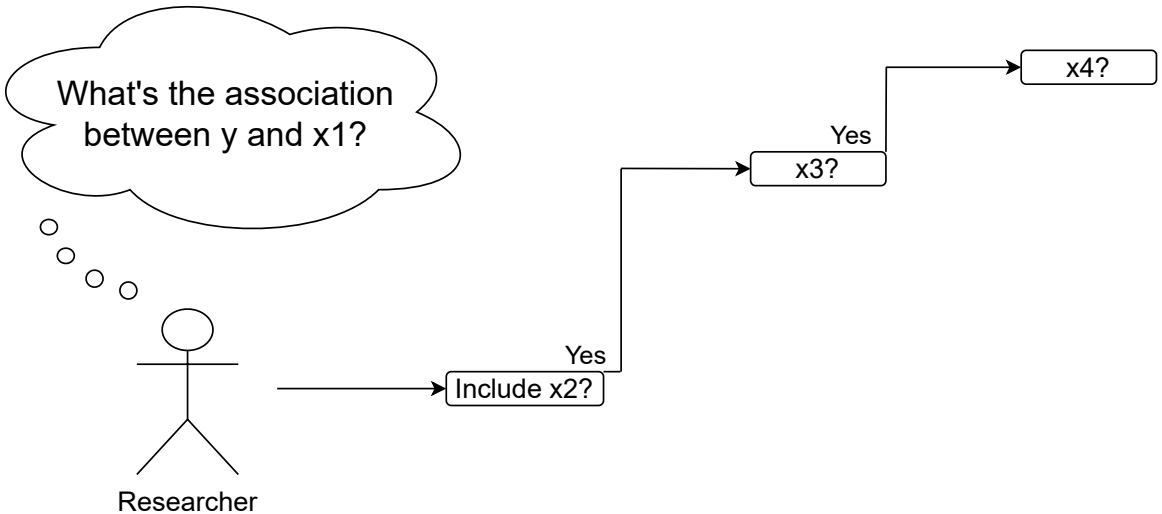

What's the association between y and x1?
Include x2?
Researcher

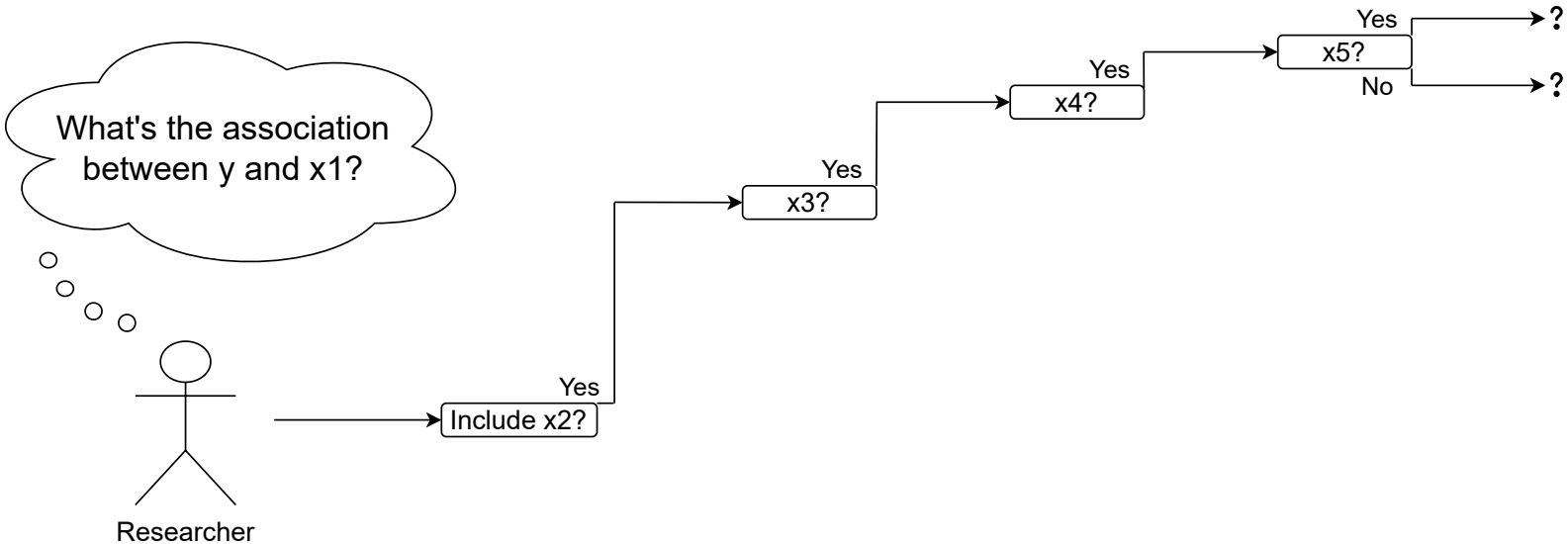

What's the association between y and x1?
x3?
Yes
Include x2?
Researcher

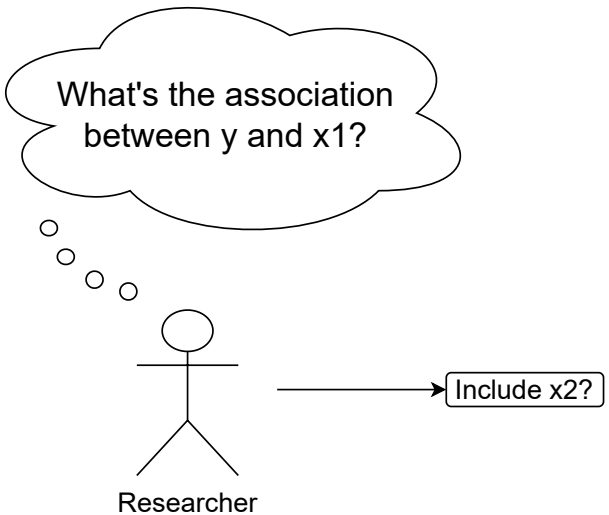

What's the association between y and x1?
Researcher
Include x2?
Yes
x3?
No
x4?

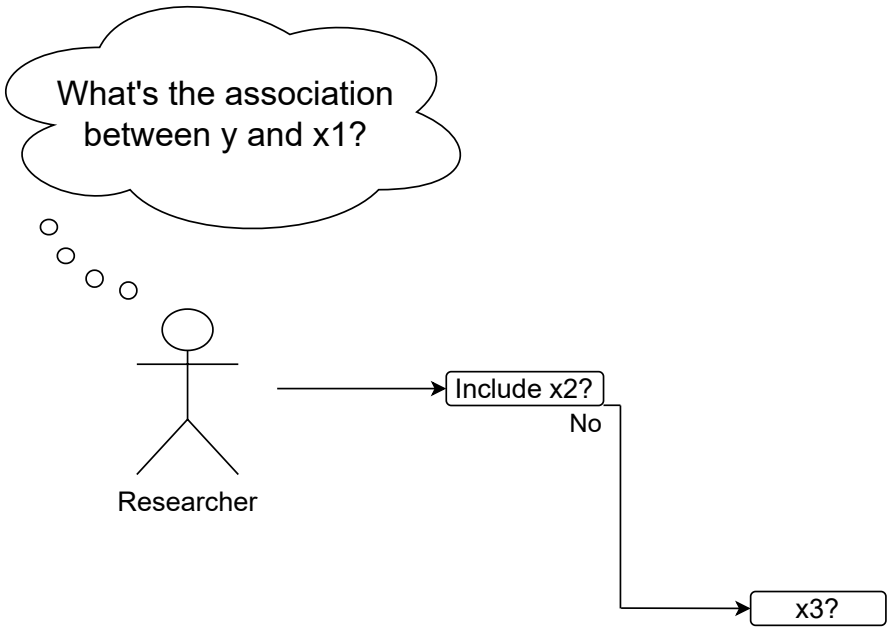

What's the association between y and x1?
Researcher
Include x2?
Yes
x3?
No
x4?
Yes
x5?

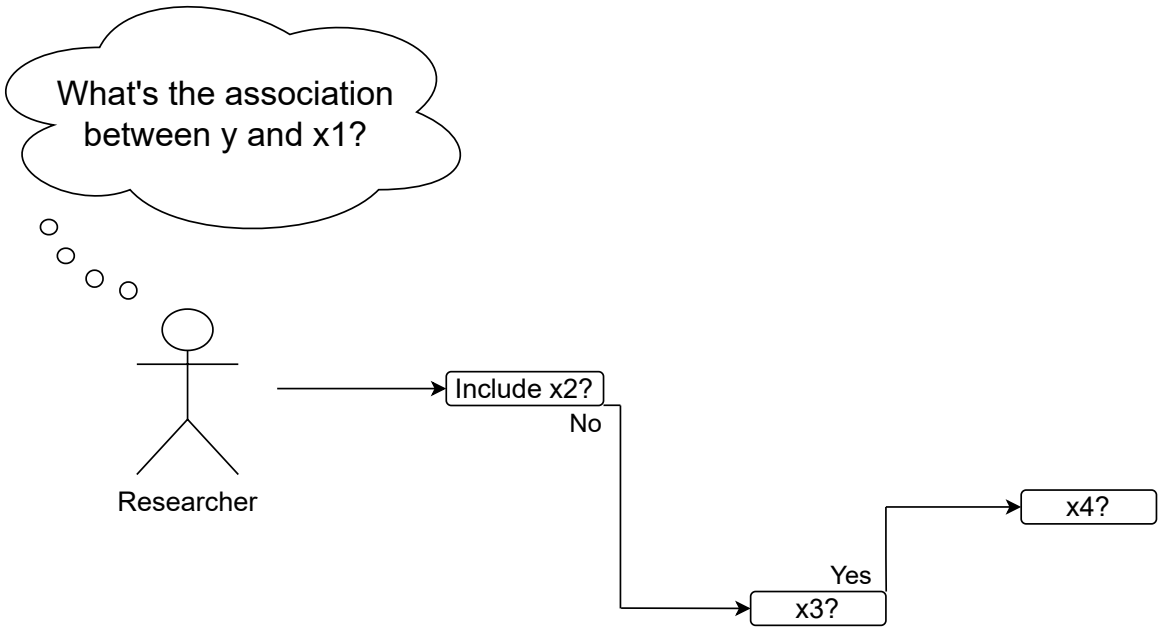

What's the association between y and x1?
Researcher
Include x2?
Yes
x3?
No
x4?
Yes
x5?
Yes
No
?
?

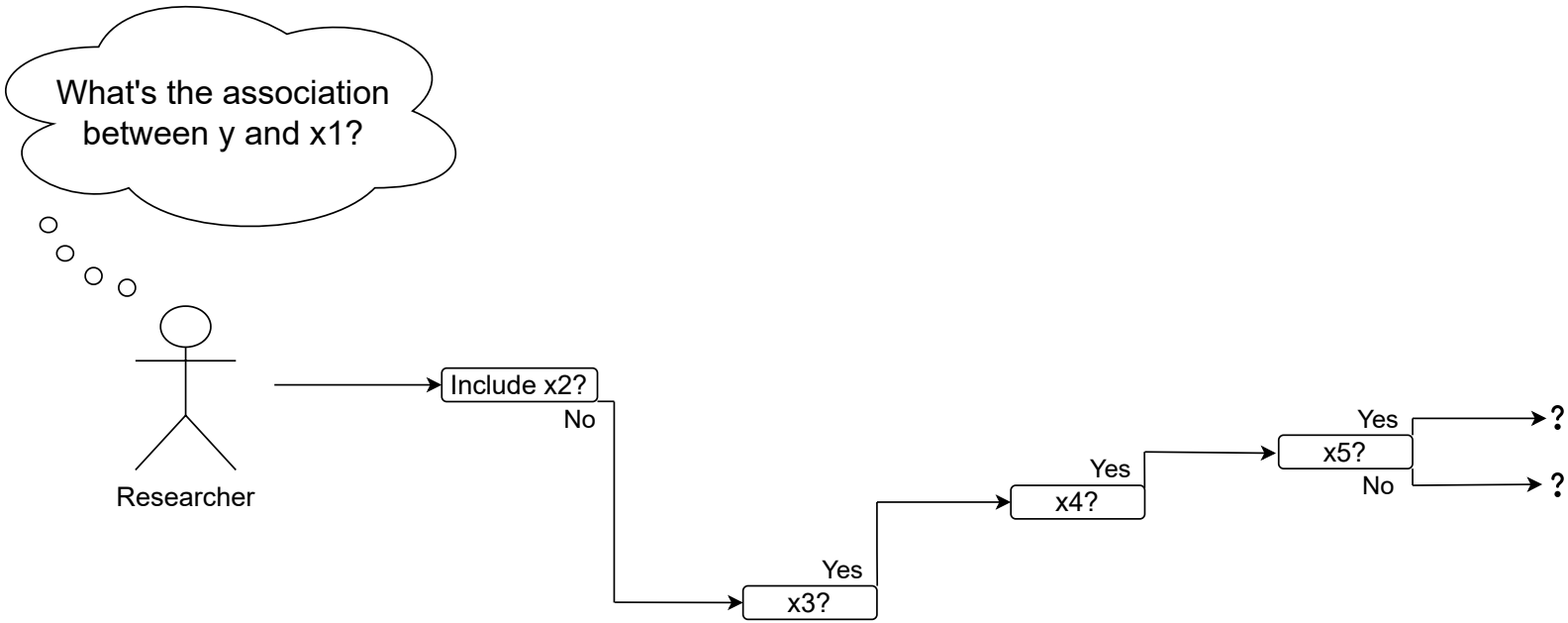

What's the association between y and x1?
Researcher
Include x2?
Yes
x3?
Yes
x4?

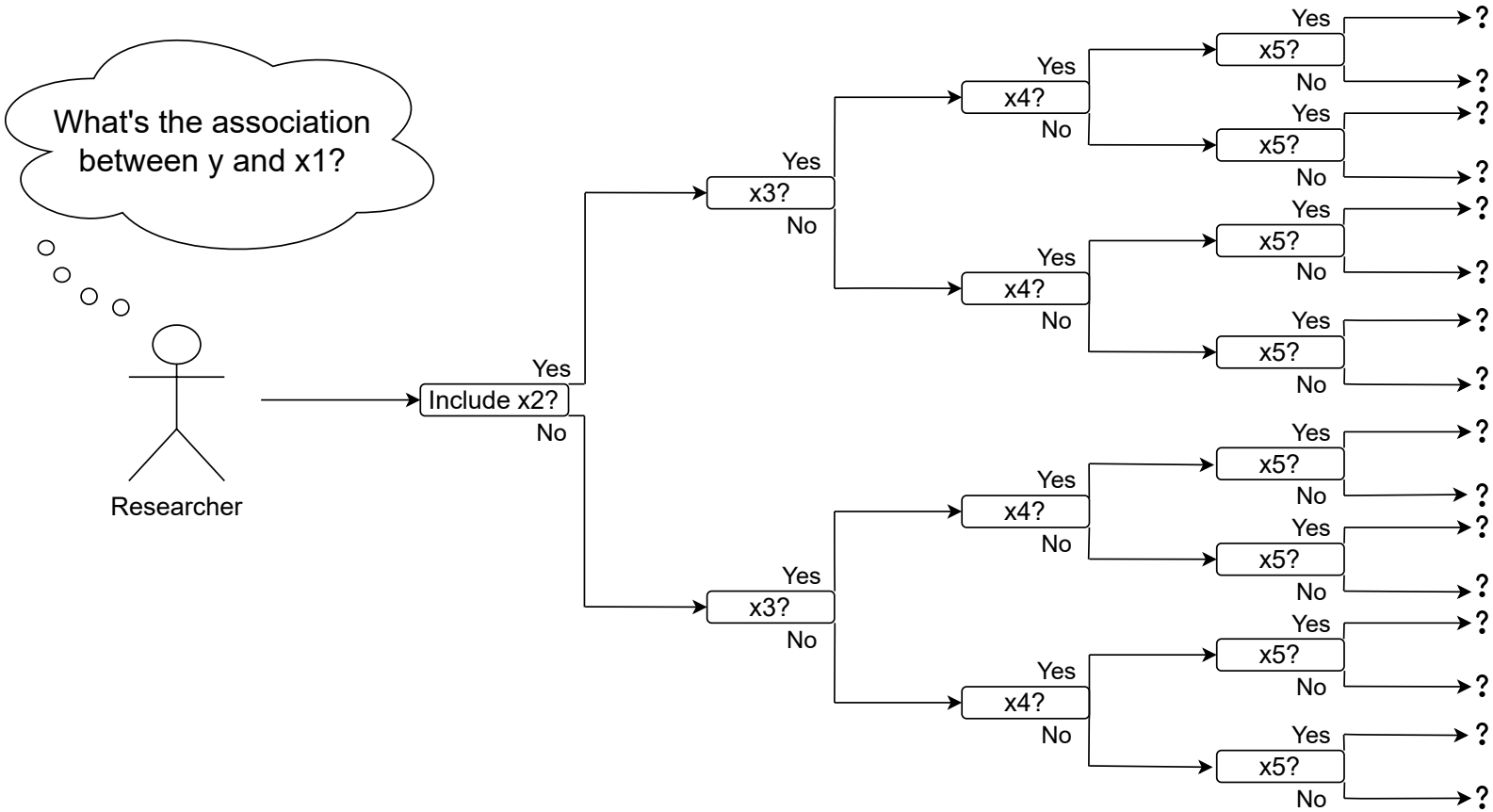

What's the association between y and x1?
Researcher
Include x2?
Yes
x3?
Yes
x4?
Yes
x5?
Yes
No
?
?


What's the association between y and x1?
Include x2?
Researcher


What's the association between y and x1?
Researcher
Include x2?
No
x3?


What's the association between y and x1?
Researcher
Include x2?
No
x4?
Yes
x3?


What's the association between y and x1?
Researcher
Include x2?
No
Yes
x3?
Yes
x4?
Yes
x5?
No
?
?


What's the association between y and x1?
Researcher
Include x2?
Yes
No
x3?
x4?
x5?
?

**This type of problem has enormous implications for p-hacking, HARKing, and the validity of the scientific record.**

**In their new book, Young and Cumberworth (2025) document the absence of systematic robustness checks in a huge number of recent and high profile Sociology articles.**

# Introducing the 'Feasible Model Space'

- From this simple choice of four auxiliary control variables, we have 16 possible models.
  - In general, this is $2^{n+1}$.
  - The '$+1$' comes from models with or without a constant.
  - With a choice of eight possible controls: $2^9 = 512$ models.
  - This can quickly get large.
- In an age of high performance computing: no problem!
- The main challenge: systematically reporting all these results.

# Introducing the 'Feasible Model Space'

- What's critical here is that we consider only a 'feasible model space'.
- We should **not** be including in this space noisy variables.
- This is where the role of Sociological Theory becomes critical.
- This allows us to include potential variables from theory.
- Systematic inclusion/omission allows holistic tests of theory.
- This does not prohibit us from having *one* preferred model.
  - This could be the one that is classically built in isolation.
  - But, now we can compare it with all others.

nature human behaviour



Resource | Published: 27 July 2020

# Specification curve analysis

Uri Simonsohn, Joseph P. Simmons & Leif D. Nelson

*Nature Human Behaviour* **4**, 1208–1214 (2020) | Cite this article

**8225** Accesses | **237** Citations | **68** Altmetric | Metrics

A Publisher Correction to this article was published on 09 October 2020

This article has been updated

Sociological Methods & Research

Impact Factor: **6.5** / 5-Year Impact Factor: **7.1**

Available access | Research article | First published online October 23, 2015 | Request permissions

# Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis

Cristobal Young and Katherine Holsteen View all authors and affiliations

Volume 46, Issue 1 | https://doi.org/10.1177/0049124115610347

Contents | PDF/EPUB | Cite | Share options | Information, rights and permissions | Metr

## Abstract

Model uncertainty is pervasive in social science. A key question is how robust empirical results are to sensible changes in model specification. We present a new approach and applied statistical software for computational multimodel analysis. Our approach proceeds in two steps: First, we estimate the modeling distribution of estimates across all combinations of possible controls as well as specified functional form issues, variable definitions, standard error calculations, and estimation commands. This allows analysts to present their core, preferred estimate in the context of a distribution of plausible estimates. Second, we develop a model influence analysis showing how each model ingredient affects the coefficient of interest. This shows which model assumptions, if any, are critical to obtaining an empirical result. We demonstrate the architecture and interpretation of multimodel analysis using data on the union wage premium, gender dynamics in mortgage lending, and tax flight migration among U.S. states. These illustrate how initial results can be strongly robust to alternative model specifications or remarkably dependent on a knife-edge specification.

# Limitations of Existing Tools

This is great!

- Not all researchers are, nor need to be, expert programmers.

However, these existing tools are limited.

- They don't include any real indication of model fit.
- They simply visualise distributions of the estimand.
- There is no out-of-sample validity.
- They place a uniform likelihood over all models.

We set ourselves the task:

**Can we create a reliable tool to conduct this/adjacent analysis?**

# Big Data in Sociology: Opportunities

Tools to deal with these large datasets more important than ever:

- Ubiquitous digital exhaust yields unprecedented sample sizes.
- Linkable administrative records enable rich longitudinal cohorts.
- Fine-grained spatiotemporal data uncover population heterogeneity.
- Natural experiments at scale strengthen causal identification.
- Real-time measurement enables adaptive, evidence-informed policy.

CHRIS ANDERSON SCIENCE JUN 23, 2008 12:00 PM

# The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

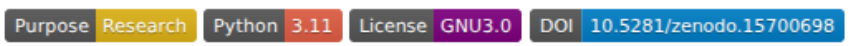
**Illustration: Marian Bantjes "All models are wrong, but some are useful." So proclaimed statistician George Box 30 years ago, and he was right. But what choice did we have? Only models, from cosmological equations to theories of human behavior, seemed to be able to consistently, if imperfectly, explain the world around us. Until now. Today companies [...]**

**We also want to build a reliable tool that respects theory.**


HIGH PERFORMANCE COMPUTING
MODEL
MODEL
MODEL
MODEL
SOCIOLOGY THEORY

README | Code of conduct | Contributing | GPL-3.0 license

# RobustiPy




Welcome to the home of `RobustiPy`, a Python library for the creation of a more robust and stable model space. RobustiPy does a large number of things, included but not limited to: high dimensional visualisation, Bayesian Model Averaging, bootstrapped resampling, (in- and)out-of-sample model evaluation, model selection via Information Criterion, explainable AI (via SHAP), and joint inference tests (as per Simonsohn et al. 2019).

Full documentation is available on Read the Docs. The first release is indexed onto Zenodo here.

`RobustiPy` performs Multiversal/Specification Curve Analysis which attempts to compute most or all reasonable specifications of a statistical model, understanding a specification as a single attempt to create an estimand of interest, whether through a particular choice of covariates, hyperparameters, data cleaning decisions, and so forth.

# Introducing RobustiPy: An efficient next generation multiversal library with model selection, averaging, resampling, and explainable AI

Daniel Valdenegro Ibarra[1,3], Jiani Yan[1,2,3,4], Duiyi Dai[1,3], and Charles Rahal[1,3,5]

[1]Leverhulme Centre for Demographic Science, University of Oxford
[2]Department of Sociology, University of Oxford
[3]ESRC Centre for Care, University of Oxford
[4]Wolfson College, University of Oxford
[5]Nuffield College, University of Oxford

October 8, 2025

**Abstract**

Scientific inference is often undermined by the vast but rarely explored "multiverse" of defensible modelling choices, which can generate results as variable as the phenomena under study. We introduce `RobustiPy`, an open-source Python library that systematizes multiverse analysis and model-uncertainty quantification at scale. `RobustiPy` unifies bootstrap-based inference, combinatorial specification search, model selection and averaging, joint-inference routines, and explainable AI methods within a modular, reproducible framework. Beyond exhaustive specification curves, it supports rigorous out-of-sample validation and quantifies the marginal contribution of each covariate. We demonstrate its utility across five simulation designs and ten empirical case studies spanning economics, sociology, psychology, and medicine—including a re-analysis of widely cited findings with documented discrepancies. Benchmarking on ∼672 million simulated regressions shows that `RobustiPy` delivers state-of-the-art computational efficiency while expanding transparency in empirical research. By standardizing and accelerating robustness analysis, `RobustiPy` transforms how researchers interrogate sensitivity across the analytical multiverse, offering a practical foundation for more reproducible and interpretable computational science.

**Keywords:** *Model Uncertainty*, *Statistical Software*, *Open Science*

# RobustiPy: The formal problem

Consider the key association between variables $Y$ and $X$, where a set of covariates $\mathbf{Z}$ can influence the relationship between the former as follows:

$$Y = F(X, \mathbf{Z}) + \varepsilon. \tag{2}$$

Let's now observe that:

- Usually $Y$ and $X$ are imprecisely defined latent variables.
- Likewise, the set of covariates $\mathbf{Z}$ can also be composed of imprecisely defined variables from which the number of them can be unknown and/or non-finite.
- Finally, $F()$ is an unknown data generating function.

# RobustiPy: The formal problem

Let's define the set of *reasonable* operationalisations of $Y$, $X$, $\mathbf{Z}$ and $F()$ as $\overleftrightarrow{Y}$, $\overleftrightarrow{X}$, $\overleftrightarrow{\mathbf{Z}}$ and $\overleftrightarrow{F()}$. Then, we have:

$$\overleftrightarrow{Y}_{k_Y} = \overleftrightarrow{F}_{k_f}(\overleftrightarrow{X}_{k_X}, \overleftrightarrow{\mathbf{Z}}_{k_\mathbf{Z}}) + \varepsilon. \tag{3}$$

- Eq. **??** corresponds to a single possible *specification* of the $\Pi$ set of all possible specifications.
- The total number of specifications can then be calculated as $2^N$ where $N$ is $n_{\overleftrightarrow{Y}} + n_{\overleftrightarrow{X}} + n_{\overleftrightarrow{\mathbf{Z}}} + n_{\overleftrightarrow{F()}}$.
- For example, an analysis with two functional forms, two target variables, two predictors and five covariates will yield a specification space $\Pi$ of $2^{10}$ or 1024 possible specifications.

# RobustiPy: The formal problem

We can also split the computation based on the operationalised variables in Eq. **??** by creating a 'reasonable specification space' for each: $\Pi_{\overleftrightarrow{Y}}$, $\Pi_{\overleftrightarrow{F()}}$, $\Pi_{\overleftrightarrow{X}}$, $\Pi_{\overleftrightarrow{\mathbf{Z}}}$. The whole specification space can be obtained again as follows:

$$\Pi = \Pi_{\overleftrightarrow{Y}} \times \Pi_{\overleftrightarrow{F()}} \times \Pi_{\overleftrightarrow{X}} \times \Pi_{\overleftrightarrow{\mathbf{Z}}} \tag{4}$$

- Any random and independently selected sample $\pi$ from $\Pi$ would lead to a reasonable approximation of $Y = F(X, \mathbf{Z})$.
- The main problem with current research practice is that the sample of $\Pi$ reported by the researchers is not random.

**The goal of RobustiPy is to automate this process to make it as easy as possible!**

# Functionality and Examples

- RobustiPy undertakes five novel types of common research design.
- It comes with:
  - Five simulated examples
  - Ten empirical examples
  - Full time profiling

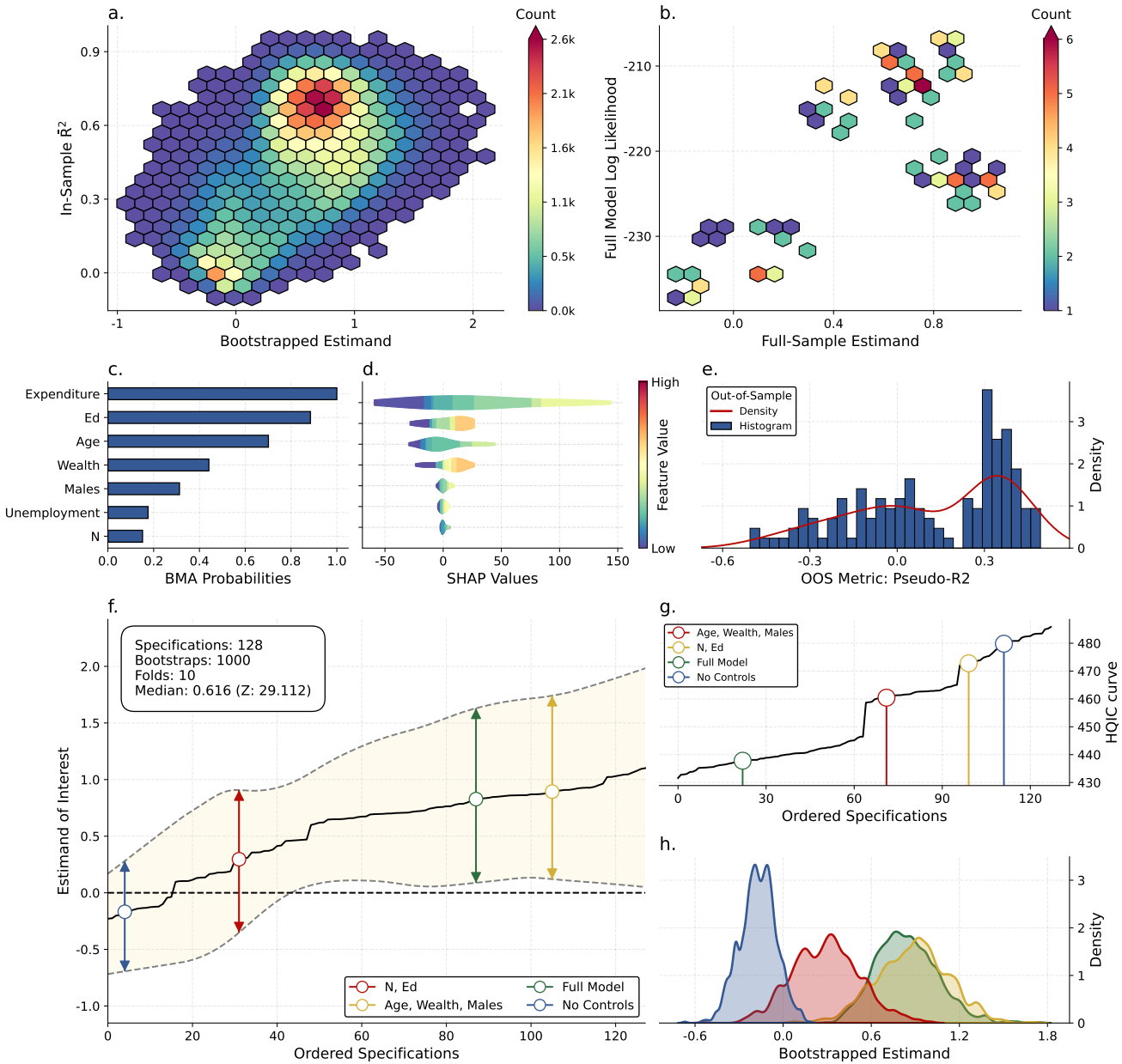
# Vanilla Computation

- Baseline case: single outcome $Y$, single predictor $X$, and varying controls $Z$.
- Data generating process:

$$Y = \alpha + \beta X + \gamma Z + \varepsilon$$

- Specification space reduces to $\Pi_{\overleftrightarrow{Z}}$ only.
- Estimation: OLS with $k$-fold CV and bootstrap resampling.
- Outputs: distribution of $\hat{\beta}$ across $2^{|Z|}$ models, with null vs full model.
- A link to the first empirical example can be found **here**!

==========
1. Model Summary
==========
Model: OLS Robust
Inference Tests: Yes
Dependent variable: R
Independent variable: Inequality
Number of possible controls: 7
...
==========
2. Model Robustness Metrics
==========
2.1 Inference Metrics
==========
Median $\beta$ (specs, no resampling): 0.6976 (null-calibrated p: 0)
Min Adj-$R^2$: -0.0310, Specs: [‘Age’, ‘Unemployment’]
...
Share $\beta$ <0 & significant (specs, no resampling): 0.0000
Share $\beta$ <0 & significant (bootstraps $\times$ specs): 0.0163 | null-calibrated p: 0.001998
Stouffer's Z = 29.1123 (two-sided p = 2.51e-186)
==========
2.2 In-Sample Metrics (Full Sample)
==========
Min AIC: 507.5432, Specs: ['Unemployment', 'Expenditure', 'N', 'Wealth', 'Males', 'Age', 'Ed']
Max Adj-$R^2$: 0.6922, Specs: ['Expenditure', 'Wealth', 'Males', 'Age', 'Ed']
Min Adj-$R^2$: -0.0310, Specs: ['Age', 'Unemployment']
==========
2.3 Out-Of-Sample Metrics (pseudo-r2 averaged across folds)
==========
Max Average: 0.4925, Specs: ['Age', 'Unemployment', 'Expenditure', 'Ed']
Min Average: -0.5057, Specs: ['Age', 'Unemployment', 'Ed', 'Males']
Mean Average: 0.12
Median Average: 0.2026


a.
In-Sample $\bar{R}^2$
Bootstrapped Estimand
Count
b.
Full Model Log Likelihood
Full-Sample Estimand
Count
c.
Expenditure
Ed
Age
Wealth
Males
Unemployment
N
BMA Probabilities
d.
SHAP Values
High
Feature Value
Low
e.
Out-of-Sample
Density
Histogram
OOS Metric: Pseudo-R2
Density
f.
Specifications: 128
Bootstraps: 1000
Folds: 10
Median: 0.616 (Z: 29.112)
Estimand of Interest
Ordered Specifications
N, Ed
Age, Wealth, Males
Full Model
No Controls
g.
Age, Wealth, Males
N, Ed
Full Model
No Controls
HQIC curve
Ordered Specifications
h.
Density
Bootstrapped Estimand

**Importantly, note the opportunity to highlight individual specifications!**

**These could/(should?) be ones driven by theory!**

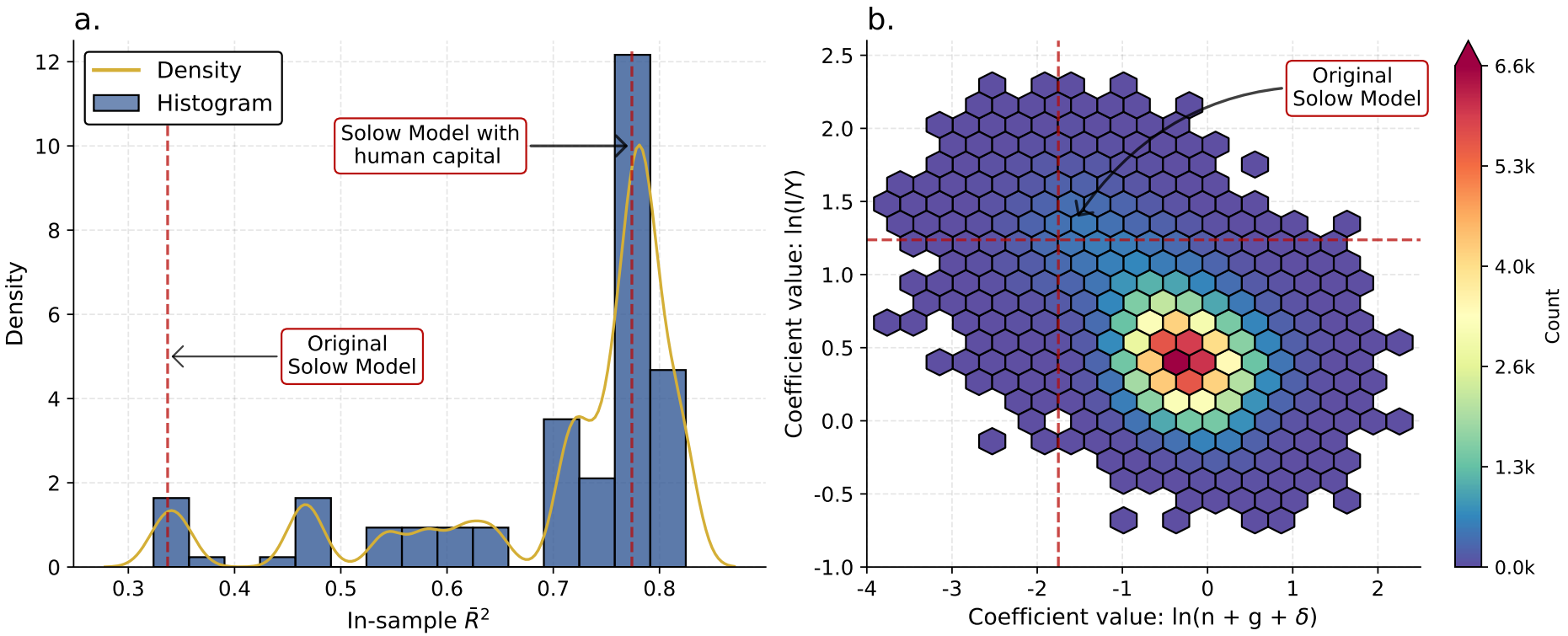
**The 'feasible model space' is also determined by theory!**

# Array of Never-Changing Variables

- Some predictors are fixed (always included in $\overleftrightarrow{X}$).
- Reduces computational cost by shrinking the specification space.
- Empirical motivation: theoretically indispensable covariates.
- Specification space:

$$\Pi = \Pi_{\overleftrightarrow{Z}} \quad \text{with fixed } X_f$$

- Results: interpretable estimates with reduced variance and runtime.
- A link to the second empirical example can be found **here**!

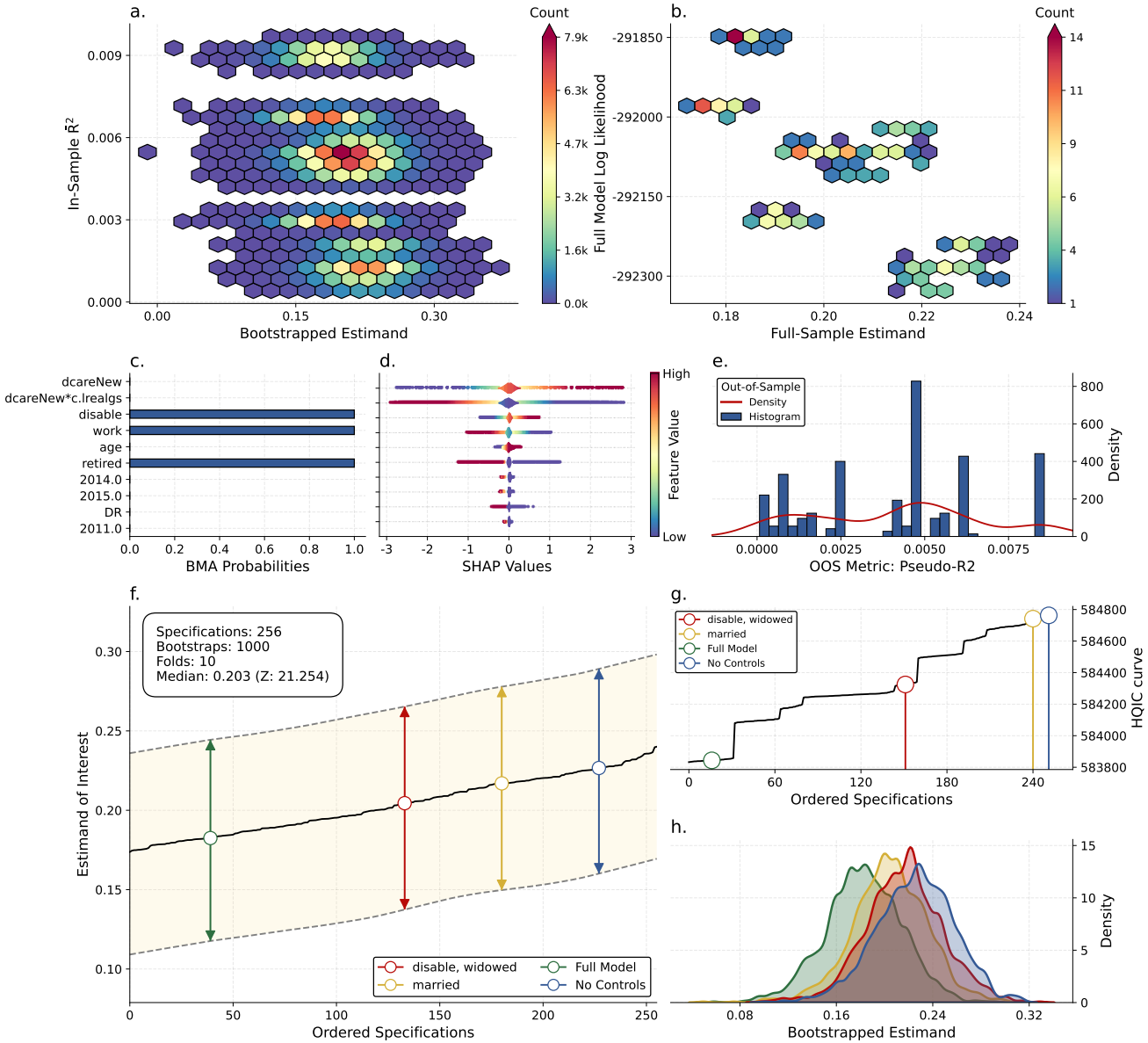

a.
Density
Histogram
Solow Model with human capital
Original Solow Model
Density
In-sample $\bar{R}^2$
b.
Original Solow Model
Coefficient value: ln(I/Y)
Coefficient value: ln(n + g + δ)
Count
6.6k
5.3k
4.0k
2.6k
1.3k
0.0k

# Fixed Effects OLS

- Designed for panel/longitudinal data with grouping variable $g$.
- Within-transformation removes group-specific heterogeneity:

$$Y_{it} - \bar{Y}_i = \beta(X_{it} - \bar{X}_i) + \varepsilon_{it}$$

- Implemented by demeaning variables by $g$.
- $\Pi_{\underset{Z}{\Leftrightarrow}}$ still varied, but estimation is within-groups.
- Useful for causal inference in repeated-measures data.
- A link to the third empirical example can be found **here**!


a.
In-Sample $\bar{R}^2$
Bootstrapped Estimand
Count
Full Model Log Likelihood
b.
Full-Sample Estimand
c.
dcareNew
dcareNew*c.lrealgs
disable
work
age
retired
2014.0
2015.0
DR
2011.0
BMA Probabilities
d.
SHAP Values
Feature Value
High
Low
e.
Out-of-Sample
Density
Histogram
OOS Metric: Pseudo-R2
f.
Specifications: 256
Bootstraps: 1000
Folds: 10
Median: 0.203 (Z: 21.254)
Estimand of Interest
Ordered Specifications
disable, widowed
married
Full Model
No Controls
g.
HQIC curve
Ordered Specifications
h.
Bootstrapped Estimand
Density

# Binary Dependents

- Logistic regression estimator for binary $Y \in \{0, 1\}$.
- Model:

$$\Pr(Y = 1 \mid X, Z) = \frac{1}{1 + \exp(-(\beta X + \gamma Z))}$$

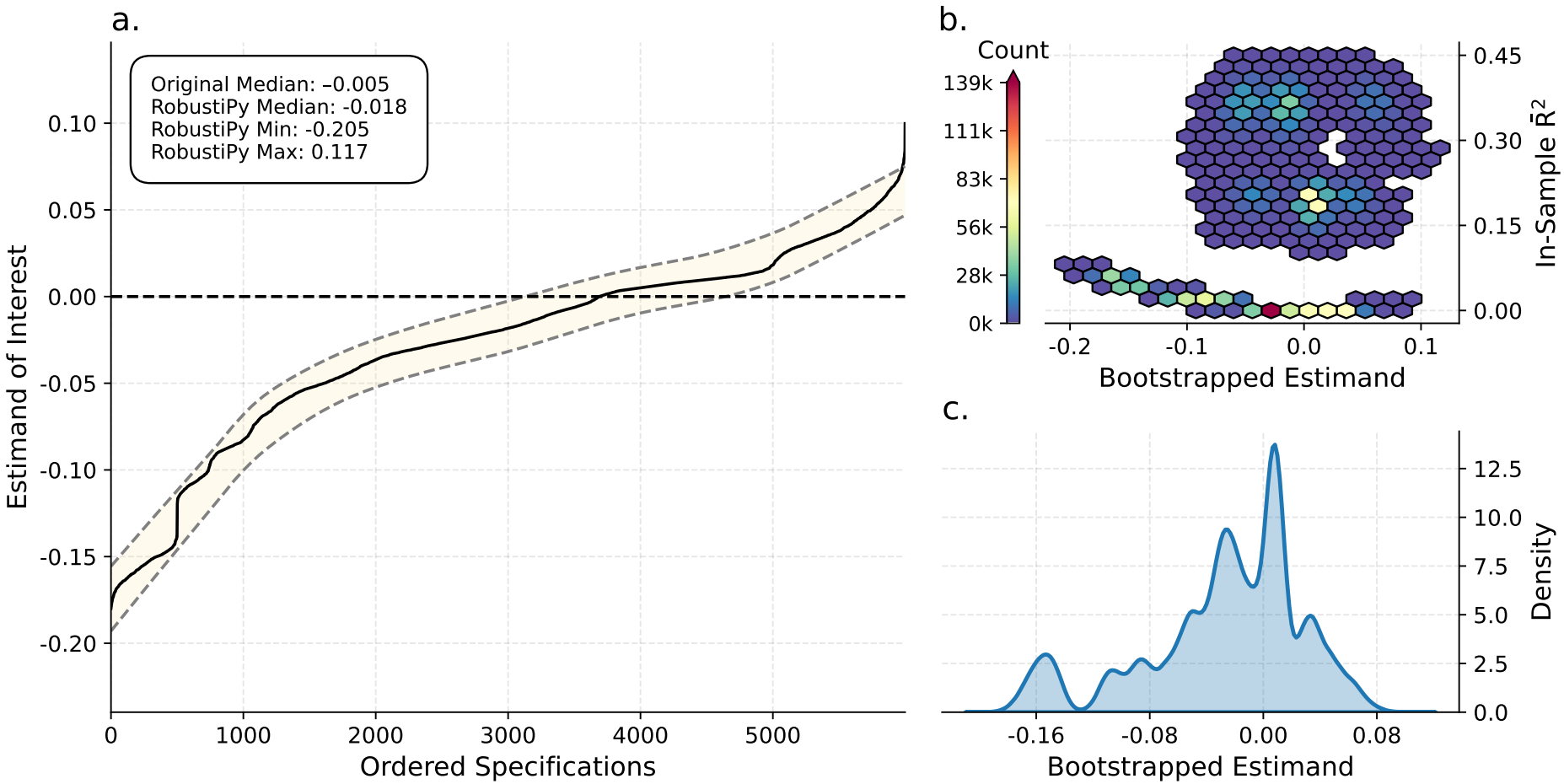
- $\Pi_{\underset{Z}{\leftrightarrow}}$ explored as before, with log-likelihood fit.
- Supports odds ratio interpretation of $\hat{\beta}$.
- Alternative: linear probability model with OLSRobust.

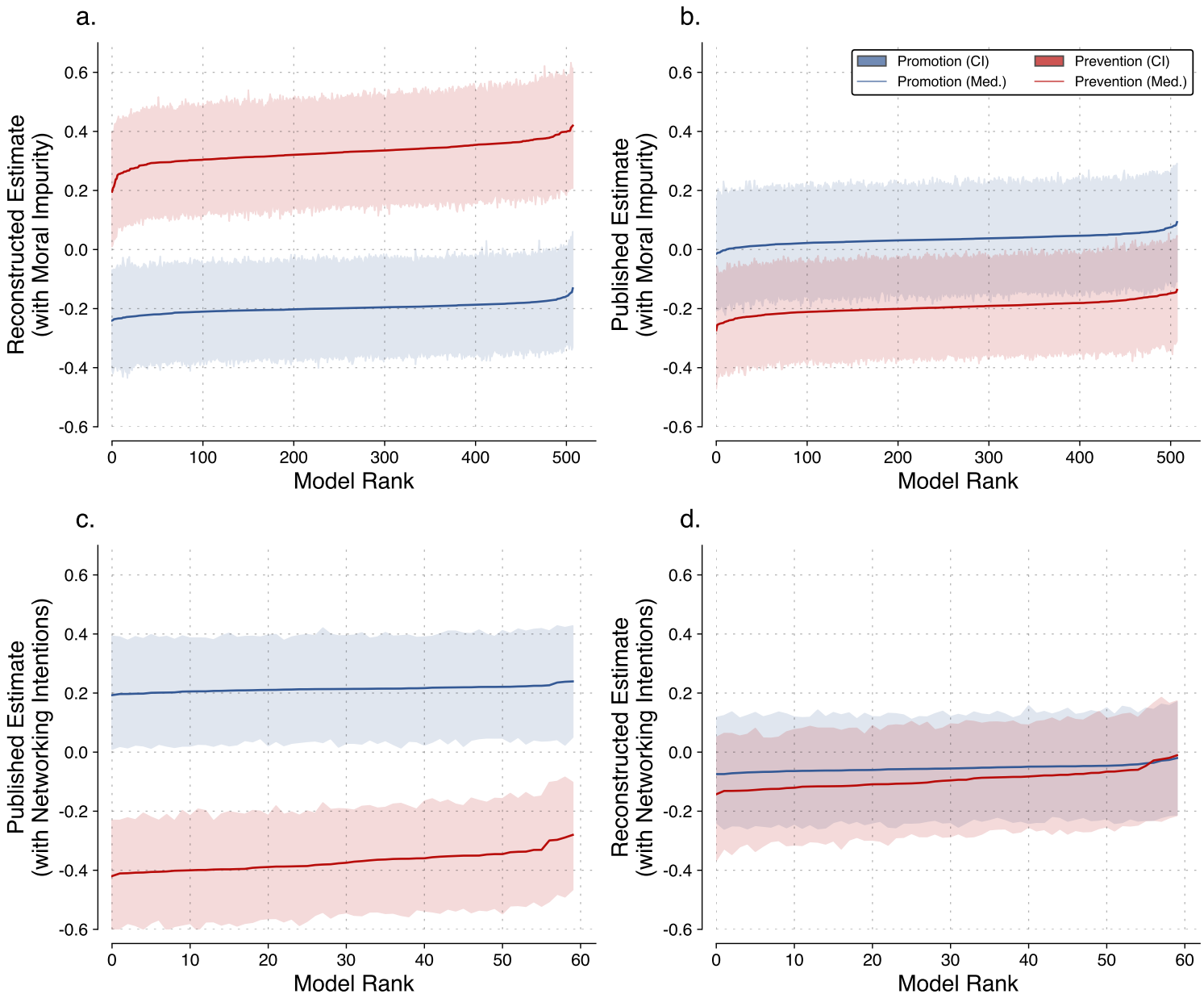
# Multiple Dependents

- Allows multiple operationalisations of $Y$.
- Construct composites by averaging $z$-scored outcomes:

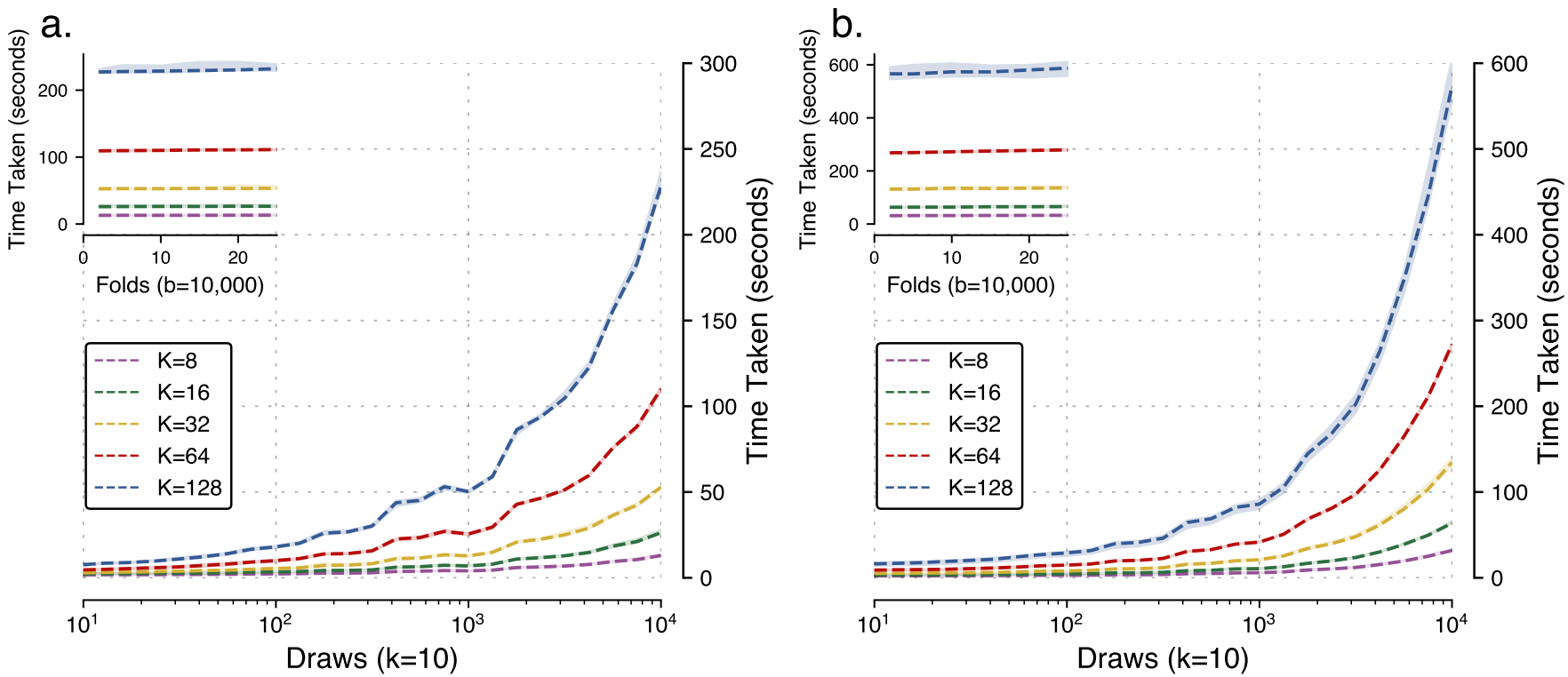
$$Y_{i,S}^{\text{comp}} = \frac{1}{|S|} \sum_{j \in S} \tilde{Y}_i^{(j)}$$

- Specification space enlarges: $\Pi_{\overleftrightarrow{Y}} \times \Pi_{\overleftrightarrow{Z}}$.
- Provides robustness to measurement uncertainty in the outcome.
- Outputs: specification curves across all dependent definitions/composites.
- Links to fourth and fifth empirical examples found **here** and **here**.


a.
Original Median: –0.005
RobustiPy Median: -0.018
RobustiPy Min: -0.205
RobustiPy Max: 0.117
Estimand of Interest
Ordered Specifications
b.
Count
139k
111k
83k
56k
28k
0k
In-Sample $\bar{R}^2$
Bootstrapped Estimand
c.
Density
Bootstrapped Estimand


a.
Reconstructed Estimate
(with Moral Impurity)
Model Rank
b.
Published Estimate
(with Moral Impurity)
Promotion (CI)
Promotion (Med.)
Prevention (CI)
Prevention (Med.)
Model Rank
c.
Published Estimate
(with Networking Intentions)
Model Rank
d.
Reconstructed Estimate
(with Networking Intentions)
Model Rank

# Time Profiling



# Discussion Questions

1. What types of uncertainty does RobustiPy *not* capture?
2. How does the Gauss-Markov theorem feel about this approach?
   - How about Angrist and Pischke?
3. What other types of major challenges researchers face?
4. What is the current status of teaching model uncertainty?
5. How trusted are academic 'experts' right now in general?
6. Is RobustiPy a tool that can be used for evil, as well as good?
7. How might quantum computing benefit this space?
8. Optimal weighting of specifications for policy decisions?

Sincere thanks to my **wonderful** co-authors!