Identification of and correction for publication bias

Isaiah Andrews    Maximilian Kasy

March 17, 2017
Introduction

- Fundamental requirement of science: replicability
- Different researchers should reach same conclusions
- Methodological conventions should ensure this (e.g., randomized experiments)
- Replicability often appears to fail, e.g.
  - Experimental economics (Camerer et al., 2016)
  - Experimental psychology (Open Science Collaboration, 2015)
  - Medicine (Ionnidias, 2005)
  - Cell Biology (Begley et al, 2012)
  - Neuroscience (Button et al, 2013)
Introduction

- Possible explanation: selective publication of results
- Due to:
  - Researcher decisions
  - Journal selectivity
- Possible selection criteria:
  - Statistically significant effects
  - Confirmation of prior beliefs
  - Novelty
- Consequences:
  - Conventional estimators are biased
  - Conventional inference does not control size
Introduction
Our contributions

1. Nonparametric **identification of selectivity** in the publication process, using
   a) Replication studies: Absent selectivity, original and replication estimates should be symmetrically distributed
   b) Meta-studies: Absent selectivity, distribution of estimates for small sample sizes should be noised-up version of distribution for larger sample sizes

2. **Corrected inference** when selectivity is known
   a) Median unbiased estimators
   b) Confidence sets with correct coverage
   c) Allow for nuisance parameters and multiple dimensions of selection
   d) Bayesian inference accounting for selection

3. **Applications** to
   a) Experimental economics
   b) Experimental psychology
   c) Effects of minimum wages on employment
   d) Effects of de-worming
Outline

1 Introduction
2 Setup
3 Identification
4 Bias-corrected inference
5 Applications
6 Conclusion
Assume there is a population of latent studies indexed by $i$

- True parameter value in study $i$ is $\Theta_i^*$
  - $\Theta_i^*$ drawn from some population $\Rightarrow$ empirical Bayes perspective
  - Different studies may recover different parameters

- Each study reports findings $X_i^*$
  - Distribution of $X_i^*$ given $\Theta_i^*$ known

- A given study may or may not be published
  - Determined by both researcher and journal: we don’t try to disentangle

- Probability of publication $P(D_i = 1|X_i^*, \Theta_i^*) = p(X_i^*)$

Published studies are indexed by $j$
Definition (General sampling process)

**Latent (unobserved) variables:** \((D_i, X^*_i, \Theta^*_i)\), jointly i.i.d. across \(i\)

\[
\begin{align*}
\Theta^*_i & \sim \mu \\
X^*_i | \Theta^*_i & \sim f_{X^*|\Theta^*}(x|\Theta^*_i) \\
D_i | X^*_i, \Theta^*_i & \sim \text{Ber}(p(X^*_i))
\end{align*}
\]

**Truncation:** We observe i.i.d. draws of \(X_j\), where

\[
l_j = \min\{i : D_i = 1, i > l_{j-1}\}
\]

\[
\Theta_j = \Theta^*_{l_j} \quad X_j = X^*_{l_j}
\]
Setup
Example: treatment effects

- Journal receives a stream of studies \( i = 1, 2, \ldots \)
- Each reporting experimental estimates \( X_i^* \) of treatment effects \( \Theta_i^* \)
- Distribution of \( \Theta_i^* \): \( \mu \)
- Suppose that \( X_i^* | \Theta_i^* \sim N(\Theta_i^*, 1) \)
- Publication probability: “significance testing,”

\[
p(X) = \begin{cases} 
0.1 & |X| < 1.96 \\
1 & |X| \geq 1.96
\end{cases}
\]

- Published studies: report estimate \( X_j \) of treatment effect \( \Theta_j \)
Setup

Example continued – Publication bias

- Left: median bias of $\hat{\theta}_j = X_j$
- Right: true coverage of conventional 95% confidence interval
Outline

1. Introduction
2. Setup
3. Identification
4. Bias-corrected inference
5. Applications
6. Conclusion
Identification
Identification of the selection mechanism $p(\cdot)$

- Key unknown object in model: publication probability $p(\cdot)$
- We propose two approaches for identification:
  1. Replication experiments:
      - replication estimate $X^r$ for the same parameter $\Theta$
      - selectivity operates only on $X$, but not on $X^r$
  2. Meta-studies:
      - Variation in $\sigma^*$, where $X^* \sim N(\Theta^*, \sigma^{*2})$
      - Assume $\sigma^*$ is (conditionally) independent of $\Theta^*$ across latent studies $i$
      - Standard assumption in the meta-studies literature; validated in our applications by comparison to replications

- Advantages:
  1. Replications: Very credible
  2. Meta-studies: Widely applicable
Identification

Intuition: identification using replication studies

- Left: no truncation
  \[ \Rightarrow \text{areas } A \text{ and } B \text{ have same probability} \]

- Right: \[ p(Z) = 0.1 + 0.9 \cdot 1(|Z| > 1.96) \]
  \[ \Rightarrow A \text{ more likely than } B \]
Identification
Approach 1: Replication studies

Definition (Replication sampling process)

- **Latent variables:** as before,
  \[ \Theta_i^* \sim \mu \]
  \[ X_i^* | \Theta_i^* \sim f_{X^*|\Theta^*}(x | \Theta_i^*) \]
  \[ D_i | X_i^*, \Theta_i^* \sim Ber(p(X_i^*)) \]

- **Additionally: replication draws,**
  \[ X_i^{*r} | X_i^*, D_i, \Theta_i^* \sim f_{X^*|\Theta^*}(x | \Theta_i^*) \]

- **Observability:** as before,
  \[ l_j = \min\{i : D_i = 1, i > l_{j-1}\} \]
  \[ \Theta_j = \Theta_{l_j} \]
  \[ (X_j, X_j^{*r}) = (X_{l_j}^*, X_{l_j}^{*r}) \]
Identification

Theorem (Identification using replication experiments)

Assume that the support of $f_{X_i^*, X_i^{*r}}$ is of the form $A \times A$ for some set $A$. Then $p(\cdot)$ is identified on $A$ up to scale.

Intuition of proof:

- Marginal density of $(X, X^r)$ is

$$f_{X, X^r}(x, x^r) = \frac{p(x)}{E[p(X_i^*)]} \int f_{X_i^*|\Theta^*}(x|\theta_i^*) f_{X_i^{*r}|\Theta^*}(x^r|\theta_i^*) d\mu(\theta_i^*).$$

- Thus, for all $a, b$, if $p(a) > 0$,

$$\frac{p(b)}{p(a)} = \frac{f_{X, X^r}(b, a)}{f_{X, X^r}(a, b)}.$$
Identification

Practical complication

- Replication experiments follow the same protocol
  ⇒ estimate same effect $\Theta$

- But often different sample size
  ⇒ different variance ⇒ symmetry breaks down

- Additionally: replication sample size often determined based on power calculations given initial estimate

- $p(\cdot)$ is still identified (up to scale):
  - Assume $X$ normally distributed
  - Intuition: Conditional on $X, \sigma$, (de-)convolve $X'$ with normal noise to get symmetry back
  - $\mu$ is identified as well
Identification

Intuition: identification using meta-studies

- Left: no truncation
  \[ \Rightarrow \text{dist for higher } \sigma \text{ noised up version of dist for lower } \sigma \]

- Right: \[ p(Z) = 0.1 + 0.9 \cdot 1(|Z| > 1.96) \]
  \[ \Rightarrow \text{“missing data” inside the cone} \]
Identification

Approach 2: meta-studies

Definition (Independent $\sigma$ sampling process)

\[
\begin{align*}
\sigma_i^* &\sim \mu_\sigma \\
\Theta_i^*|\sigma_i^* &\sim \mu_\Theta \\
X_i^*|\Theta_i^*, \sigma_i^* &\sim N(\Theta_i^*, \sigma_i^{*2}) \\
D_i|X_i^*, \Theta_i^*, \sigma_i^* &\sim Ber(p(X_i^*/\sigma_i^*))
\end{align*}
\]

We observe i.i.d. draws of $(X_j, \sigma_j)$, where

\[
l_j = \min\{i : D_i = 1, \ i > l_{j-1}\}
\]

\[
(X_j, \sigma_j) = (X_{l_j}^*, \sigma_{l_j}^*)
\]

Define $Z^* = \frac{X^*}{\sigma^*}$ and $Z = \frac{X}{\sigma}$
Identification

**Theorem (Nonparametric identification using variation in \( \sigma \))**

*Suppose that the support of \( \sigma \) contains a neighborhood of some point \( \sigma_0 \). Then \( p(\cdot) \) is identified up to scale.*

**Intuition of proof:**

- Conditional density of \( Z \) given \( \sigma \) is

\[
    f_{Z|\sigma}(z|\sigma) = \frac{p(z)}{E[p(Z^*)|\sigma]} \int \phi(z - \theta/\sigma) d\mu(\theta).
\]

- Thus

\[
    \frac{f_{Z|\sigma}(z|\sigma_2)}{f_{Z|\sigma}(z|\sigma_1)} = \frac{E[p(Z^*)|\sigma = \sigma_1]}{E[p(Z^*)|\sigma = \sigma_2]} \cdot \frac{\int \phi(z - \theta/\sigma_2) d\mu(\theta)}{\int \phi(z - \theta/\sigma_1) d\mu(\theta)}.
\]

- Recover \( \mu \) from right hand side, then recover \( p(\cdot) \) from first equation.
Outline

1. Introduction
2. Setup
3. Identification
4. Bias-corrected inference
5. Applications
6. Conclusion
Bias-corrected inference

- Once we know $p(\cdot)$, can correct inference for selection
- For simplicity, here assume $X, \Theta$ both 1-dimensional
- Density of published $X$ given $\Theta$:

$$f_{X|\Theta}(x|\theta) = \frac{p(x)}{E[p(X^*)|\Theta^* = \theta]} \cdot f_{X^*|\Theta^*}(x|\theta)$$

- Corresponding cumulative distribution function: $F_{X|\Theta}(x|\theta)$
Bias-corrected inference
Corrected frequentist estimators and confidence sets

- We are interested in bias, and the coverage of confidence sets
  - Condition on $\theta$: standard frequentist analysis
- Define $\hat{\theta}_\alpha (\mathbf{x})$ via
  $$ F_{X|\Theta}(x|\hat{\theta}_\alpha (\mathbf{x})) = \alpha $$
- Can show that
  $$ P\left(\hat{\theta}_\alpha (\mathbf{X}) \leq \theta | \theta\right) = \alpha \ \forall \theta $$
- Median-unbiased estimator: $\hat{\theta}_{\frac{1}{2}} (\mathbf{X})$ for $\theta$
- Equal-tailed level $1 - \alpha$ confidence interval:
  $$ \left[ \hat{\theta}_{\frac{\alpha}{2}} (\mathbf{X}), \hat{\theta}_{1 - \frac{\alpha}{2}} (\mathbf{X}) \right] $$
Bias-corrected inference

Example: treatment effects

Let us return to the treatment effect example discussed above.

Again assume \( X^* | \Theta^* \sim N(\Theta^*, 1) \) and

\[
p(X) = 0.1 + 0.9 \cdot 1(|X| > 1.96)
\]
Bias-corrected inference

Example continued – corrected confidence sets for $\beta_p = 0.1$
Outline

1. Introduction
2. Setup
3. Identification
4. Bias-corrected inference
5. Applications
6. Conclusion
Applications

Replications of Lab Experiments in Economics

- Camerer et al. (2016)
- Sample: all 18 between-subject laboratory experimental papers published in AER and QJE between 2011 and 2014
- Scatterplot next slide:
  - $Z = X / \sigma$: normalized initial estimate
  - $Z^r = X^r / \sigma$: replicate estimate
  - Initial estimates normalized to be positive
Applications

Economics Lab Experiments: Original and Replication Z Statistics
Applications

Economics Lab Experiments: Estimates of Selection model

- Model:

\[ \Theta^* \sim N(0, \tau^2) \]

\[ p(Z) \propto \begin{cases} \beta_p \mid Z \mid < 1.96 \\ 1 \mid Z \mid \geq 1.96 \end{cases} \]

- Estimates:

<table>
<thead>
<tr>
<th>\tau</th>
<th>\beta_p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.354</td>
<td>0.100</td>
</tr>
<tr>
<td>(0.750)</td>
<td>(0.091)</td>
</tr>
</tbody>
</table>

- Interpretation: insignificant (at the 5 % level) results about 10% as likely to be published as significant results
Applications

Economics Lab Experiments: Adjusted Estimates

Kuziemko et al. (QJE 2014)
Ambrus and Greiner (AER 2012)
Abeler et al. (AER 2011)
Chen and Chen (AER 2011)
Ifcher and Zarghamee (AER 2011)
Ericson and Fuster (QJE 2011)
Kirchler et al. (AER 2012)
Bartling et al. (AER 2012)
Huck et al. (AER 2011)
de Clippel et al. (AER 2014)
Duffy and Puzzello (AER 2014)
Charness and Dufwenberg (AER 2011)
Fehr et al. (AER 2013)
Kirchler et al. (AER 2012)
Ericson and Fuster (QJE 2011)
Ifcher and Zarghamee (AER 2011)
Chen and Chen (AER 2011)
Abeler et al. (AER 2011)
Ambrus and Greiner (AER 2012)
Kuziemko et al. (QJE 2014)

Original Estimates
Adjusted Estimates
Applications

Economics Lab Experiments: Adjusted Estimates

Kuziemko et al. (QJE 2014)
Ambrus and Greiner (AER 2012)
Abeler et al. (AER 2011)
Chen and Chen (AER 2011)
Ifcher and Zarghamee (AER 2011)
Ericson and Fuster (QJE 2011)
Kirchler et al (AER 2012)
Fehr et al. (AER 2013)
Duffy and Puzzello (AER 2014)
Charness and Dufwenberg (AER 2011)
Friedman and Oprea (AER 2012)
Kogan et al. (AER 2011)
Dulleck et al. (AER 2011)
Fudenberg et al. (AER 2012)
de Clippel et al. (AER 2014)
Huck et al. (AER 2011)
Bartling et al. (AER 2012)

Original Estimates
Adjusted Estimates
Replication Estimates
Applications

Economics Lab Experiments: Meta-study Approach
Applications

Economics Lab Experiments: Meta-study Results

- Model:

\[ \Theta^* \sim N(0, \tilde{\tau}^2) \]

\[ p(X/\sigma) \propto \begin{cases} 
\beta_p & |X/\sigma| < 1.96 \\
1 & |X/\sigma| \geq 1.96 
\end{cases} \]

- Recall replication-based estimates:

<table>
<thead>
<tr>
<th>\tau</th>
<th>\beta_p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.354</td>
<td>0.100</td>
</tr>
<tr>
<td>(0.750)</td>
<td>(0.091)</td>
</tr>
</tbody>
</table>

- Meta-study based estimates (only \( \beta_p \) comparable):

<table>
<thead>
<tr>
<th>\tilde{\tau}</th>
<th>\beta_p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.299</td>
<td>0.045</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.045)</td>
</tr>
</tbody>
</table>
Applications
Replications of Lab Experiments in Psychology

- Open Science Collaboration (2015)
- 270 contributing authors
- Sample: 100 out of 488 articles published 2008 in
  - Psychological Science
  - Journal of Personality and Social Psychology
  - Journal of Experimental Psychology: Learning, Memory, and Cognition
Applications

Experiments in Psychology: Original and Replication Z Statistics

![Graph showing scatter plot with Z and Z^r axes, with points labeled A and B.]
Applications

Experiments in Psychology: Estimates of Selection Model

- Model:

  \[ \Theta^* \sim N(0, \tau^2) \]

  \[ p(Z) \propto \begin{cases} 
  \beta_{p1} & |Z| < 1.64 \\
  \beta_{p2} & 1.64 \leq |Z| < 1.96 \\
  1 & |Z| \geq 1.96 
\end{cases} \]

- Estimates:

<table>
<thead>
<tr>
<th>\tau</th>
<th>\beta_{p1}</th>
<th>\beta_{p2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.252</td>
<td>0.021</td>
<td>0.294</td>
</tr>
<tr>
<td>(0.195)</td>
<td>(0.012)</td>
<td>(0.128)</td>
</tr>
</tbody>
</table>

- Results insignificant at the 10% level 2% as likely to be published as results significant at 5% level

- Results significant at the 5% level over three times as likely to be published as results significant at 10% level
Applications

Psychology Lab Experiments: Meta-studies Approach
Applications
Psychology Lab Experiments: Estimates of Meta-studies Selection Model

- Model:

\[ \Theta^* \sim N(0, \tau^2) \]

\[ p(Z) \propto \begin{cases} 
\beta_{p1} & |Z| < 1.64 \\
\beta_{p2} & 1.64 \leq |Z| < 1.96 \\
1 & |Z| \geq 1.96 
\end{cases} \]

- Recall replication-based estimates:

<table>
<thead>
<tr>
<th>\tau</th>
<th>\beta_{p1}</th>
<th>\beta_{p2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.252</td>
<td>0.021</td>
<td>0.294</td>
</tr>
<tr>
<td>(0.195)</td>
<td>(0.012)</td>
<td>(0.128)</td>
</tr>
</tbody>
</table>

- Meta-study based estimates (only \( \beta_p \) comparable):

<table>
<thead>
<tr>
<th>\tilde{\tau}</th>
<th>\beta_{p1}</th>
<th>\beta_{p2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.252</td>
<td>0.025</td>
<td>0.375</td>
</tr>
<tr>
<td>(0.041)</td>
<td>(0.015)</td>
<td>(0.166)</td>
</tr>
</tbody>
</table>
Applications

Meta-study of the Effect of Minimum Wages on Employment

- Wolfson and Belman (2015)
- Elasticity of employment w.r.t. the minimum wage
  \[ X > 0 \Leftrightarrow \text{negative employment effect} \]
- 1000 estimates from 37 studies using U.S. data that were circulated after 2000, either as articles in journals or as working papers
- For some: more than 1 estimate per study
Estimates of selection model

- **Model:**

\[ \Theta^* \sim N(\bar{\theta}, \tau^2) \]

\[ p(X/\sigma) \propto \begin{cases} 
    \beta_{p1} & X/\sigma < -1.96 \\
    \beta_{p2} & -1.96 \leq X/\sigma < 0 \\
    \beta_{p3} & 0 \leq X/\sigma < 1.96 \\
    1 & X/\sigma \geq 1.96 
\end{cases} \]

- **Estimates:**

<table>
<thead>
<tr>
<th>( \bar{\theta} )</th>
<th>( \bar{\tau} )</th>
<th>( \beta_{p,1} )</th>
<th>( \beta_{p,2} )</th>
<th>( \beta_{p,3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.024</td>
<td>0.122</td>
<td>0.225</td>
<td>0.424</td>
<td>0.738</td>
</tr>
<tr>
<td>(0.053)</td>
<td>(0.038)</td>
<td>(0.118)</td>
<td>(0.207)</td>
<td>(0.291)</td>
</tr>
</tbody>
</table>

- Bias in favor of estimates which find minimum wage reduces employment
Applications

Meta-Study of the Effects of Deworming

- Croke et al. (2016)
- Follow procedures outlined in the “Cochrane Handbook for Systematic Reviews of Interventions”
- Randomized controlled trials of deworming that include child body weight as an outcome
- 22 estimates from 20 studies
Applications

Meta-Study of the Effects of Deworming
Applications

Deworming: Estimates of selection model

- **Model:**

  \[ \Theta^* \sim N(\bar{\theta}, \tau^2) \]

  \[ p(X) \propto \begin{cases} 
  \beta_p & |X/\sigma| < 1.96 \\
  1 & |X/\sigma| \geq 1.96 
  \end{cases} \]

- **Estimates:**

<table>
<thead>
<tr>
<th>( \bar{\theta} )</th>
<th>( \tilde{\tau} )</th>
<th>( \beta_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.190</td>
<td>0.343</td>
<td>2.514</td>
</tr>
<tr>
<td>(0.120)</td>
<td>(0.128)</td>
<td>(1.872)</td>
</tr>
</tbody>
</table>
Selectivity in the publication process is a potentially serious problem for statistical inference.

We non-parametrically identify the form of selectivity:

- Using replication studies:
  Original and replication estimates would be symmetrically distributed, absent selectivity
- Using meta-studies:
  Under an independence assumption, higher-variance estimates distribution would be noised-up version of lower-variance estimate distribution, absent selectivity
Conclusion

- Easy correction for selectivity, if form is known:
  - Median unbiased estimators
  - Equal-tailed confidence sets with correct coverage

- Empirical findings:
  - Selectivity on significance in experimental economics, experimental psychology
  - Selectivity towards negative employment effects in minimum wage literature
  - Selectivity toward insignificant effects in meta-study for de-worming (but noisy estimate!)
Thank you!