Universidad del Rosario | REBEL Rosario Experimental and Behavioral Economics Lab

Introduction: Experimental Economics –the present, the past and the future

Diego Aycinena

2020 Bogotá Experimental Economics Workshop
Universidad del Rosario, Bogotá, Colombia

This event is possible thanks to the generous support provided by our main sponsors: the *International Foundation for Research in Experimental Economics* (IFREE) and the Department of Economics at Universidad del Rosario.

ifree

REBEL
Rosario Experimental and Behavioral Economics Lab

Outline

- Where are we now?
  - Experimental Economics: methods, conventions and criticisms
- How did we get here?
  - History and development of experiments as research method in economics
- Where are we heading?
  - What are the main challenges and opportunities ahead?

REBEL
Rosario Experimental and Behavioral Economics Lab

Where are we now?

Experimental methods in economics

Diego Aycinena

REBEL
Rosario Experimental and Behavioral Economics Lab

What are experiments?

- Methods to generate data under controlled conditions
    - An important empirical tool for testing theories and generating causal based knowledge
    - Experimental data: deliberatly generated data in a controlled environment

- Experiments foster a healthy interaction between theory and empirical testing
    - An economic experiment generates, in a controlled environment, the economic situation that we want to analyze in order to subsequently be able to make variants of it and compare them.

REBEL
Rosario Experimental and Behavioral Economics Lab

Friedmand & Sunder (1994) "Experimental Methods: A Primer for Economists"

Types of Data

	Happenstance	Experimental
Field	Macro data macro (GDP, inflation, Exchange rate, asset returns, etc.) Administrative data (standardized test scores, crime statistics, income and tax statistics, etc.) Firms & HH's surveys (i.e. LSMS)	NFE, RCT's & IE experiments (education, tax compliance, savings & microfinance, remittances, agricultural productivity, etc.) FFE & *nudge* experiments (organ donation, wáter & energy conservation, pension fund contributions, etc.)
Lab	?	Experimental games (UG, DG, PGG, TG, ME-WL G, etc.) Auctions, asset markets, complex market design and mehanism design Individual decisions (time, risk, honesty) Other (voting & political economy, corruption, health, neuroecon, emotions etc.)

REBEL
Rosario Experimental and Behavioral Economics Lab

Friedmand & Sunder (1994) "Experimental Methods: A Primer for Economists"

Types of Data

	Happenstance	Experimental
Field	Macro data macro (GDP, inflation, Exchange rate, asset returns, etc.) Administrative data (standardized test scores, crime statistics, income and tax statistics, etc.) Firms & HH's surveys (i.e. LSMS)	NFE, RCT's & IE experiments (education, tax compliance, savings & microfinance, remittances, agricultural productivity, etc.) FFE & *nudge* experiments (organ donation, wáter & energy conservation, pension fund contributions, etc.)
Lab	?	Experimental games (UG, DG, PGG, TG, ME-WL G, etc.) Auctions, asset markets, complex market design and mehanism design Individual decisions (time, risk, honesty) Other (voting & political economy, corruption, health, neuroecon, emotions etc.)

REBEL
Rosario Experimental and Behavioral Economics Lab

What are experiments?

- Methods to generate data under controlled conditions
  - An important empirical tool for testing theories and generating causal based knowledge
- Experiments foster a healthy interaction between theory and empirical testing
  - An economic experiment generates, in a controlled environment, the economic situation that we want to analyze in order to subsequently be able to make variants of it and compare them.
- Two main advantages of experimental data over happenstance data:
  - Control
  - Replication

REBEL
Rosario Experimental and Behavioral Economics Lab

Control (I)

- Smith (1976, p. 275) defines control as the essence of experimental methodology
- Direct control: keep constant or vary as treatments
  - Incentives
    - Payoffs
  - Institutions
    - i.e. rules: auctions, markets, voting, communication, punishment & retaliation, etc.
  - Environments
    - Values, costs, resource endowment, production function, information, etc.
  - Complete direct control may not always be possible

REBEL
Rosario Experimental and Behavioral Economics Lab

Control (II)

- What if complete/direct control is not possible?
  - i.e. altruism, spite and social preferences, social norms and beliefs, understanding (common knowledge?)
  - Randomization as indirect control
    - Avoids self-selection or correlation of treatment with individual characteristicas
    - Exogenous (random) assignment of treatments allows controlled variation → Important for establishing causal relationships
  - Measurement
    - Measure social preferences, risk preferences, beliefs, understanding, etc.
  - Control (and/or measure) nuissance variables
    - Isolate focus variables

REBEL
Rosario Experimental and Behavioral Economics Lab

Replication

- Allows to bulid knowledge by relying on previous experimental findings
    - Facilitates a cumulative and systematic process of experimental learning
- Allows to check for robustness of findings, text for experimenter effects, explore *edge of validity*, etc.
    - Gives researchers an improved incentive to do things right
    - Availability of data, instructions, programs, protocols and procedures, etc.
        - Important to design an experiment that can be repeated by other researchers as this puts a safeguard in science; scientific knowledge must be replicable
- Status of replicability?
    - Recent evidence suggests replicability is far from the scientific benchmark (more on this later)

REBEL
Rosario Experimental and Behavioral Economics Lab

Smith, V.L. (1982) "Microeconomic Systems as an Experimental Science" *AER*

Microeconomic Systems (Smith, 1982)

(1) Environment:

- *consists of the collection of all agents' characteristics*
  - *i.e. individual demand (willingness-to-pay) and supply (willingness-to-accept) schedules.*

(2) Institution:

- *language (messages* or actions) of communication
- rules of communication, allocation, cost imputation, and order of moves
  - i.e., bids by buyers, offers by sellers, acceptances by either

(3) Behavior:

- *agent choices of messages or actions* given the environment and the institutions relating choices to allocations.

REBEL
Rosario Experimental and Behavioral Economics Lab

Smith, V.L. (1989) "Theory, Experiment and Economics" *JEP*.

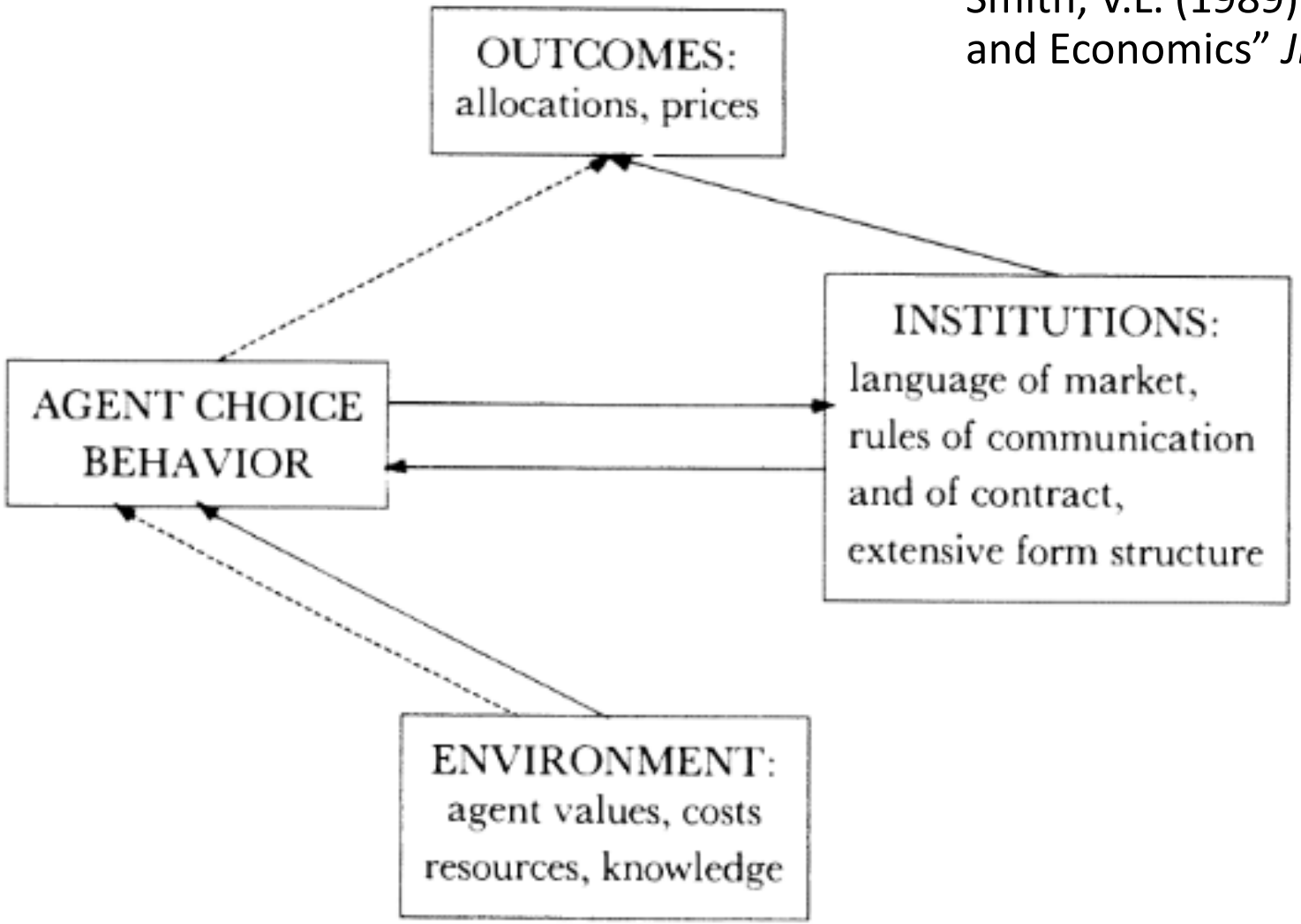


Fig. 1. Institutions in economic theory

REBEL
Rosario Experimental and Behavioral Economics Lab

Types of Experiments

- (Traditional) Lab Experiments

- Artefactual Field Experiments
  - same as a lab experiment but with a non-standard subject pool
    - Binswanger (1980), Karlan (2005), Attanasio et al. (2012), Callen et al. (2014) y Jakiela & Ozier (2015), Friebel et al., (2018)
- Framed Field Experiments
  - same as AFE but with field context in either the commodity, task, or information set that the subjects can use
    - Cox, et al. (2016a, 2016b), Cardenas (2000, 2004), Velez et al. (2009, 2010), Janssen et al. (2013), Hill & Viceisza (2012), Barr et al. (2009), Ambler (2015), Bah & Batista (2018)
- Natural Field Experiments
  - same as a framed field experiment but where the environment is one where the subjects naturally undertake these tasks and where the subjects do not know that they are in an experiment
    - Armantier & Boly (2011, 2013), Ferraro & Price (2013); Schultz et al. (2016), Hallsworth et al. (2017), Banerjee et al., (2016)

REBEL
Rosario Experimental and Behavioral Economics Lab

Experimental design terminology I

- Treatment assignment:
  - Between subjects: some (random) subjects receive one treatment, others receive a different treatment → each subjects assigned to only one treatment
  - Within subjects: each subject assigned to more than one treatment
    - Dual trial: both treatments at the same time
    - Crossover [design]: Treatments assigned in sequence
      - Usually vary order and control for order effects
- Response method:
  - Direct response: subject makes direct (non-contingent) choice/decision
  - Strategy method: subject makes a decision that is contingent on the choice of other subject(s), i.e. reveals strategy profile
    - Other procedures: CIL, etc.

REBEL
Rosario Experimental and Behavioral Economics Lab

Experimental design terminology II

- Matching protocols:
  - Partners: always play with same group
  - Strangers: randomly re-matched before playing each game
  - Perfect strangers: subjects do not play with the same subjects more than once

- Anonimity:
  - Single-blind procedure: anonymity between subjects
  - Double-blind procedure: anonymity between subjects and between experimenter and subjects

REBEL
Rosario Experimental and Behavioral Economics Lab

How to conduct an experiment (in 1 slide)

- Start with a precise and clearly defined research question
  - Define focus variable
- Choose an experimental design that answers your question
  - Define treatments
    - Treatments: between-subjects vs within-subjects?
  - Direct response or strategy method?
  - One shor or repeated trials
    - Matching protocol?
  - Choose parameters and payment stakes
  - Determine number of required (indep.) observations
    - Power analysis
- Write instructions, protocols, software, session script, etc.
- Collect and analyze data

REBEL
Rosario Experimental and Behavioral Economics Lab

Current Experimental Conventions

- Use monetary incentives (performancepay)
  - Advantages: more effort / attention / less noise
  - Disadvantages: its expensive and limits stakes
- Do not use deception
  - Advantages: lowers costs, makes it easier to study rare situations, and makes it easier to design an experiment
  - Disadvantages: you lose control if subjects do not believe the instructions and try to outguess the experimenter, imposes an externality on other researchers
    - Ultimately it is an empirical issue, but there is a strong norm in experimental economics
- Use neutral framing
  - Advantages of concrete wording: can help subjects understand the experiment and bring the experiment closer to research question
  - Disadvantages of concrete wording: you lose control since you don't know how subjects perceive their role (e.g. subjects might role-play a manager's behavior)

REBEL
Rosario Experimental and Behavioral Economics Lab

Common Concerns I

- Artificiality
  - Not new… (argued by Friedman & Walis when critizing Thurstone, 1931)
  - Artificial setting…
  - …with real people making real money-relevant decisions
  - Is this relevant for testing theories?
    - Our theories and models are general enough that they do not specify that they only hold in naturally occurring settings
    - If a theory "fails" in the artificial setting of the lab, we can examine the reasons for it (and learn more!)
      - …instead of discarding the results due to "artificiality"

REBEL
Rosario Experimental and Behavioral Economics Lab

Common Concerns II

- External validity or Generalizability
  - "Sure, this applies in the lab, with students; but what about outside of the lab?"
    - Common assumption: effects found in the lab unlikely to hold outside of the lab
  - Empirical matter
    - Gneezy and List (2006) -, Kube et al. (2012) Camerer (2015) +, Herbst and Mas (2015) +
  - Not exclusive to experiments; applies to any empirical study
    - Orr et al. (2019)-, Meager (2019)+
  - Bottom line: sure, we need to think (and test!) how generalizable are experimental results (as any empirical results); but this is no reason to not do experiments.
    - It is a reason to do *more* experiments to deliberately test conditions under which results hold (and learn more).

REBEL
Rosario Experimental and Behavioral Economics Lab

Common Concerns III

- Are students real people?
  - Representativeness of subjects
  - Students are very commonly used
    - Not because we want to learn something about students
    - Or because we believe students (who participate in the lab) are a representative sample of humanity
    - But because they are convenient
      - Low opportunity cost
      - Rapid learning curve
  - Relevant for comparative statics? Can be tested
    - Not important (qualitative/direction effect) in many situations
      - No difference in 9/13 in Frechette, 2015 review; only in 1/13 were pros closer to theory than students.
    - E.g., Weitzel et al. (2019): Finance professionals are not immune to bubbles in experimental asset markets
      - bubble-drivers—capital inflows or high initial capital supply— affect students as well as professionals; bubble-moderators yield efficient markets in both groups

REBEL
Rosario Experimental and Behavioral Economics Lab

Common Concerns

- External validity or Generalizability
- Artificiality
- Are students real people?
- Others (stakes, sample sizes, scrutiny, self-selection, etc.)
- Concerns should be taken seriously, not dismissed
  - How important? Depends on the nature of the research question
- Experimental results should not be dismissed because of potential concerns
  - We need (additional work) to examine and address these concerns
- Controlled variation is important for causal knowledge
- The lab offers cheap and direct control of (many) relevant variables
  - Lab Experiments Are a Major Source of Knowledge in the Social Sciences: if the objective is to learn and generate knowledge, more (not less) lab experiments should be conducted to address and examine these potential concerns
  - Experiments are no magic bullet, they are complements of other empirical methods

REBEL
Rosario Experimental and Behavioral Economics Lab

How did we get here?

History of experimental methods in economics

Diego Aycinena

Economics: An Experimental Science?

- "Unfortunately, we can seldom test particular predictions in the social sciences by experiments explicitly designed to eliminate what are judged to be the most important disturbing influences. […] The necessity of relying on uncontrolled experience rather than on controlled experiment makes it difficult to produce dramatic and clear-cut evidence to justify the acceptance of tentative hypothesis."
  - (Friedman, 1953, p. 10)

Economics: An Experimental Science?

J.S. Mill said it first:

"The first difficulty which meets us in the attempt to apply experimental methods for ascertaining the laws of social phenomena, is that we are without the means of making artificial experiments. Even if we could contrive experiments at leisure, and try them without limit, we should do so under immense disadvantage; … But it is unnecessary to consider the logical objections which would exist to the conclusiveness of our experiments, since we palpably never have the power of trying any. We can only watch those which nature produces, or which are produced for other reasons."

(Mill 1872, p. 881)

REBEL
Rosario Experimental and Behavioral Economics Lab

Economics: An Experimental Science?

- "Economists cannot make use of controlled experiments to settle their differences"
  - (Robinson 1979, p. 1319)

- "Economics … cannot perform the controlled experiments of chemists or biologists because [it] cannot easily control other important factors. Like astronomers or meteorologists, [it] generally must be content largely to observe."
  - (Samuelson and Nordhaus, 1985, p. 8)

REBEL
Rosario Experimental and Behavioral Economics Lab

Economics as an Experimental Science

- Currently → experimental methods widely used in economics
  - Over 200 experimental Labs around the world
- How did we get here?
  - How did we go from believing experiments were imposible (or useless) to being common and prevalent?
- How did it all start?

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments?

- Chamberlin (1948)
  - Imperfect competition in markets
- Thurstone (1931)
  - Experiments exploring preferences over diferent goods trying to identify indifferent curves
  - Wallis & Friedman critized the artificiality and hypothetical nature of decisions
- Bernoulli (1783)
  - Experiments over St. Petersburg's paradox
- Sparks in the dark: all isolated instances of experiments
  - Did not originate a cumulative process of learning
    - "Chamberlin's experiments were a lone outlier in the methodological spectrum then current in economics." (Smith 1992, p. 242)
- Firsts vs. *Last firsts*? (Roth, 1993)

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments

- "History suggests that a discipline becomes experimental *when innovators develop techniques for conducting relevant experiements.*" (Sunder & Friedman, p.1)
  - Theoretical/conceptual framework, methodology

- Advances in microeconomic theory:
  - Game theory & EUT,
  - Industrial organization and market structure

REBEL
Rosario Experimental and Behavioral Economics Lab

History of Experimental Economicss

- Pre-history (?-1948): Sparks in the dark
- Origins (1949-1967's): Smoke & Heat
  - Santa Monica Conference
  - S. Siegel (with L. Fouraker, M. Shubik)
  - H. Sauermann & R. Selten
  - V. L. Smith
- Foundations and development (1967's-1985's): Fire
  - The workshops
  - Associations
- Maturity and Growth (1986-): Light
  - ESA
  - ExEcon, JESA

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments

1. RAND (including Nash): Examining game theoretic predictions
    - Kalish, Milnor, Nash and Nehrig (1954) y Flood (1952, 1958, 1959)
- 1952 Conference on "The Design of Experiments in Decision Processes“

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments

1. RAND (including Nash): Examining game theoretic predictions
2. Sidney Siegel (with L. Fouraker, M. Shubik): Bargaining and experimental protocols
   - Siegel (1957), Siegel and Fouraker (1960), Fouraker and Siegel (1963)

- Importance of relevant payments, information conditions and detailed instructions (including instructions in research papers)

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments

1. RAND (including Nash): Examining game theoretic predictions
2. Sidney Siegel (with L. Fouraker, M. Shubik): Bargaining and experimental protocols
3. Selten and the German School: limited rationality and behavioral processes in games and decisions
   - Sauerman and Selten (1959) “Ein Oligopolexperiment”

- Developed in parallell and mostly independent from the groups in the US (1960 – mid 80’s)

REBEL
Rosario Experimental and Behavioral Economics Lab

Origins: First Experiments

1. RAND (including Nash): Examining game theoretic predictions
2. Sidney Siegel (with L. Fouraker, M. Shubik): Bargaining and experimental protocols
3. Selten and the German School: limited rationality and behavioral processes in games and decisions
4. Smith & Plott: Markets (& committees)
   - Smith (1962, 64), Smith and Plott (1978), Fiorina and Plott (1978)

- 70's y 80's VL Smith published theoretical framework behind experimental economics
  - Smith (1976), Smith (1982)
- Experimental econ grew exponentially in the 80's & 90's

REBEL
Rosario Experimental and Behavioral Economics Lab

Foundations: the workshops & associations

- The German School
  - The 1st research community of experimental economists in the world
    - formed around Heinz Sauermann (1905-1981) in Frankfurt in the 1960s
    - Experimental economics conference in 1967
  - 1977 → *GfeW*, *The Society for Experimental Economics Research*
    - 1st association dedicated to experimental research in economics; became the umbrella organization for experimentalists from German-speaking countries.

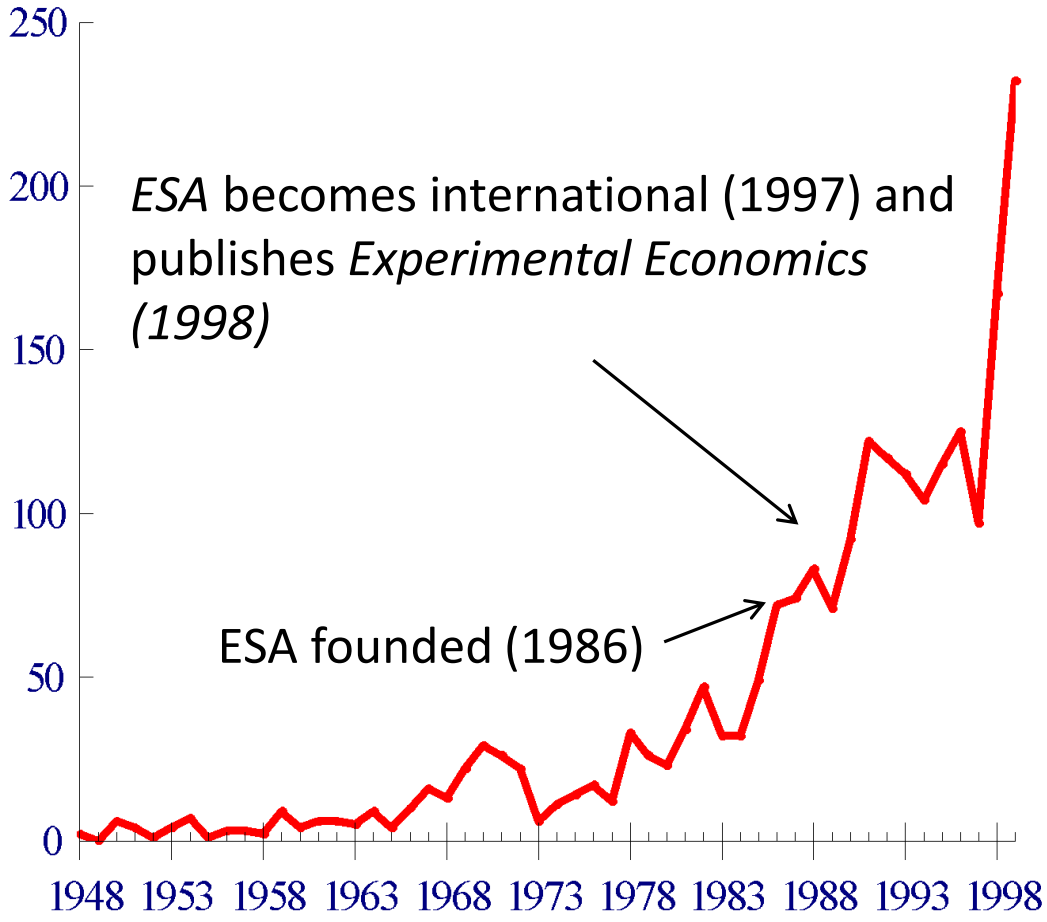
- The Tucson Workshops (at the Westward Look)
  - 1st workshop, March 18-20, 1977
    - focused on completed research
    - The papers presented in the first workshop were refereed and some were accepted for the first volume of Research in Experimental Economics (1979)
  - 2nd workshop, also entitled the NSF Experimental Economics conference, October 19-21, 1979
  - 3rd Experimental Economics Workshop, March 27-9, 1984
  - ESA born in 1986, fully international by 1997

Maturity and growth of Experiments

- 1985: JEL introduced "Experimental Economic Methods“ code
- 1986: *Economic Science Association* (*ESA*) is born
- 1997: *ESA* becomes internacional
- 1998: *Experimental Economics* by ESA

Maturity and growth of Experiments

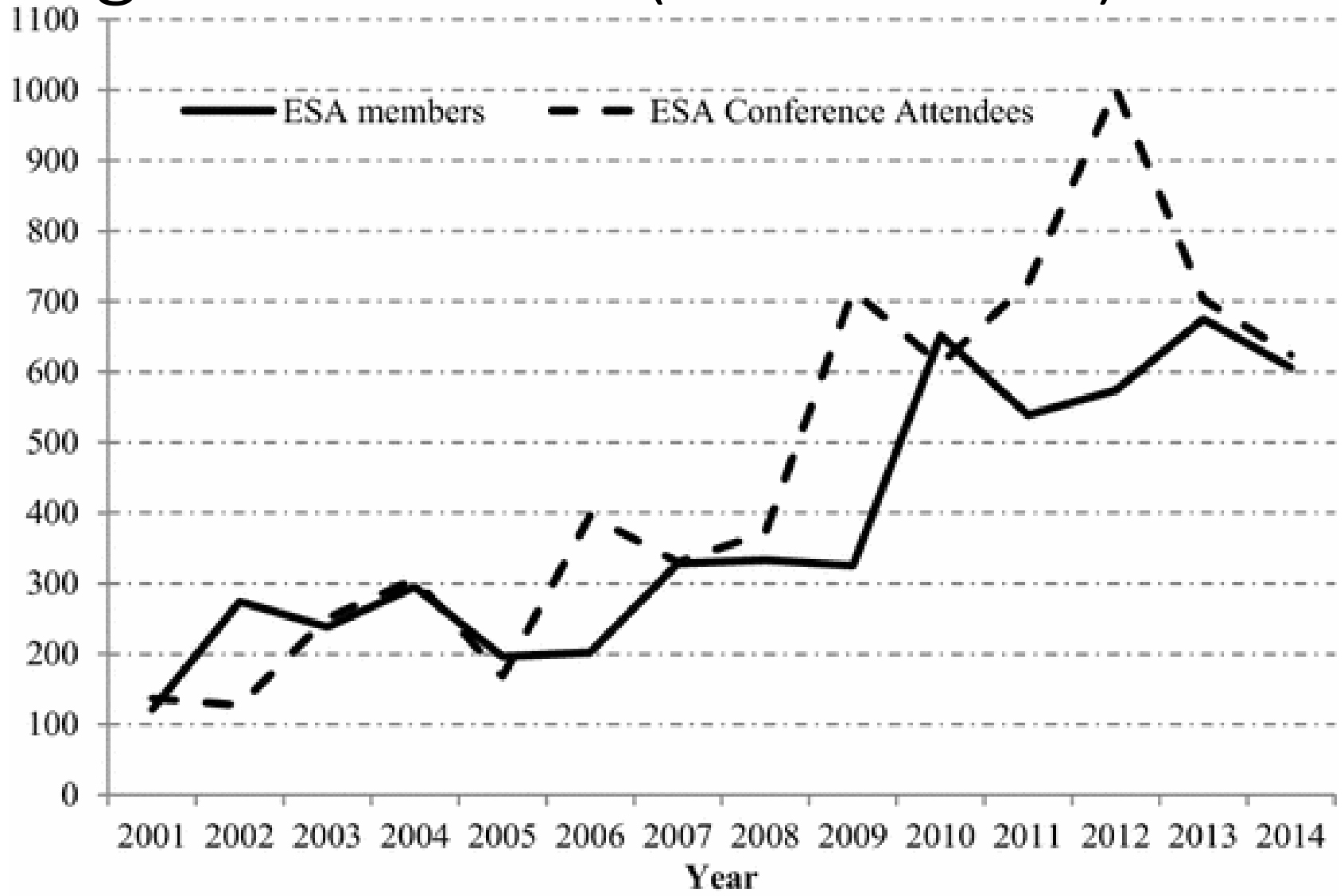
- 1985: JEL introduced "Experimental Economic Methods“ code
- 1986: *Economic Science Association* (*ESA*) is born
- 1997: *ESA* becomes internacional
- 1998: *Experimental Economics* by ESA



Maturity and growth of Experiments

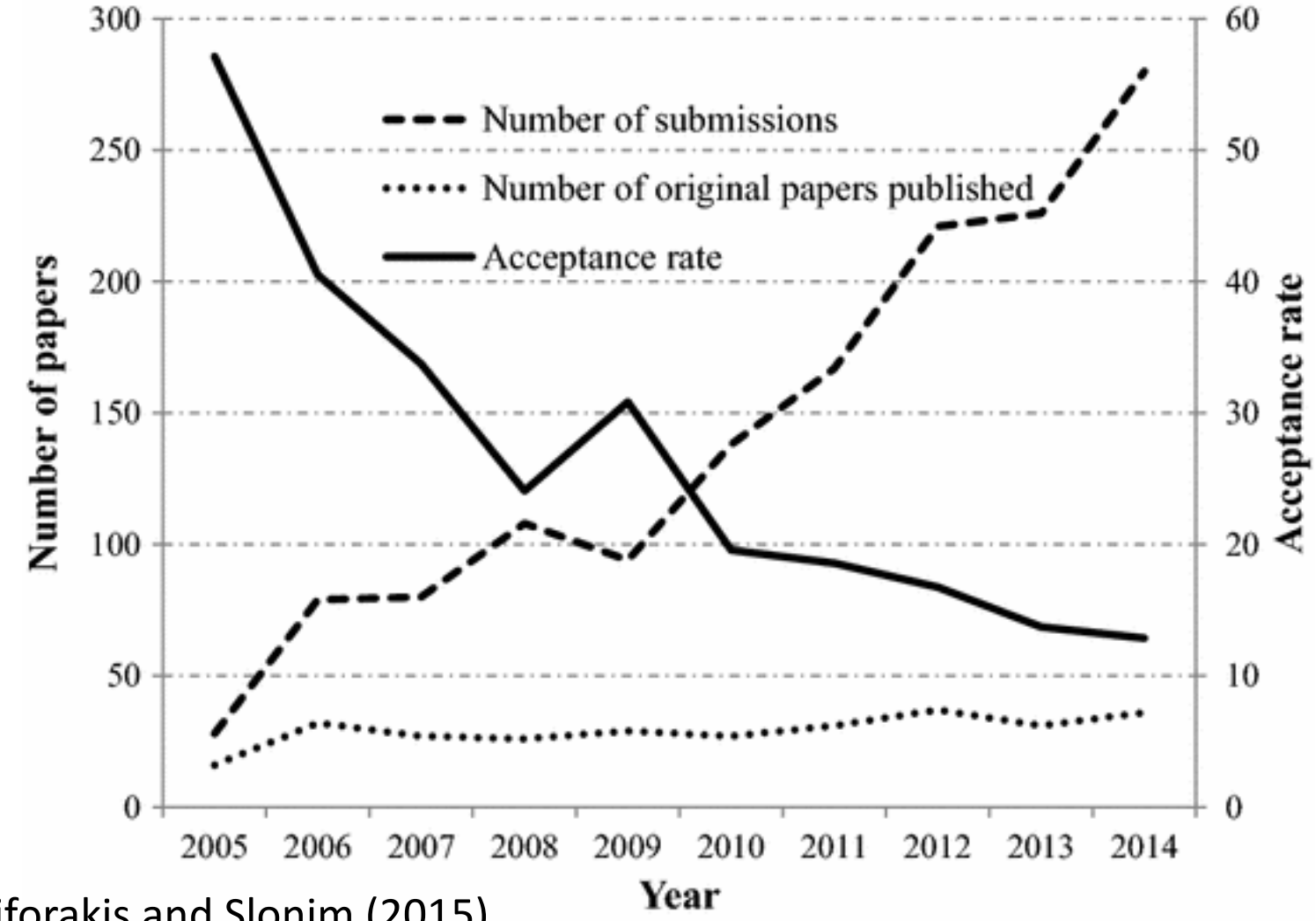
- 1985: JEL introduced "Experimental Economic Methods“ code
- 1986: *Economic Science Association* (*ESA*) is born
- 1997: *ESA* becomes internacional
- 1998: *Experimental Economics* by ESA
- 2002: Nobel to **Vernon L. Smith** (& **D. Kahneman**)
  - "*for having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms*“
- 2009: Nobel to **Elinor Ostrom** (& O. Williamson)
- 2012: Nobel to **Alvin E. Roth** (& Lloyd S. Shapley)
- 2015: *Journal of the Economic Science Association*

The growth of ESA (2001–2014)



Nikiforakis and Slonim (2015)

Publications in *Experimental Economics*



Nikiforakis and Slonim (2015)

Maturity and growth of Experiments

- 1985: JEL introduced "Experimental Economic Methods“ code
- 1986: *Economic Science Association* (*ESA*) is born
- 1997: *ESA* becomes internacional
- 1998: *Experimental Economics* by ESA
- 2002: Nobel to **Vernon L. Smith** (& **D. Kahneman**)
  - "*for having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms*“
- 2009: Nobel to **Elinor Ostrom** (& O. Williamson)
- 2012: Nobel to **Alvin E. Roth** (& Lloyd S. Shapley)
- 2015: *Journal of the Economic Science Association*
- 2018: Nobel 2018: Abhijit Banerjee, Esther Duflo & Michael Kremer
  - *"for their experimental approach to alleviating global poverty”*

Origins: Field Experiments (last firsts)

1. David Reily: Auction experimentos using mail lists over the internet
   - Lucking-Reiley (1999, 2006) y List and Lucking-Reiley (2000, 2002)

2. John List : experiments using *sports cards*
   - List and Shogren (1998), List (2002)

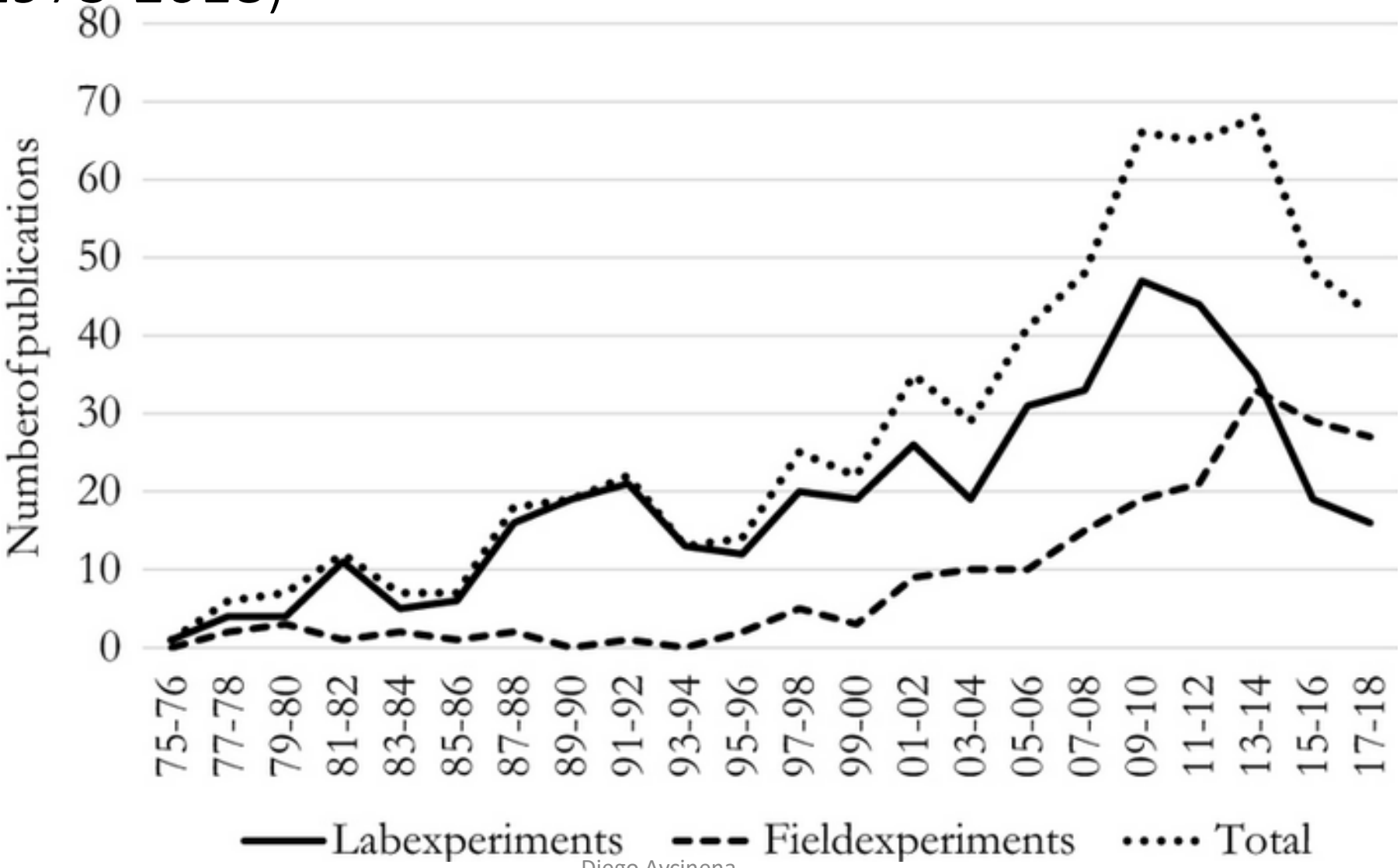
3. J-PAL : Impact evaluation (RCTs)
   - Esther Duflo y Abhijit Banerjee

- In contrast to lab experiments (motivated by advances in theory), field experiments were motivated by econometric identification problems

REBEL
Rosario Experimental and Behavioral Economics Lab

Nikiforakis and Slonim (2019)

Top 5 Experimental Publications (1975-2018)



REBEL
Rosario Experimental and Behavioral Economics Lab

Where are we going?

New challenges and opportunities for experimental economics

Diego Aycinena

REBEL
Rosario Experimental and Behavioral Economics Lab

Replicability Crisis?

- Crisis of confidence in scientific findings
  - Ioannidis, 2005; Simmons et al., 2011; Aarts et al., 2015;

Source: Openn Science Collaboration

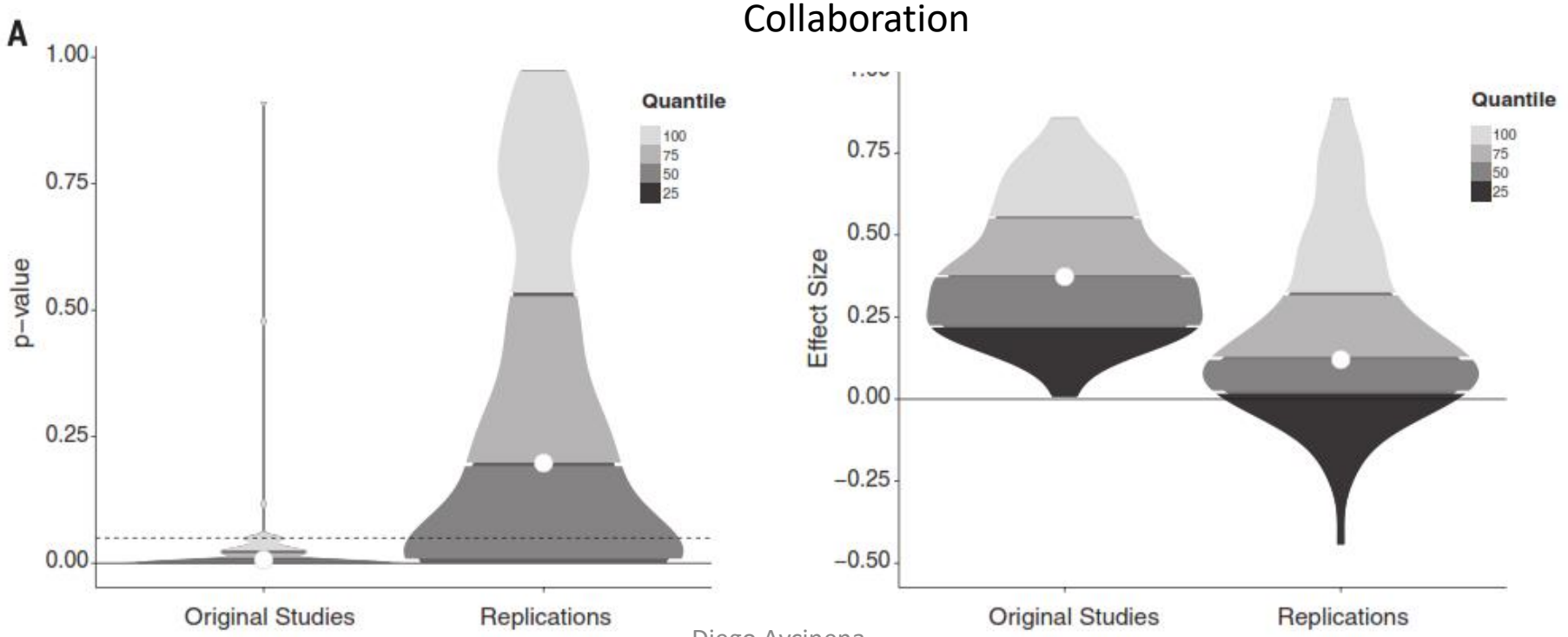


Fig. 1. Density plots of original and replication *P* values and effect sizes. (**A**) *P* values. (**B**) Effect sizes (correlation coefficients). Lowest quantiles for *P* values are not visible because they are clustered near zero.

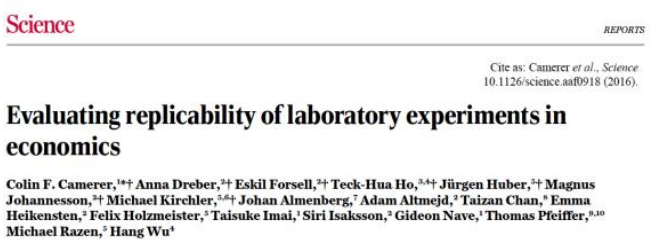
REBEL
Rosario Experimental and Behavioral Economics Lab

(Major) Replication Projects in Psychology

- Reproducibility Project: 36/100
  - 100 social/cognitive findings; pseudo-random selection from 4 top journals
- Many Labs: 10/13
  - 13 classic & contemporary effects across 36 labs (25 US, 11 Intl), N=6,344
- Many Labs 2: 14/28
  - 28 classic and contemporary psychology findings replicated in more than 60 labs each across three dozen nations and territories with N=15,305
- Many Labs 3: 3/10
  - 10 time of semester known effects (+ 10 individual diffs, & 3 data quality indicators) over the academic semester in 20 participant pools (N = 2,696) & online (N = 737)

- **Replicability in Psychology: 63/151 = 41.7%**

REBEL
Rosario Experimental and Behavioral Economics Lab


Science REPORTS

Cite as: Camerer *et al.*, *Science* 10.1126/science.aaf0918 (2016).

Evaluating replicability of laboratory experiments in economics

Colin F. Camerer,[1*†] Anna Dreber,[2†] Eskil Forsell,[2†] Teck-Hua Ho,[3,4†] Jürgen Huber,[5†] Magnus Johannesson,[2†] Michael Kirchler,[5,6†] Johan Almenberg,[7] Adam Altmejd,[3] Taizan Chan,[8] Emma Heikensten,[2] Felix Holzmeister,[5] Taisuke Imai,[1] Siri Isaksson,[2] Gideon Nave,[1] Thomas Pfeiffer,[9,10] Michael Razen,[5] Hang Wu[4]

Replicability of Lab Experiments I

- Camerer, et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433-1436.
- Replicated 18 studies published in the *American Economic Review* and the *Quarterly Journal of Economics* between 2011 and 2014.
- All replications followed predefined analysis plans that were made publicly available beforehand
- All have a statistical power of at least 90% to detect the original effect size at the 5% significance level.
- Authors ran prediction market

REBEL
Rosario Experimental and Behavioral Economics Lab

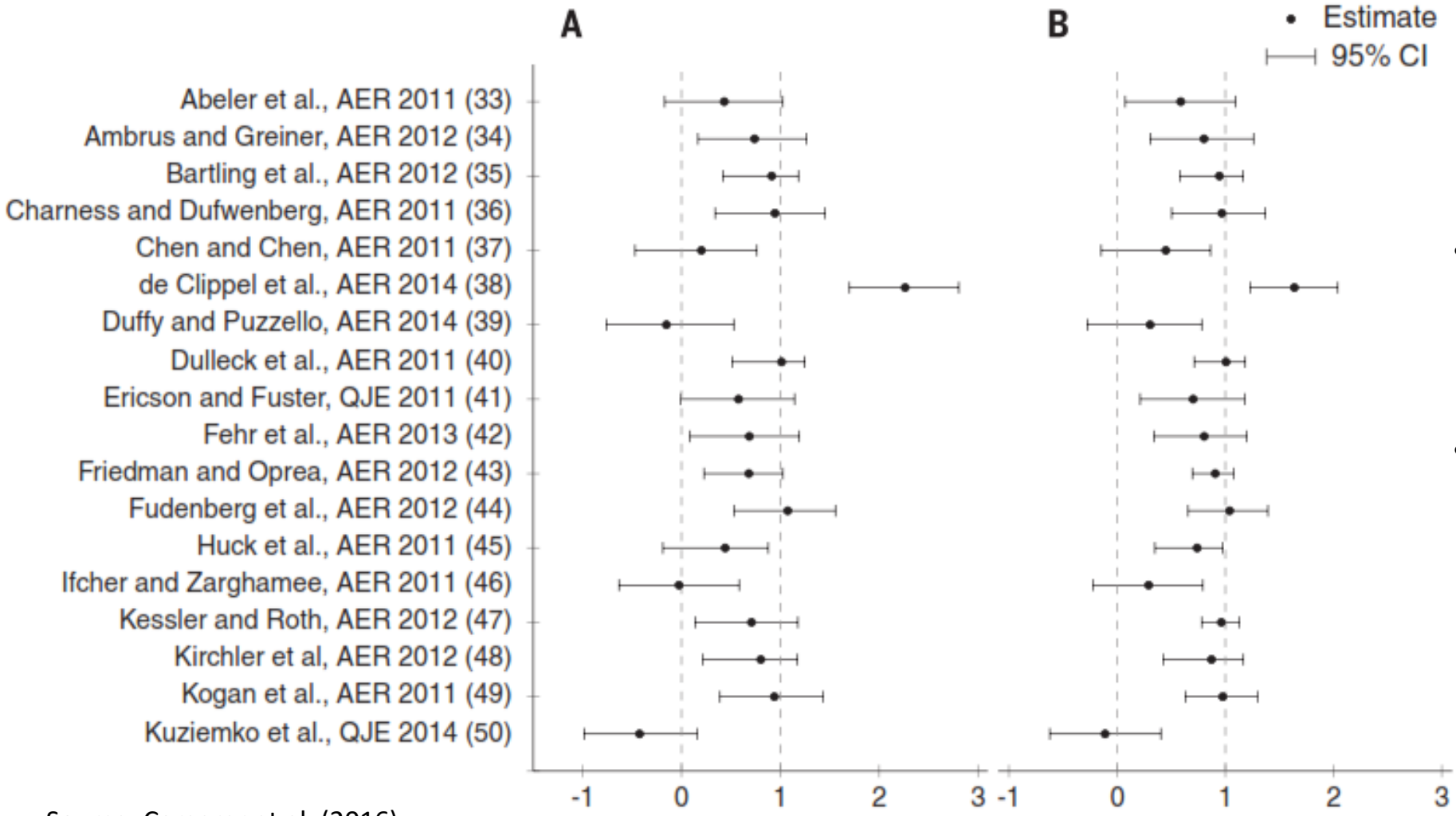
Replicability of Lab Experiments I

Science REPORTS

Cite as: Camerer *et al.*, *Science* 10.1126/science.aaf0918 (2016).

Evaluating replicability of laboratory experiments in economics

Colin F. Camerer,[1*†] Anna Dreber,[2†] Eskil Forsell,[2†] Teck-Hua Ho,[3,4†] Jürgen Huber,[5†] Magnus Johannesson,[2†] Michael Kirchler,[5,6†] Johan Almenberg,[7] Adam Altmejd,[2] Taizan Chan,[8] Emma Heikensten,[2] Felix Holzmeister,[5] Taisuke Imai,[1] Siri Isaksson,[2] Gideon Nave,[1] Thomas Pfeiffer,[9,10] Michael Razen,[5] Hang Wu[4]



- Replicated 11/18 (61%)
  - Significant effect in the same direction as in original study
  - Average replicated effect size is 66% of original
- Replicability rate varies between 66.7% and 72.5% for additional indicators

Source: Camerer et al. (2016)

REBEL
Rosario Experimental and Behavioral Economics Lab

Replicability of Experiments II

- Camerer et al. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637.
- Replicate 21 social sciences experimental studies published in *Nature* and *Science* 2010-2015.
- Follow analysis plans reviewed by the original authors and pre-registered.
- High powered: original sample powered 90% to detect an effect size =*1/2 of original*
  - Typical sample 5x original!
- Phase II sample if phase I replicates poorly

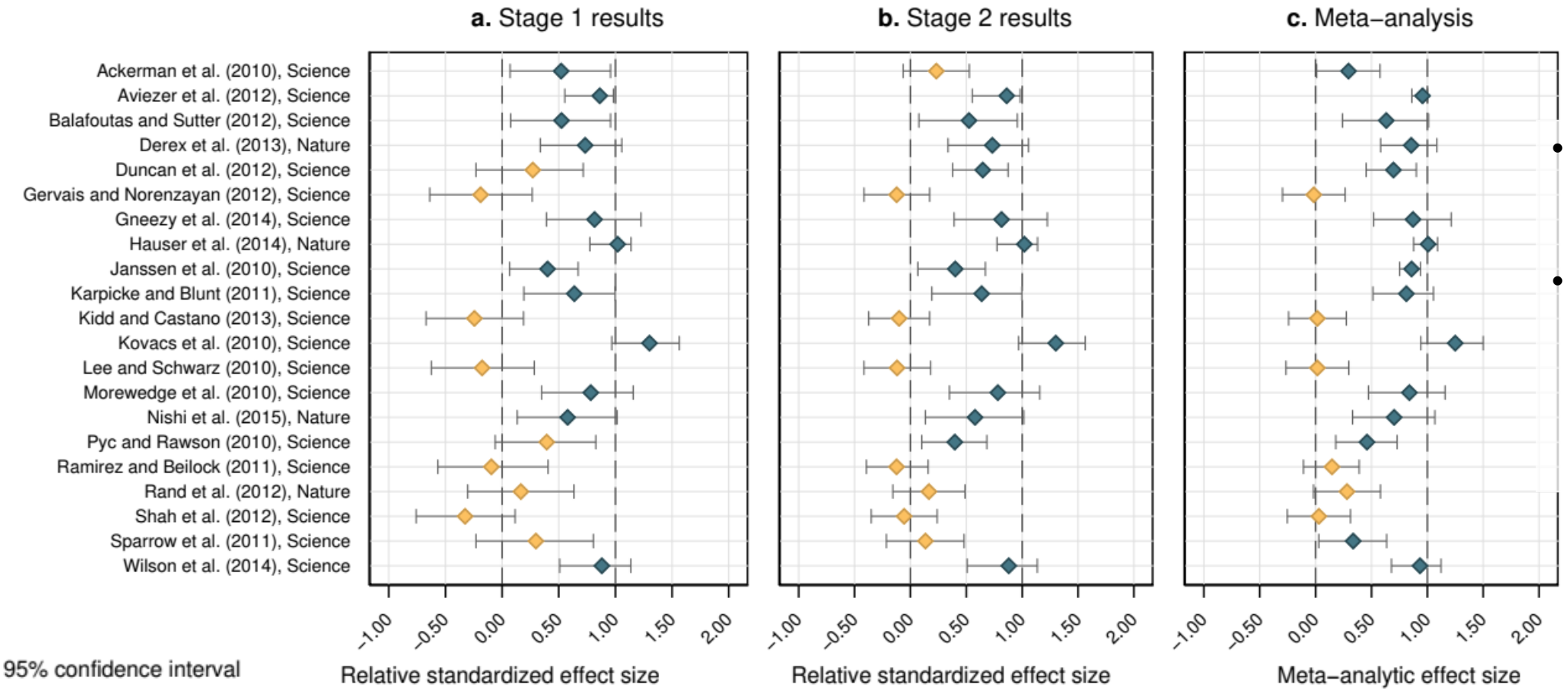
REBEL
Rosario Experimental and Behavioral Economics Lab

nature human behaviour
Letter · Published: 27 August 2018

Replicability of Experiments II

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

Colin F. Camerer, Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A. Nosek, Thomas Pfeiffer, Adam Altmejd, Nick Buttrick, Taizan Chan, Yiling Chen, Eskil Forsell, Anup Gampa, Emma Heikensten, Lily Hummer, Taisuke Imai, Siri Isaksson, Dylan Manfredi, Julia Rose, Eric-Jan Wagenmakers & Hang Wu



- Replicated 13/21 (62%) studies
- Replicability rate varies between 57% and 67% for additional indicators

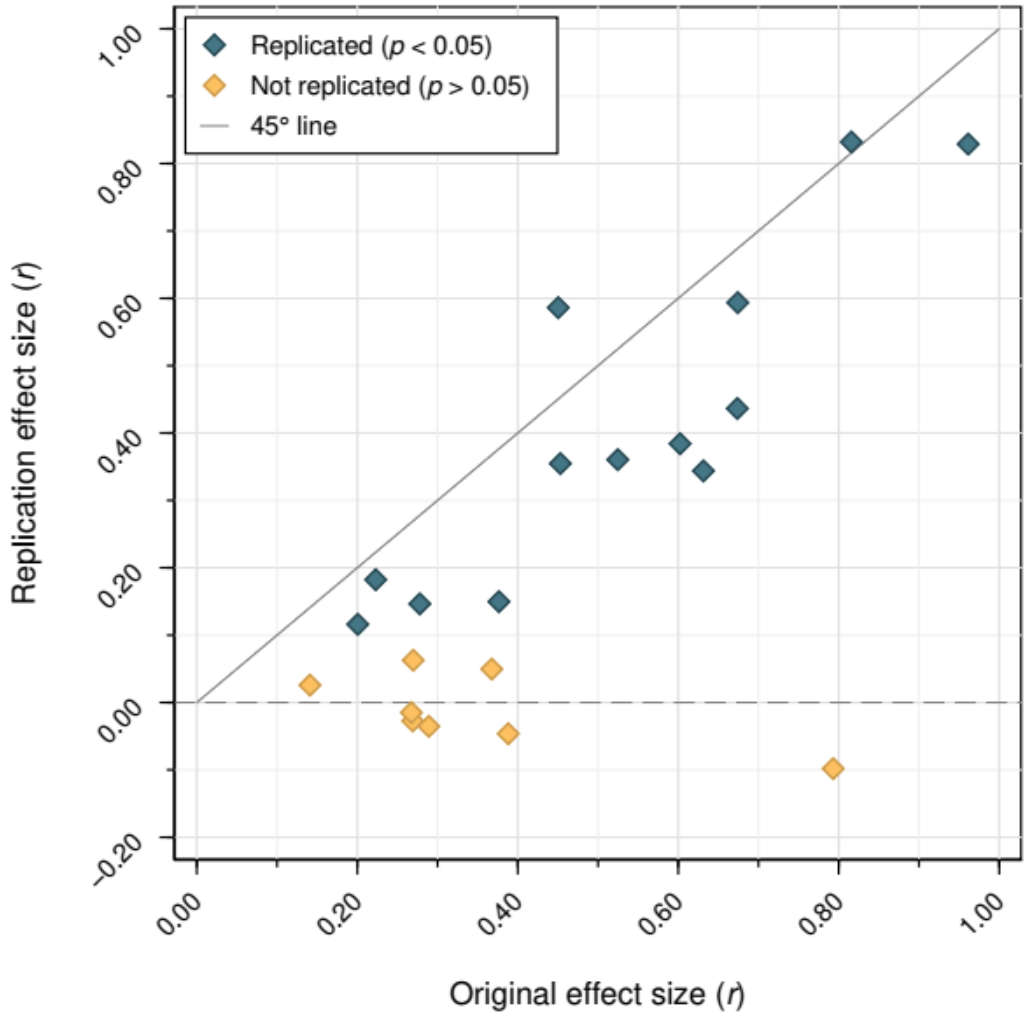
REBEL
Rosario Experimental and Behavioral Economics Lab

Replicability of Experiments II

nature human behaviour
Letter · Published: 27 August 2018

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

Colin F. Camerer, Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A. Nosek, Thomas Pfeiffer, Adam Altmejd, Nick Buttrick, Taizan Chan, Yiling Chen, Eskil Forsell, Anup Gampa, Emma Heikensten, Lily Hummer, Taisuke Imai, Siri Isaksson, Dylan Manfredi, Julia Rose, Eric-Jan Wagenmakers & Hang Wu



- Average effect size of replications is about 50% of the original.

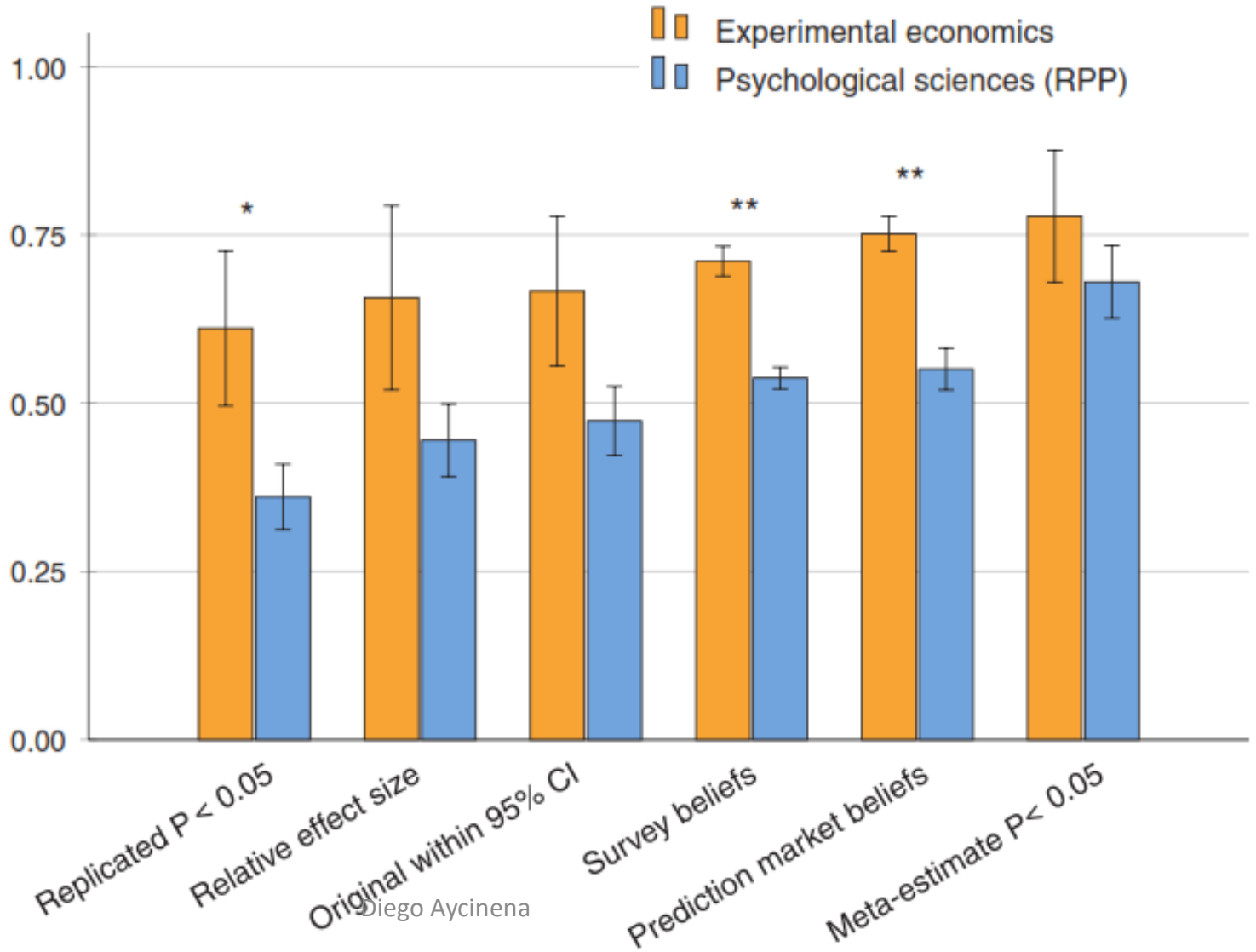
REBEL
Rosario Experimental and Behavioral Economics Lab

Science
REPORTS
Cite as: Camerer et al., Science 10.1126/science.aaf0918 (2016).

Evaluating replicability of laboratory experiments in economics

Colin F. Camerer,[1*†] Anna Dreber,[2†] Eskil Forsell,[2†] Teck-Hua Ho,[3*†] Jürgen Huber,[2†] Magnus Johannesson,[2†] Michael Kirchler,[5*†] Johan Almenberg,[7] Adam Altmejd,[3] Taizan Chan,[8] Emma Heikensten,[4] Felix Holzmeister,[5] Taisuke Imai,[1] Siri Isaksson,[9] Gideon Nave,[1] Thomas Pfeiffer,[8,10] Michael Razen,[5] Hang Wu[1]

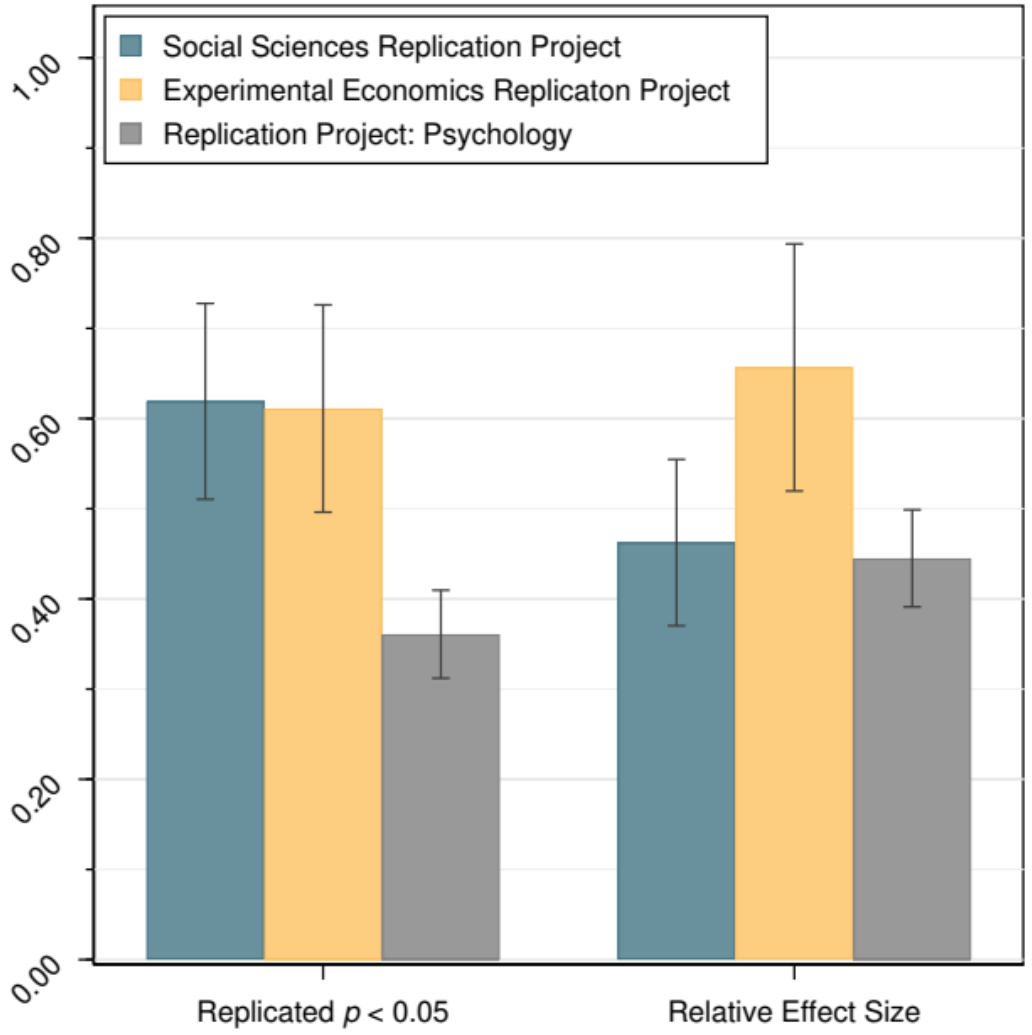
Replicability in Experimental Economics vs Psychology



Source: Camerer et al., 2016

REBEL
Rosario Experimental and Behavioral Economics Lab

Replicability Comparisons



Comparison of replicability indicators between the Social Sciences Replication Project (Camerer et al., 2018), the Experimental Economics Replication Project (Camerer et al., 2016), and the Reproducibility Project: Psychology (RPP)

Source: Camerer et al. (2018)


nature human behaviour

Letter · Published: 27 August 2018

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

Colin F. Camerer, Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A. Nosek, Thomas Pfeiffer, Adam Altmejd, Nick Buttrick, Taizan Chan, Yiling Chen, Eskil Forsell, Anup Gampa, Emma Heikensten, Lily Hummer, Taisuke Imai, Siri Isaksson, Dylan Manfredi, Julia Rose, Eric-Jan Wagenmakers & Hang Wu

REBEL
Rosario Experimental and Behavioral Economics Lab

Why Replicability Crisis? General Problems To Economics (Science)

- Forking paths: flexibility in data analysis
  - "P-hacking", data mining, fishing
- Data selection
- *HARKing*
  - Hypothesizing after results are known
- Multiple hypothesis testing

- Small samples, low power
  - Button et al., 2013
- Publication bias
  - Low incentives for replication
- Selective reporting (drawer-file problem)
- Fraud

REBEL
Rosario Experimental and Behavioral Economics Lab

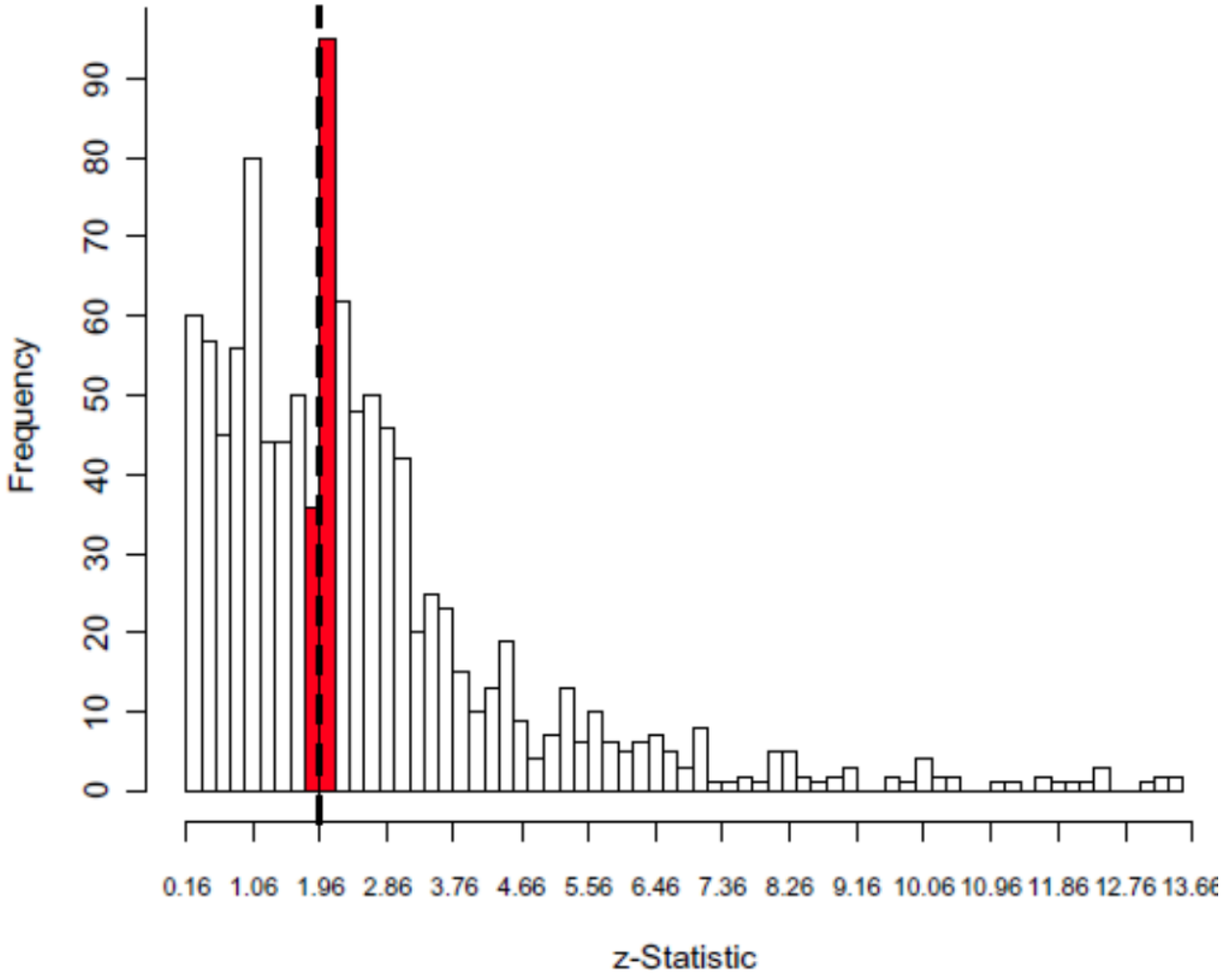
Forking paths: flexibility in data analysis

- "The econometric art as it is practiced at the computer terminal involves fitting many, perhaps thousands, of statistical models. One or several that the researcher finds pleasing are selected for reporting purposes. This searching for a model is often well intentioned, but there can be no doubt that such a specification search invalidates the traditional theories of inference. The concepts of unbiasedness, consistency, efficiency, maximum-likelihood estimation, in fact, all the concepts of traditional theory, utterly lose their meaning by the time an applied researcher pulls from the bramble of computer output the one thorn of a model he likes best, the one he chooses to portray as a rose."
    - Leamer, E. (1983, p. 36-37) “Let's take the con out of econometrics” AER 73(1)

Rosario Experimental and Behavioral Economics Lab

Gerber and Malhotra. 2008. “Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals.” *QJPS*

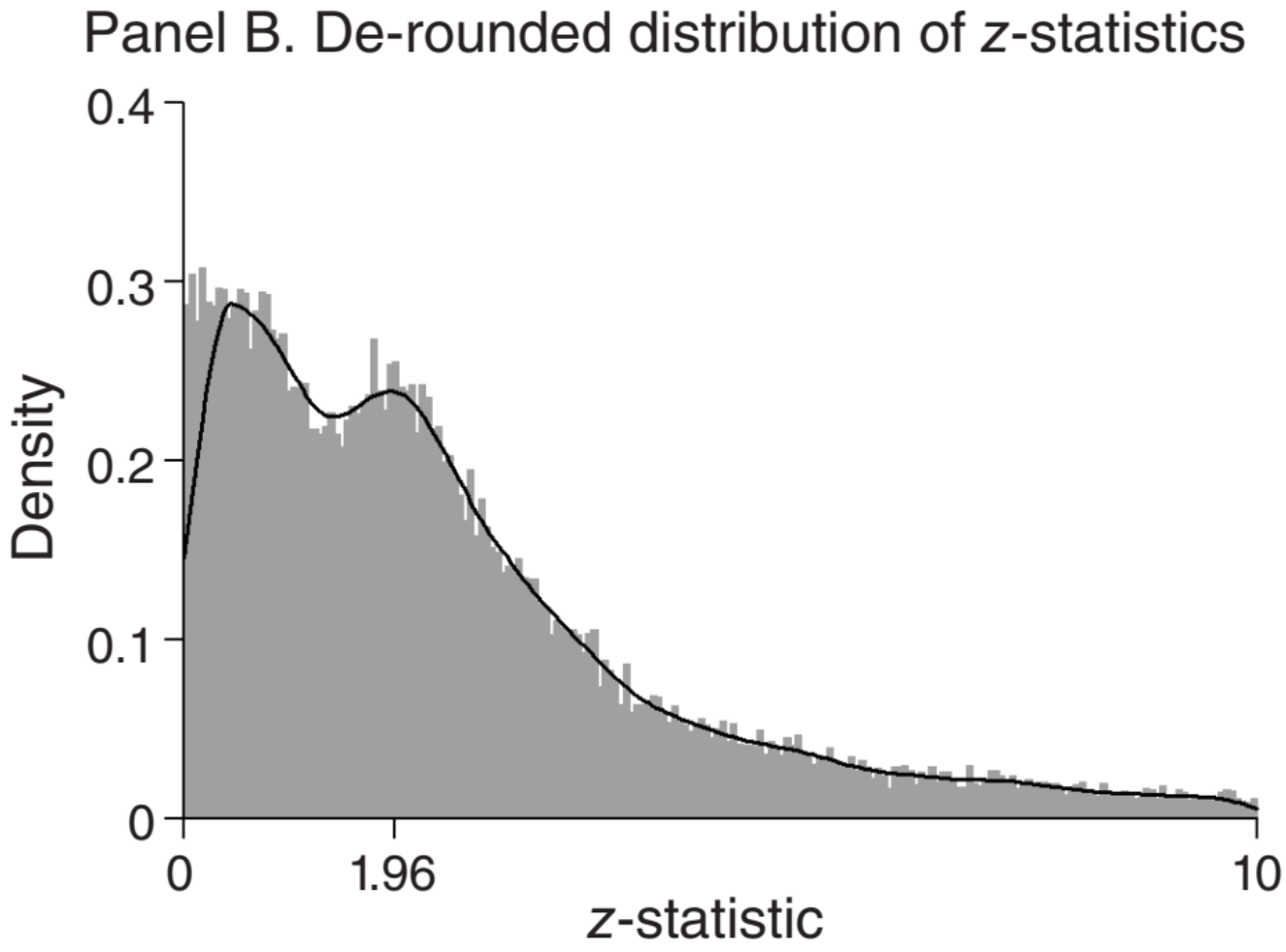
P-hacking, forking paths & flexibility in data analysis



REBEL
Rosario Experimental and Behavioral Economics Lab

Brodeur, et al (2016). “Star Wars: The Empirics Strike Back.” *AEJ Applied*

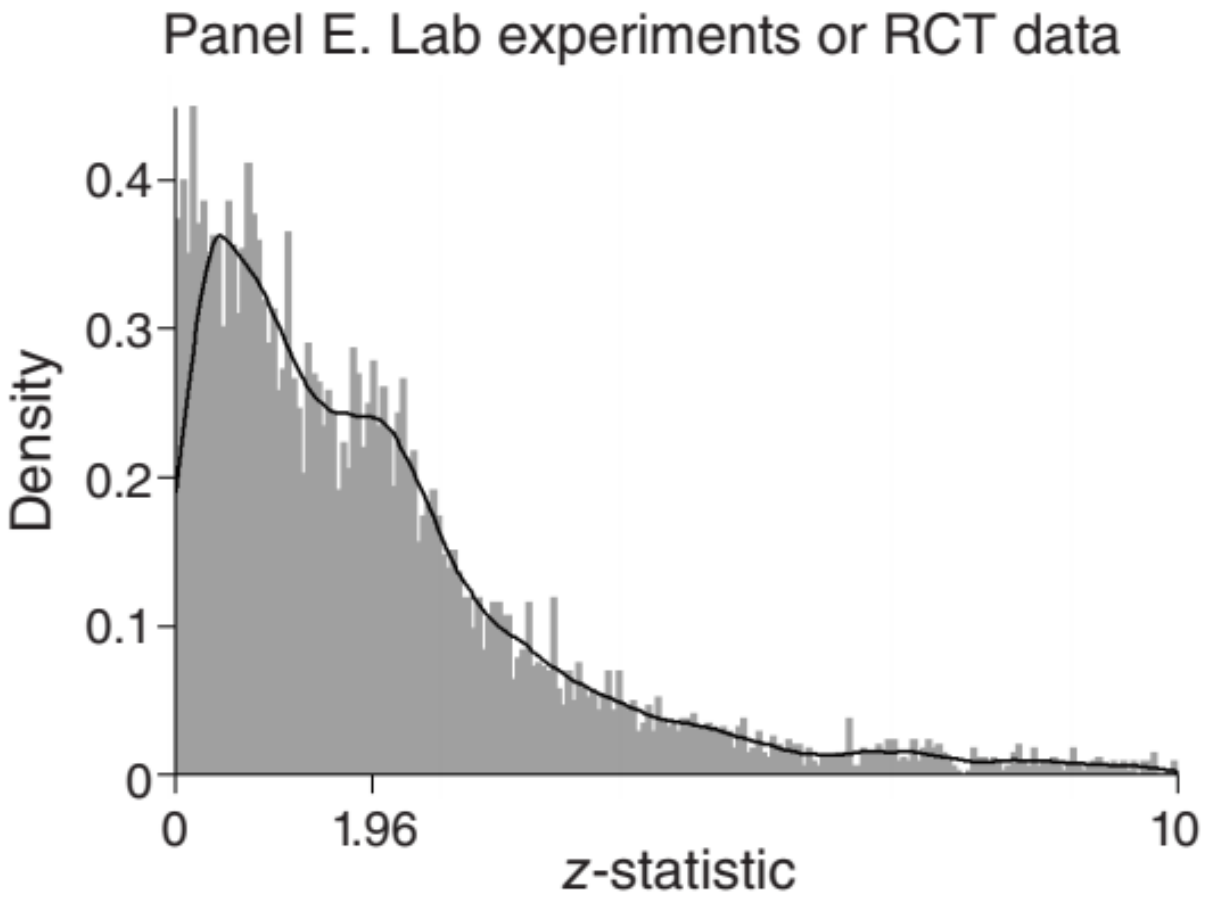
P-hacking, forking paths & flexibility in data analysis

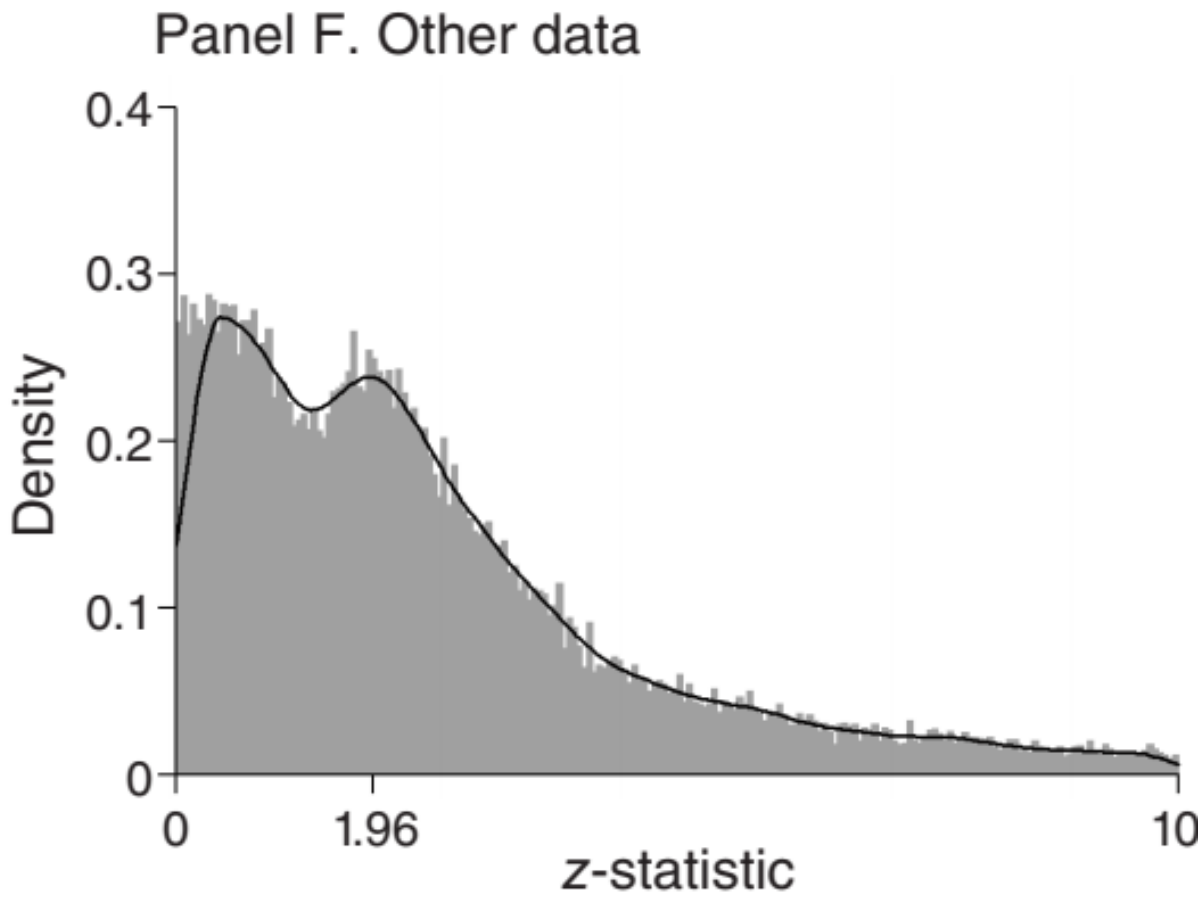


REBEL
Rosario Experimental and Behavioral Economics Lab

Brodeur, et al (2016). "Star Wars: The Empirics Strike Back." *AEJ Applied*

P-hacking: Experimental Economics

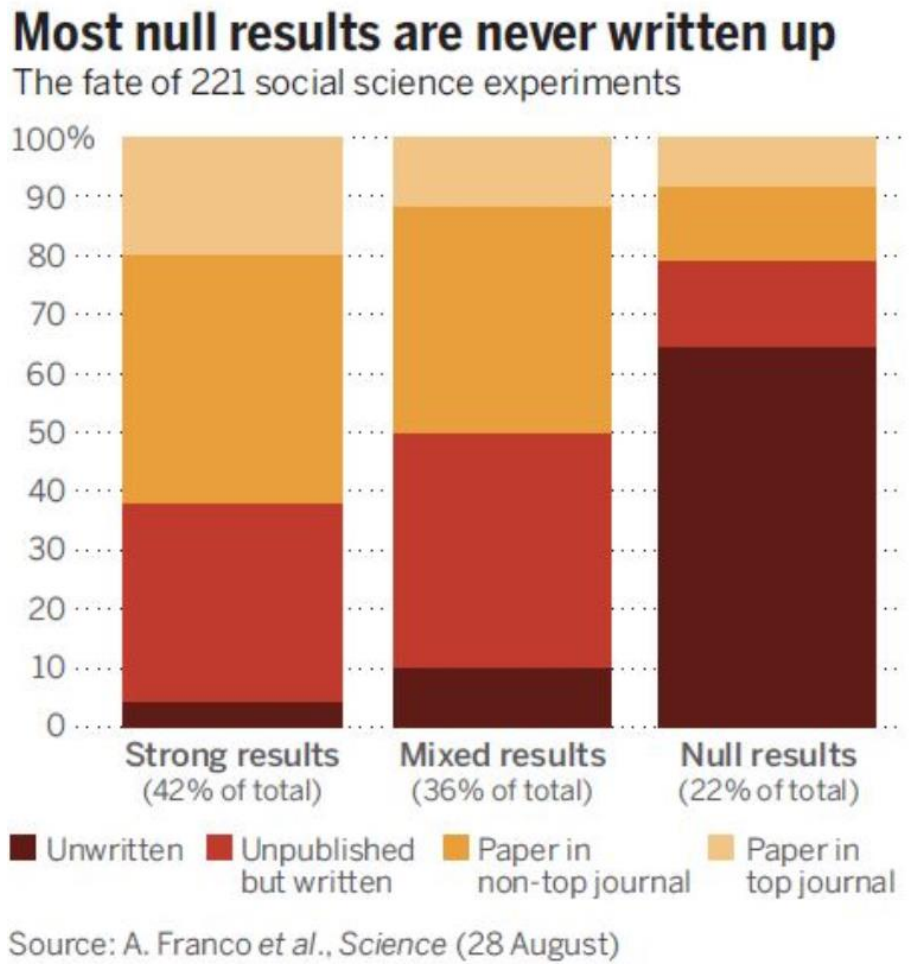


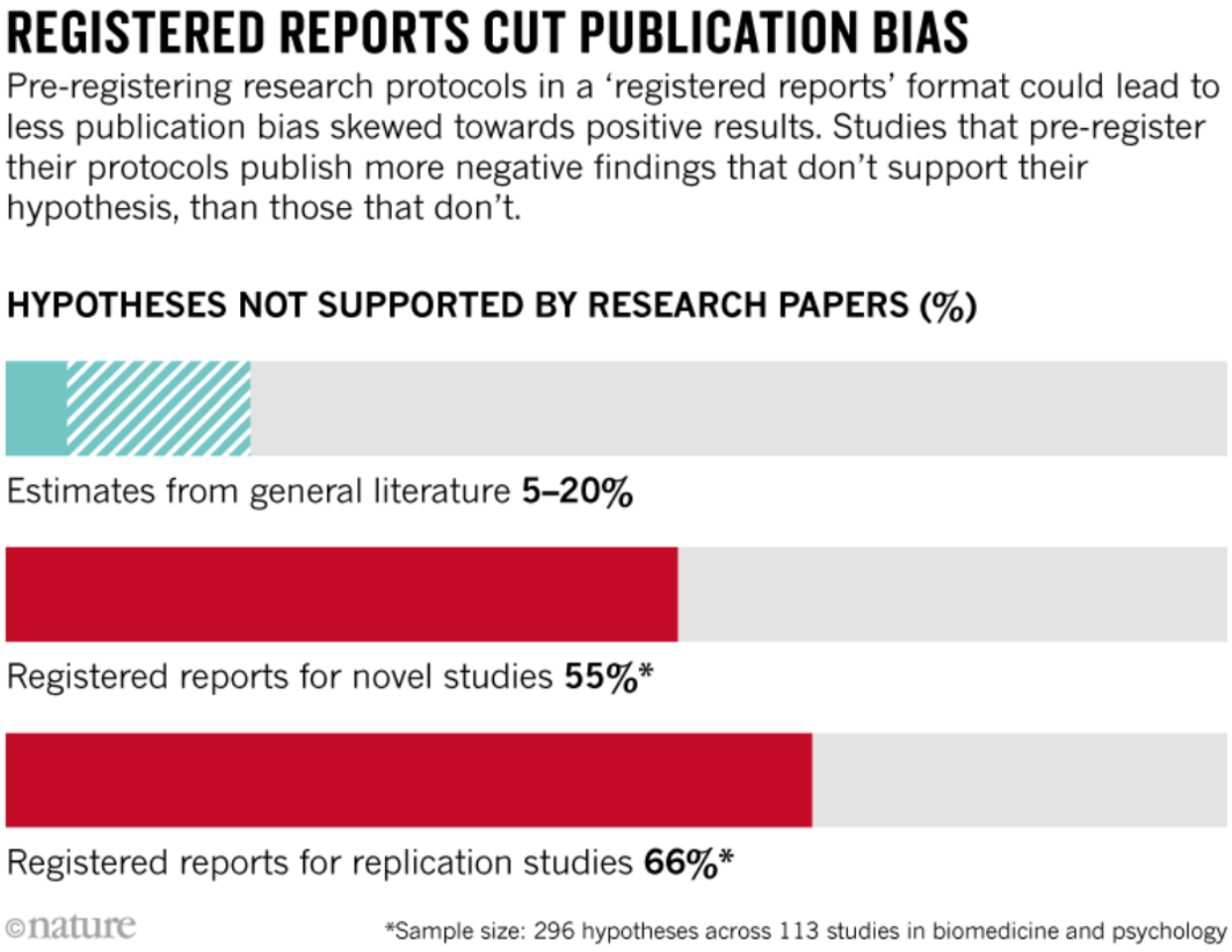


REBEL
Rosario Experimental and Behavioral Economics Lab

Franco, Malhotra, and Simonovits. (2014) "Publication Bias in the Social Sciences: Unlocking the File Drawer." *Science*

Publication Bias





REBEL
Rosario Experimental and Behavioral Economics Lab

Wouldn't more replications fixe the problems?

- Some argue none of this is really needed, all we need is more replications
  - Oversimplification?

- Potential problems:
  1. Editors: concerned with per-page citations, replications are perceived to perform less well than original studies.
     - editorial costs of allowing replication may be heavy when controversy between authors ensues.
  2. Original authors: disincentives to cooperate with scholars who wish to replicate their studies.
     - providing data and code as having a potential downside and very little upside
  3. Would-be replicators: time to undertake the replication vs. the likelihood of being published.
     - May be concerned about the implication of lack of originality,
     - Getting a reputation of having an unfavorable personality, or advancing themselves at the expense of more established authors.
- Are these problems empirically relevant?

REBEL
Rosario Experimental and Behavioral Economics Lab

Incentives for Replication in Economics

- Editors: “While all editors responded they would in principle publish a replication study that overturned the results of the original study, only 29% responded that they would consider publishing a replication study that confirmed the original study results. ”
  - Galiani et al. (2017)

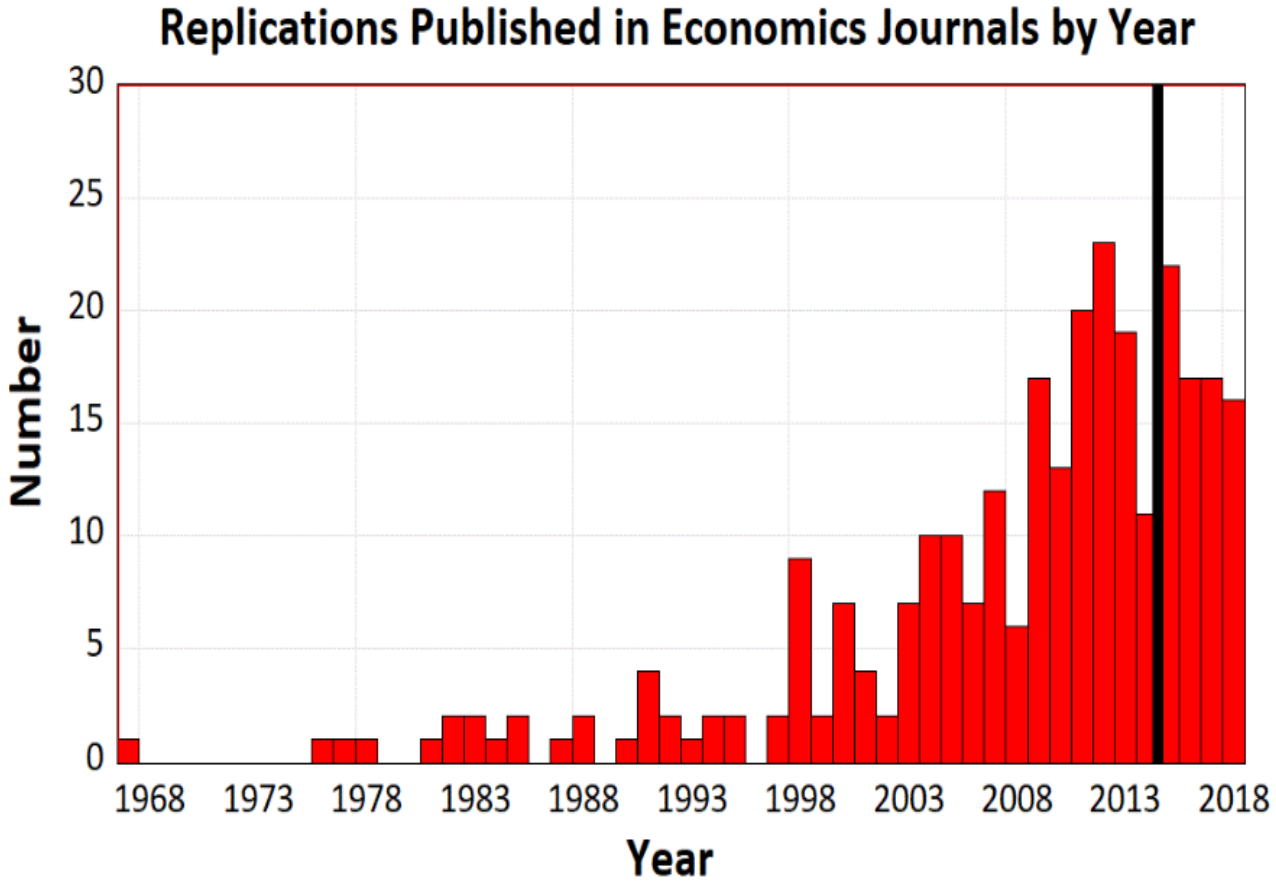
- Original Authors: It is an issue in empirical economics: 14% of 203 papers according to Galiani et al. (2018)
  - Does not seem to have been a big issue in experimental econ for Camerer et al (2016) and for Camerer et al. (2018)

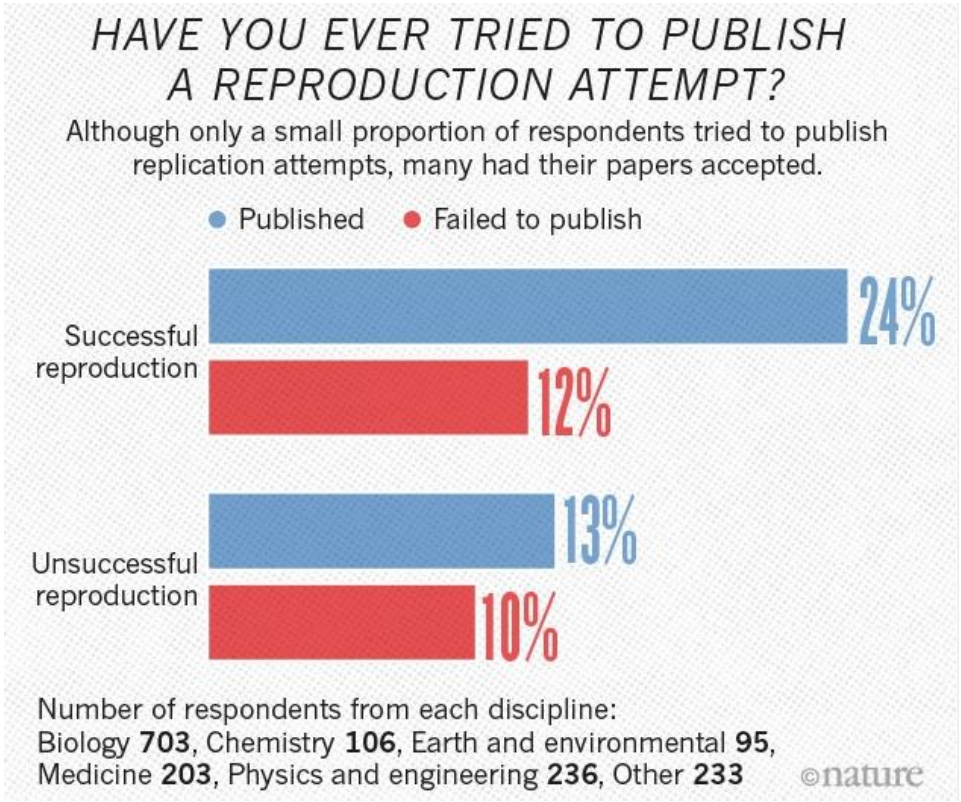
- [Potential] Replicators: “Of the 27 studies commissioned, 20 were completed, and 7 (35%) overturned some of the original results; i.e., claimed to be not able to fully replicate the original article. Only 1 was published in a peer reviewed journal and it overturned the results from the original paper.”
  - Galiani et al. (2017)

Rosario Experimental and Behavioral Economics Lab

Duvendack, Palmer-Jones, and Reed (2018) "An Update on the Progress of Replications in Economics" - https://replicationnetwork.com/2018/10/31/reed-an-update-on-the-progress-of-replications-in-economics/

Estimates of published replications





REBEL
Rosario Experimental and Behavioral Economics Lab

Incentives for replication in ExEcon?

- Replications almost omnipresent in experimental economics, implicitly.
  - Baseline treatments commonly replicate results from previous experiments
    - E.g. double auctions, asset markets, UG, DG, TG, VCM, cheating (die-rolling task), etc.

- Direct and stand alone replications far less common in experiments:
  - ExEcon (Duvendack, et al. 2015*): 9 replications (7 direct) between 1988-2018
  - JESA: 4 direct replications between 2015-2018
  - Deck, et al. (ed.) 2015. Replication in Experimental Economics (Research in Experimental Economics, Vol 18): 7 (2 direct, 5 as basis for extension)
    - "it was very difficult to get direct replications"
  - James Friedman during his dissertation stage at Yale in 1962: "Martin [Shubik] urged me to just replicate the oligopoly experiments in Bargaining Behavior, but I was unwilling to follow his scientifically sound advice because my work would then be insufficiently ´original'." --Smith (1992, p. 256)

Roth (1994) “Lets Keep the Con out of Experimental Econ.: A Methodological Note” *Empirical Economics*

Problems specific to Experimental Economics

- Planned experimental design vs. Independent trials
  - What is an experiment?
  - What is a pilot?

- “the questions of what experiments to conduct, and which experiments to report together and which separately (and which not at all) are questions of art, as opposed to matters of clearly defined practice.” (p. 286)
  - “while there can be very good reasons to carefully select experimental tasks and conditions through search or other means, the manner in which this selection is carried out is a reportable part of the experiment.” (p. 273)

REBEL
Rosario Experimental and Behavioral Economics Lab

Problems specific to Experimental Economics

- Transparency and disclosure of “failed” experiments
  - “There is room for us all to do a better job reporting what kind of pilot experiments we have conducted, and how they may have influenced the design choices made in the experiments from which the reported data were gathered. It may not always be possible or desirable to conduct pre-planned experimental designs, as the results in an early cell of an experiment may call for a change in plans. But if the process by which the data are collected is reported, the potential for miscommunication can be reduced.” (p. 287)

REBEL
Rosario Experimental and Behavioral Economics Lab

What can we do?

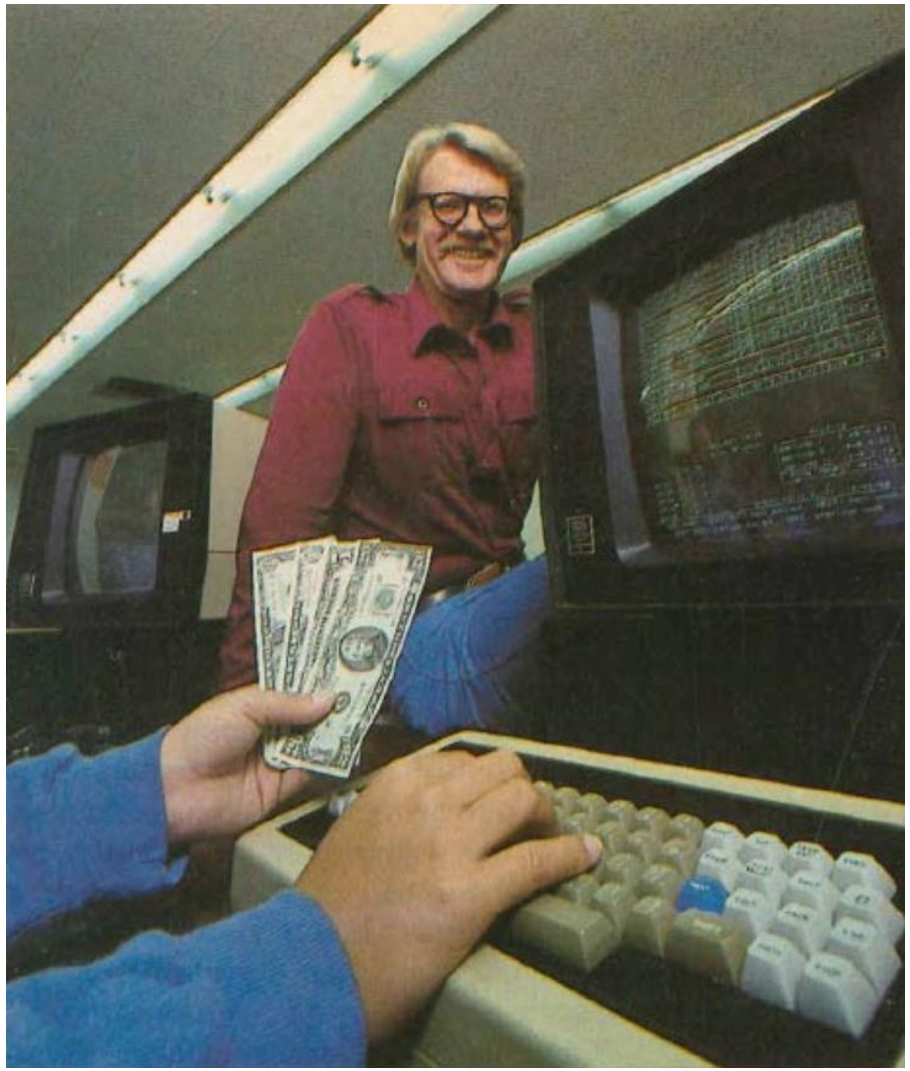
- Pre-experiment power analysis
  - Ensure powered studies
- Pre-analysis and pre-registration
  - Open Science Framework https://osf.io/
  - Note that some "policing" may be required for final papers to not deviate from the plan without explicitly stating so
- Lobby for changes in journals/associations?
  - Registered reports, pre-results review
    - JDE has currently implemented this
    - ExEcon has a special issue coming up!

Technological Advances and Experimentation

- 1960´s: The Lab!
  - Austin C. Hoggatt → the first computerized laboratory for controlled experimentation in economics, social, behavioral, and decision science
    - The Management Science Laboratory at the Center for Research in Management Science at UC Berkeley in 1964

Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
  - First computerized double auction experiments



Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
- 1990´s: (Public commercial) Internet
  - First field experiments run over the internet – D. Reiley (1999)
    - Pre WWW

Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
- 1990´s: (Public commercial) Internet
- 90´s–00's: z-Tree platform (1999) and supporting software
  - Massive reduction in costs of programming computerized experiments
    - Supporting software : ORSEE, Lab manager, E-nstructions, etc.

Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
- 1990´s: (Public commercial) Internet
- 90´s–00's: z-Tree platform (1999) and supporting software
- 1990's: Technology for fMRI
  - Neuroeconomics in 2001
    - Breiter et al. (Neuron, 2001), McCabe et al. (PNAS, 2001)

Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
- 1990´s: (Public commercial) Internet
- 90´s–00's: z-Tree platform (1999) and supporting software
- 1990's: Technology for fMRI
- 2010's: New Platforms
  - oTree, Amazon Mechanical Turk, Qualtrics, etc.
    - Reduction in costs of programming experiments that can be run on any platform, on the field and with different populations

Technological Advances and Experimentation

- 1960´s: The Lab!
- 1980´s: PLATO platform
- 1990´s: (Public commercial) Internet
- 90´s–00's: z-Tree platform (1999) and supporting software
- 1990's: Technology for fMRI
- 2010's: New Platforms (oTree, AMT, Qualtrics, etc.)
- 2010's: Combining new tolos + new methods = new possibilities
  - Choice-Process Data (beyond choice *outcomes*)
    - Eye-tracking, mouse-tracking, response time
  - Emotions: measuring and inducing
    - Skin conductance, face-reader, VR, etc.

REBEL
Rosario Experimental and Behavioral Economics Lab

Experimental Economics: Past, Present & Future

- Experiments in Economics has come a long way:
  - A time when the consensus was that experiments in economics were not possible
  - From isolated groups exploring the use of experiments (1950s y 60's)…
    - Working in parallel with no/limited knowledge of each other
      - Epicenters in Germany and the US (Tucson)
    - …to the systematic development of methods and tools for experiments (1970s y 80s)… ESA (1986), Plato, Smith (1977, 1982), ExEcon (1998)
  - …to the recognition and acceptance of experimental methods in economics (2002-2012)
  - …to the challenges (replicability) of today, and the opportunities of tomorrow

Universidad del Rosario | REBEL Rosario Experimental and Behavioral Economics Lab

Introduction: Experimental Economics –the present, the past and the future

Diego Aycinena

2020 Bogotá Experimental Economics Workshop
Universidad del Rosario, Bogotá, Colombia

This event is possible thanks to the generous support provided by our main sponsors: the *International Foundation for Research in Experimental Economics* (IFREE) and the Department of Economics at Universidad del Rosario.

ifree