

# Transparencia, Replicabilidad, y Credibilidad en la Ciencia

Haciendo Economía I  
Econ 2205

Ignacio Sarmiento-Barbieri

Universidad de los Andes

October 30, 2024

# Plan para hoy

- ▶ Ser capaz de buscar buena literatura y leer críticamente es una habilidad clave para los economistas
- ▶ Nos vamos a centrar en la búsqueda de Literatura (Tarea)
- ▶ y en el manejo de datos
  - 1 Como se suelen manipular para confundir e inducir conclusiones erróneas
  - 2 La importancia de transparencia y reproducibilidad para la credibilidad en investigación
  - 3 El material de la clase viene de [www.callingbullshit.org](http://www.callingbullshit.org) de Carl T. Bergstrom and Jevin West
- ▶ Recomendación para el tiempo libre: “How to Lie with Statistics” de Darrell Huff (1954)
- ▶ Tangencial: The Signal and the Noise: Why So Many Predictions Fail de Nate Silver (2020)

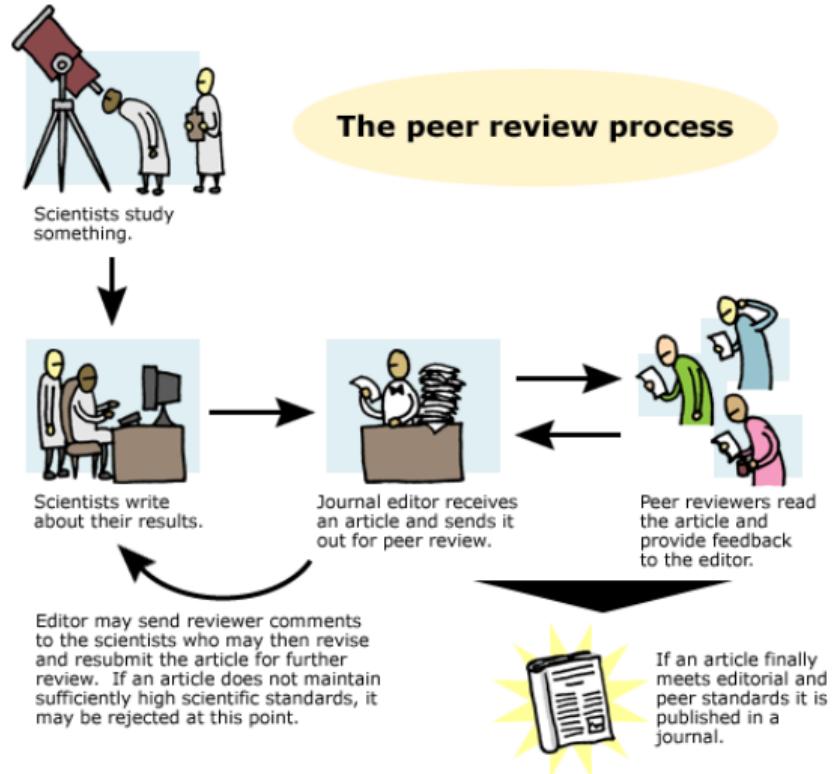
# Donde buscar literatura

## Revistas o Journals

- ▶ Como en muchas otras disciplinas, las nuevas investigaciones en economía suelen aparecer primero en revistas académicas.
- ▶ Los artículos enviados a revistas para su posible publicación generalmente son enviados por el editor a dos o tres árbitros (o revisores) para su evaluación.
- ▶ Los árbitros son expertos en el tema del trabajo.
- ▶ Para garantizar una evaluación imparcial, la mayoría de las revistas utilizan un proceso de evaluación "ciego", en el que no se informa al autor sobre la identidad de los evaluadores.
- ▶ El tiempo desde la presentación original hasta la publicación suele ser de al menos un año.

# Donde buscar literatura

## Revisión de pares



# Donde buscar literatura

## Revisión de pares

Dear Dr. Sarmiento Barbieri,

Last last year, I requested your assistance in reviewing a manuscript that I believe fell within your area of interested. Based on the reviews I received, I initially rejected the paper in its previous form and requested the authors undergo significant revisions. The authors have undertake those changes and submitted a substantially revised manuscript that addresses many of your concerns. I would like to invite you to review this revision again if you have the capacity.

The abstract for this manuscript is included below. You should treat this invitation, the manuscript and your review as confidential. You must not share your review or information about the review process with anyone without the agreement of the editors and authors involved, even after publication. This also applies to other reviewers' "comments to author" which are shared with you on decision (and vice versa).

Please respond to this invitation at your earliest opportunity.

# Donde buscar literatura

## Working Papers

- ▶ Dados los tiempos largos de publicación algunos autores publican lo que se conoce como Working Papers o Documentos de Trabajo.
- ▶ Estos son trabajos que todavía no han sido evaluados por pares expertos. Por lo tanto hay que tener cuidado y leerlos criticamente.

## Working Papers

New research by NBER affiliates, circulated for discussion and comment. The NBER distributes more than 1,200 working papers each year. These papers have not been peer reviewed. Papers issued more than 18 months ago are open access. More recent papers are available without charge to affiliates of subscribing academic institutions, employees of NBER Corporate Associates, government employees in the US, journalists, and residents of low-income countries. The NBER is working to make its web content accessible to all. Please address any concerns or requests for assistance to [webaccessibility@nber.org](mailto:webaccessibility@nber.org).

## Documentos CEDE

### BROWSE BY

[Faculties and Programs](#) [Authors](#) [Titles](#) [Theme](#) [Types of contents](#)

Search within this collection:

## Recent Submissions



"Un nuevo comienzo": El impacto del Programa Vivienda Gratuita sobre la calidad de vida de los hogares beneficiarios

Camacho-González, Adriana; Caputo-Leyva, Jorge Enrique; Sánchez-Torres Fabio (2022-04)

# Rankings e indicadores

| Ranking H5 | Journal                                      | h5-index | h5-median | Score CLm |
|------------|--|----------|-----------|-----------|
| 1          | American Economic Review                     | 158      | 246       | 98.1      |
| 2          | The Quarterly Journal of Economics           | 110      | 226       | 100       |
| 3          | The Review of Financial Studies              | 108      | 167       | 80.6      |
| 4          | Journal of Political Economy                 | 102      | 180       | 96.2      |
| 5          | The Journal of Finance                       | 99       | 182       | 72.2      |
| 6          | Journal of Economic Perspectives             | 86       | 157       | 30.5      |
| 7          | Econometrica                                 | 83       | 140       | 95.7      |
| 8          | The Review of Economic Studies               | 81       | 125       | 81        |
| 9          | Review of Economics and Statistics           | 74       | 110       | 74.1      |
| 10         | The Economic Journal                         | 71       | 112       | 64.5      |
| 11         | Journal of Public Economics                  | 69       | 99        | 62        |
| 12         | Journal of Development Economics             | 69       | 94        | 42.7      |
| 13         | Journal of International Economics           | 68       | 105       | 61.5      |
| 14         | Journal of Monetary Economics                | 63       | 100       | 75.8      |
| 15         | Economic Modelling                           | 61       | 80        | 11.3      |
| 16         | Journal of Economic Literature               | 60       | 144       | 28.3      |
| 17         | European Economic Review                     | 60       | 92        | 55.2      |
| 18         | Journal of Labor Economics                   | 59       | 119       | 55.1      |
| 19         | Journal of the European Economic Association | 58       | 82        | 57        |
| 20         | Journal of International Money and Finance   | 56       | 84        | NA        |

# Rankings e indicadores

- ▶ The h-index: It is the largest number  $h$  such that  $h$  articles published in 2015-2019 have at least  $h$  citations each. E.g. the h5-index of a journal is 158 means that 158 articles published in the journal in 2015-2019 have received at least 185 citations each article.
- ▶ Combes and Linnemer (2010) Score: This index combines (sophisticated) citation indexes, field of specialization normalized indexes, and a h-index based on Google Scholar citations.
- ▶ The Scimago Journal Rank (SJR) is based on the transfer of prestige from a journal to another one; such prestige is transferred through the references that a journal do to the rest of the journals and to itself.

1 Literatura en Economía

2 Transparencia y reproducibilidad en Economía

3 Manipular para confundir

- Manipular las gráficas
- Manipular las muestras
- Manipular cifras

4 Mensaje Final y Próxima Clase

# Transparencia y reproducibilidad en Economía

Is Economics Research Replicable?  
Sixty Published Papers from Thirteen Journals Say  
“Usually Not”

Andrew C. Chang\* and Phillip Li†

September 4, 2015

# Controversia de Reinhart & Rogoff

The New York Times

Opinion

OP-ED CONTRIBUTORS

## Debt, Growth and the Austerity Debate

By Carmen M. Reinhart and Kenneth S. Rogoff

April 25, 2013



CAMBRIDGE, Mass.

IN May 2010, we published an academic paper, [“Growth in a Time of Debt.”](#) Its main finding, drawing on data from 44 countries over 200 years, was that in both rich and developing countries, high levels of government debt — specifically, gross public debt equaling 90 percent or more of the nation’s annual economic output — was associated with notably lower rates of growth.



# Controversia de Reinhart & Rogoff

Growth in a Time of Debt

Carmen M. Reinhart and Kenneth S. Rogoff

NBER Working Paper No. 15639

January 2010, Revised January 2010

JEL No. E2,E3,E6,F3,F4,N10

## ABSTRACT

We study economic growth and inflation at different levels of government and external debt. Our analysis is based on new data on forty-four countries spanning about two hundred years. The dataset incorporates over 3,700 annual observations covering a wide range of political systems, institutions, exchange rate arrangements, and historic circumstances. Our main findings are: First, the relationship between government debt and real GDP growth is weak for debt/GDP ratios below a threshold of 90 percent of GDP. Above 90 percent, median growth rates fall by one percent, and average growth falls considerably more. We find that the threshold for public debt is similar in advanced and emerging economies. Second, emerging markets face lower thresholds for external debt (public and private)—which is usually denominated in a foreign currency. When external debt reaches 60 percent of GDP, annual growth declines by about two percent; for higher levels, growth rates are roughly cut in half. Third, there is no apparent contemporaneous link between inflation and public debt levels for the advanced countries as a group (some countries, such as the United States, have experienced higher inflation when debt/GDP is high). The story is entirely different for emerging markets, where inflation rises sharply as debt increases.

# Controversia de Reinhart & Rogoff

## Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff

Thomas Herndon, Michael Ash and Robert Pollin\*

We replicate Reinhart and Rogoff (2010A and 2010B) and find that selective exclusion of available data, coding errors and inappropriate weighting of summary statistics lead to serious miscalculations that inaccurately represent the relationship between public debt and GDP growth among 20 advanced economies. Over 1946–2009, countries with public debt/GDP ratios above 90% averaged 2.2% real annual GDP growth, not –0.1% as published. The published results for (i) median GDP growth rates for the 1946–2009 period and (ii) mean and median GDP growth figures over 1790–2009 are all distorted by similar methodological errors, although the magnitudes of the distortions are somewhat smaller than with the mean figures for 1946–2009. Contrary to Reinhart and Rogoff's broader contentions, both mean and median GDP growth when public debt levels exceed 90% of GDP are not dramatically different from when the public debt/GDP ratios are lower. The relationship between public debt and GDP growth varies significantly by period and country. Our overall evidence refutes RR's claim that public debt/GDP ratios above 90% consistently reduce a country's GDP growth.

*Key words:* Public debt, Austerity

*JEL classifications:* E60, E62, E65

# Incapacidad para replicar resultados

|  | Number of articles (requests) | Attempted reprod. | Successful reprod. | Reprod. rate   attempted | Reprod. rate   empirical article |
|--|-------------------------------|-------------------|--------------------|--------------------------|----------------------------------|
| Dewald et al. (1986)<br>before policy change | 62                            | 5                 | 3                  | 60.0%                    | 4.8%                             |
| Dewald et al. (1986)<br>after policy change  | 92                            | 3                 | 2                  | 66.7%                    | 2.2%                             |
| McCullough et al. (2006)                     | 193                           | 62                | 14                 | 22.6%                    | 7.3%                             |
| Chang & Li (2017)                            | 67                            | 59                | 29                 | 49.2%                    | 43.3%                            |

Source: <https://hdsr.mitpress.mit.edu/pub/fgpmpj11/release/3>

# Incapacidad para replicar resultados

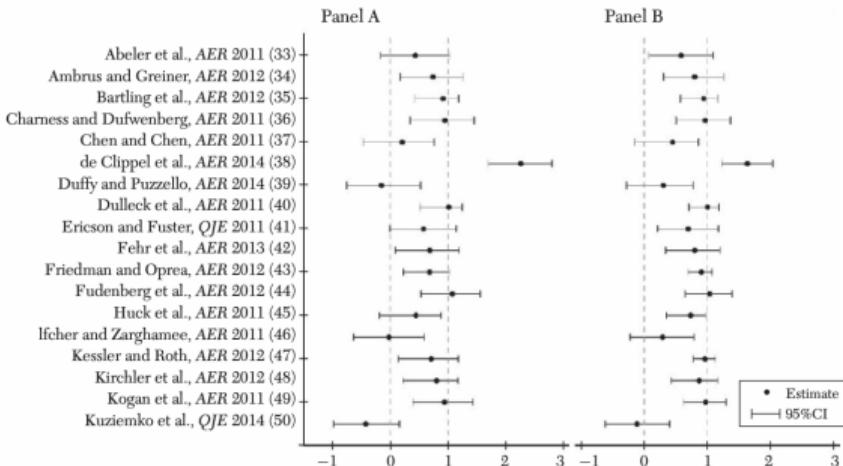


Figure 3. Replicability in Experimental Economics

*Notes:* Figure from Camerer et al. (2016). Reprinted with permission from AAAS. Panel A: Plotted are 95% CIs of replication effect sizes (standardized to correlation coefficients). The standardized effect sizes are normalized so that 1 equals the original effect size (fig. S1 in Camerer et al. 2016 shows a non-normalized version). Eleven replications have a significant effect in the same direction as in the original study [61.1%; 95% CI = (36.2%, 86.1%)]. The 95% CI of the replication effect size includes the original effect size for twelve replications [66.7%; 95% CI = (42.5%, 90.8%)]; if one also includes the study in which the entire 95% CI exceeds the original effect size, this increases to thirteen replications [72.2%; 95% CI = (49.3%, 95.1%)]. AER denotes the *American Economic Review* and QJE denotes the *Quarterly Journal of Economics*. Panel B: Meta-analytic estimates of effect sizes, combining the original and replication studies. Plotted are 95% CIs of combined effect sizes (standardized to correlation coefficients). The standardized effect sizes are normalized as in panel A (where again fig. S1 shows a non-normalized version). Fourteen studies have a significant effect in the same direction as the original study in the meta-analysis [77.8%; 95% CI = (56.5%, 99.1%)].

# Incapacidad para replicar resultados

## [98] Evidence of Fraud in an Influential Field Experiment About Dishonesty

---

Posted on August 17, 2021 by Uri, Joe, & Leif

This post is co-authored with a team of researchers who have chosen to remain anonymous. They uncovered most of the evidence reported in this post. These researchers are not connected in any way to the papers described herein.

Source: <http://datacolada.org/98>

# Incapacidad para replicar resultados

In 2012, Shu, Mazar, Gino, Ariely, and Bazerman published a three-study paper in PNAS ([.htm](#)) reporting that dishonesty can be reduced by asking people to sign a statement of honest intent *before* providing information (i.e., at the top of a document) rather than *after* providing information (i.e., at the bottom of a document). In 2020, Kristal, Whillans, and the five original authors published a follow-up in PNAS entitled, "Signing at the beginning versus at the end does not decrease dishonesty" ([.htm](#)). They reported six studies that failed to replicate the two original lab studies, including one attempt at a direct replication and five attempts at conceptual replications.

Our focus here is on Study 3 in the 2012 paper, a field experiment ( $N = 13,488$ ) conducted by an auto insurance company in the southeastern United States under the supervision of the fourth author. Customers were asked to report the current odometer reading of up to four cars covered by their policy. They were randomly assigned to sign a statement indicating, "I promise that the information I am providing is true" either at the top or bottom of the form. Customers assigned to the 'sign-at-the-top' condition reported driving 2,400 more miles (10.3%) than those assigned to the 'sign-at-the-bottom' condition.

Source: <http://datacolada.org/98>

# Evidence of Fraud

This is Table 1 in Kristal et al. (2020), reporting their re-analysis of Shu et al. (2012)

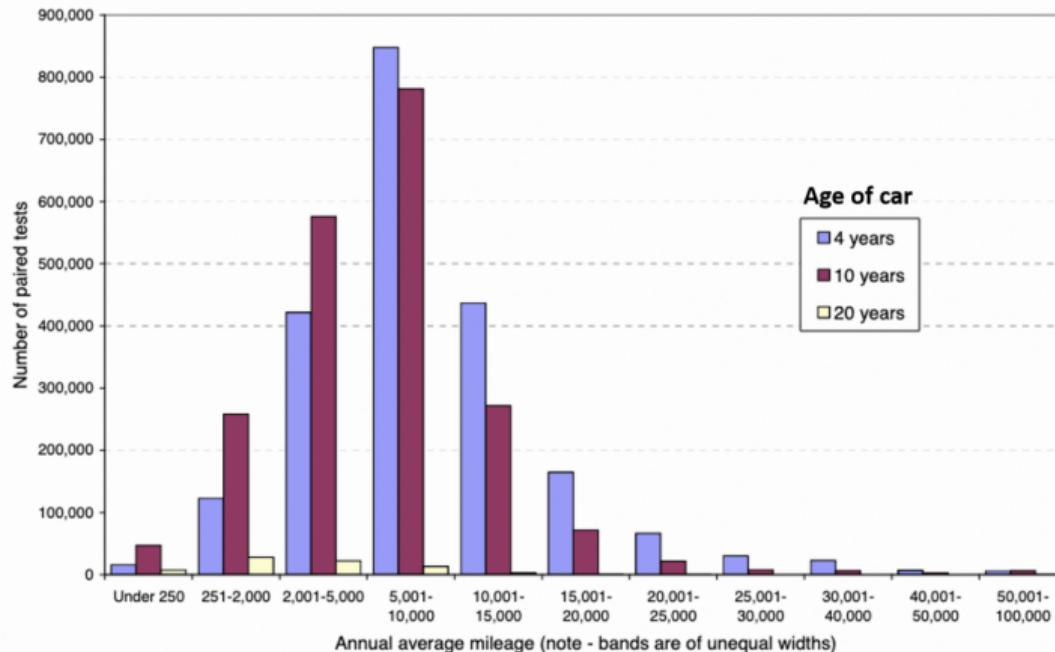
|   | Sign-at-the-bottom,<br>means (SD) | Sign-at-the-top,<br>means (SD) | Two-sided <i>t</i> test, values      |
|---|-----------------------------------|--------------------------------|--------------------------------------|
| Baseline odometer reading ( <i>t</i> 0)   | 75,034.50 (50,265.35)             | 59,692.71 (49,953.51)          | $t_{(13,474)} = 17.78, P < 0.0001$   |
| New odometer reading ( <i>t</i> 1)  | 98,705.14 (51,934.76)             | 85,791.10 (51,701.31)          | $t_{(13,475)} = 14.47, P < 0.0001$   |
| Difference in odometer<br>readings; i.e., miles driven ( <i>t</i> 1– <i>t</i> 0)* | 23,670.64 (12,621.38)             | 26,098.40 (12,253.37)          | $t_{(13,448)} = -11.331, P < 0.0001$ |

\*This row was the outcome reported in the original paper.

# Evidence of Fraud

Figure from a UK Department of Transportation Report on Distribution of Yearly Miles Driven in 2010

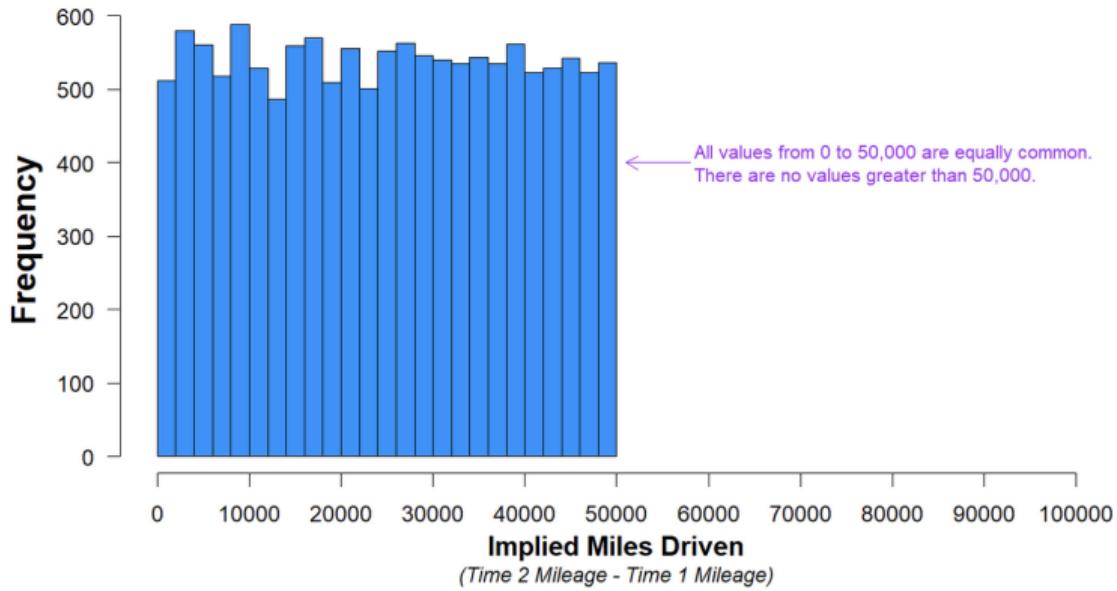
Source: <https://bit.ly/3jwSP2N>



Source: <http://datacolada.org/98>

# Evidence of Fraud

Figure 1. Histogram of Miles Driven - Car #1 (N=13,488)



Source: <http://datacolada.org/98>

# Evidence of Fraud

|    | A           | B  | C       | D             | E           |
|----|-------------|----|---------|---------------|-------------|
| 1  | condition   | id | font    | baseline_car1 | update_car1 |
| 2  | Sign Top    | 1  | Cambria | 896           | 39198       |
| 3  | Sign Bottom | 2  | Cambria | 21396         | 63511       |
| 4  | Sign Bottom | 3  | Cambria | 21340         | 37460       |
| 5  | Sign Bottom | 4  | Cambria | 23912         | 59136       |
| 6  | Sign Bottom | 5  | Calibri | 16862         | 59292       |
| 7  | Sign Top    | 6  | Calibri | 147738        | 167895      |
| 8  | Sign Bottom | 7  | Calibri | 18780         | 49811       |
| 9  | Sign Top    | 8  | Calibri | 41930         | 80323       |
| 10 | Sign Top    | 9  | Cambria | 28993         | 63707       |
| 11 | Sign Bottom | 10 | Calibri | 78382         | 127817      |
| 12 | Sign Top    | 11 | Calibri | 58500         | 81081       |

Source: <http://datacolada.org/98>

# Evidence of Fraud

|    | A           | B    | C       | D             | E             | F             | G             |
|----|-------------|------|---------|---------------|---------------|---------------|---------------|
| 1  | condition   | id   | font    | baseline_car1 | baseline_car2 | baseline_car3 | baseline_car4 |
| 10 | Sign Bottom | 5938 | Calibri | 49675         | 17709         | 27357         | 64428         |
| 11 | Sign Bottom | 1137 | Cambria | 50350         | 18421         | 27714         | 64784         |

Cambria is 675 miles more than Calibri

Cambria is 712 miles more than Calibri

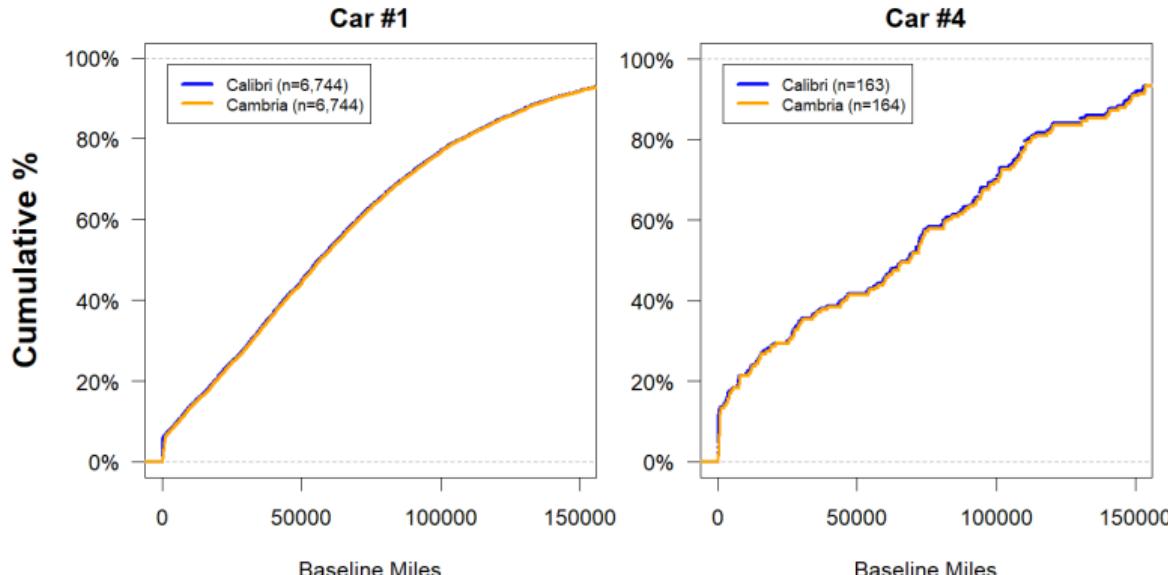
Cambria is 357 miles more than Calibri

Cambria is 356 miles more than Calibri

Source: <http://datacolada.org/98>

# Evidence of Fraud

Figure 5. CDFs of Miles Driven for Rows in Calibri vs Cambria Font



1 Literatura en Economía

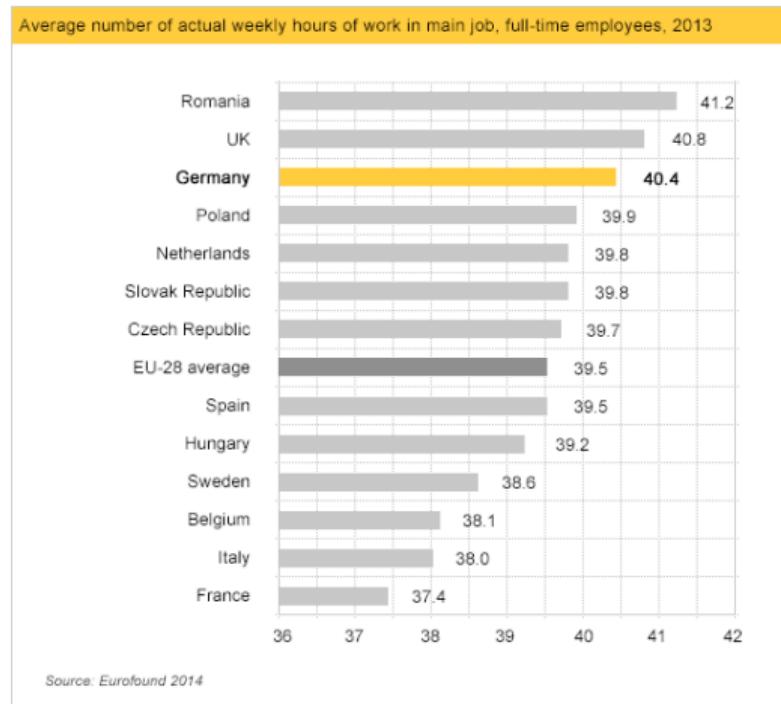
2 Transparencia y reproducibilidad en Economía

3 Manipular para confundir

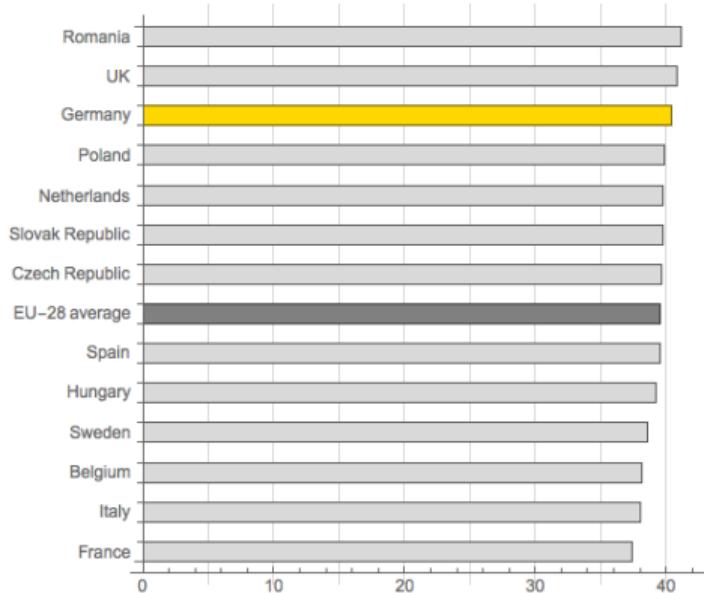
- Manipular las gráficas
- Manipular las muestras
- Manipular cifras

4 Mensaje Final y Próxima Clase

# Manipular las gráficas

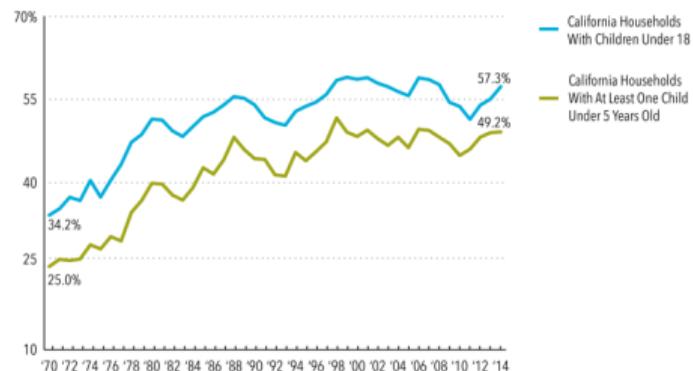


# Manipular las gráficas



# Manipular las gráficas

## More California Households Have All Parents Working, Making Access to Child Care an Important Priority Percentage of California Households Where All Parents Work, 1970 to 2014

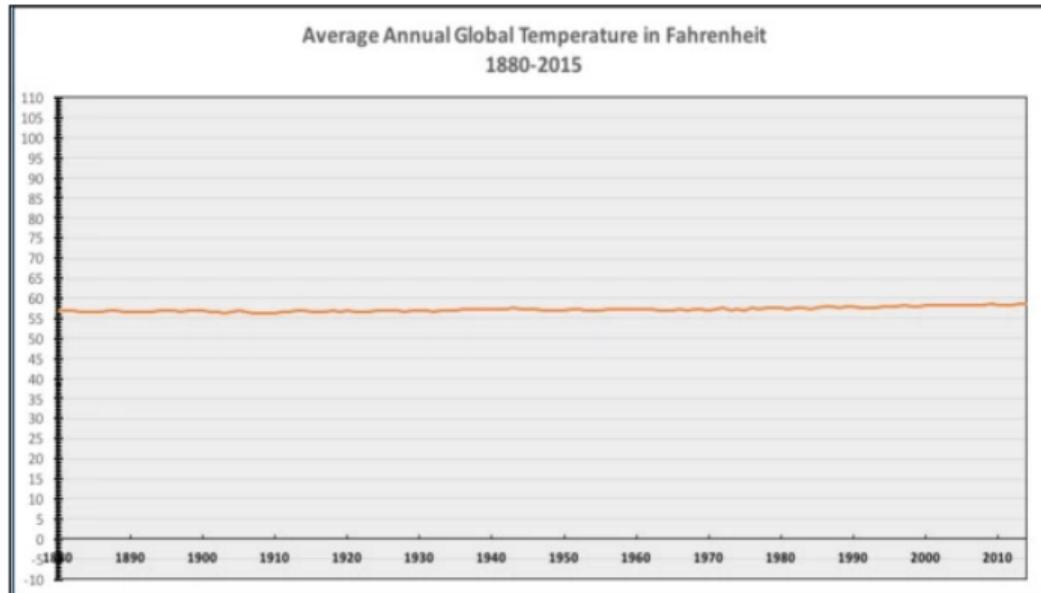


Note: A "household where all parents work" includes single-parent households and dual-earner households. Parents include stepparents and adoptive parents.  
Source: Budget Center analysis of US Census Bureau data



California Budget  
& Policy Center  
Independent Analysis. Shared Prosperity.

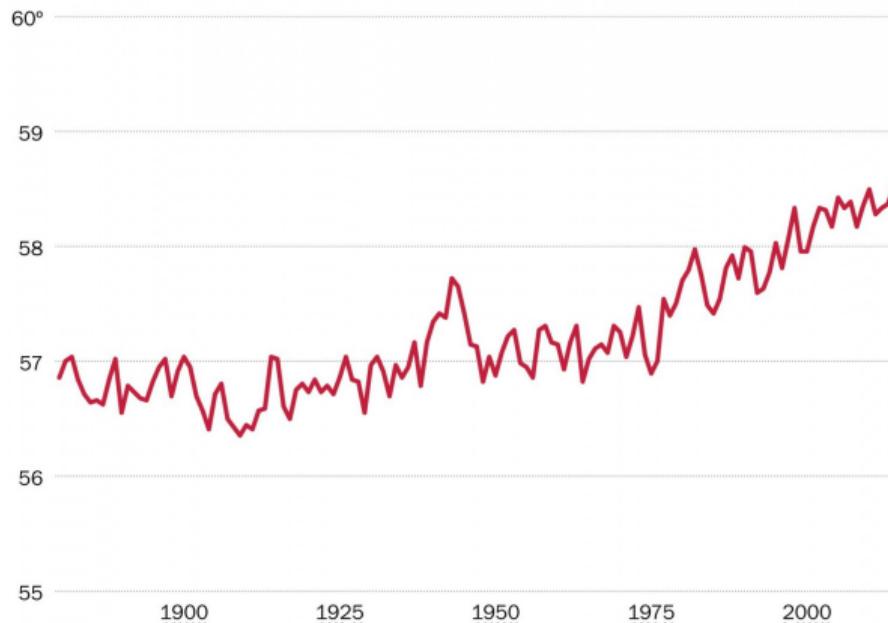
# Manipular las gráficas



# Manipular las gráficas

Average global temperature by year

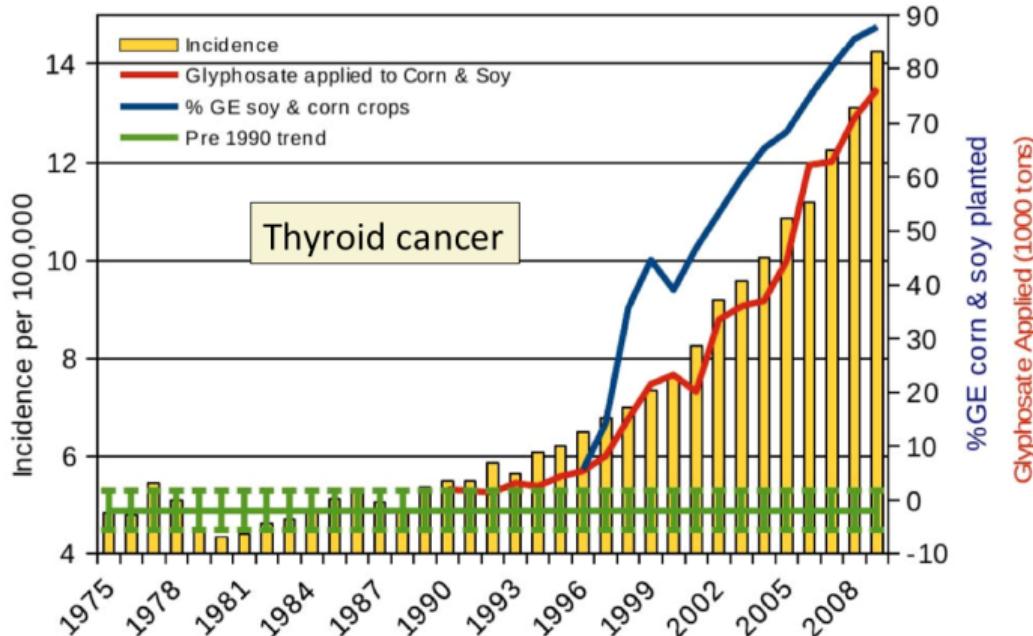
Data from NASA/GISS.



# Manipular las gráficas

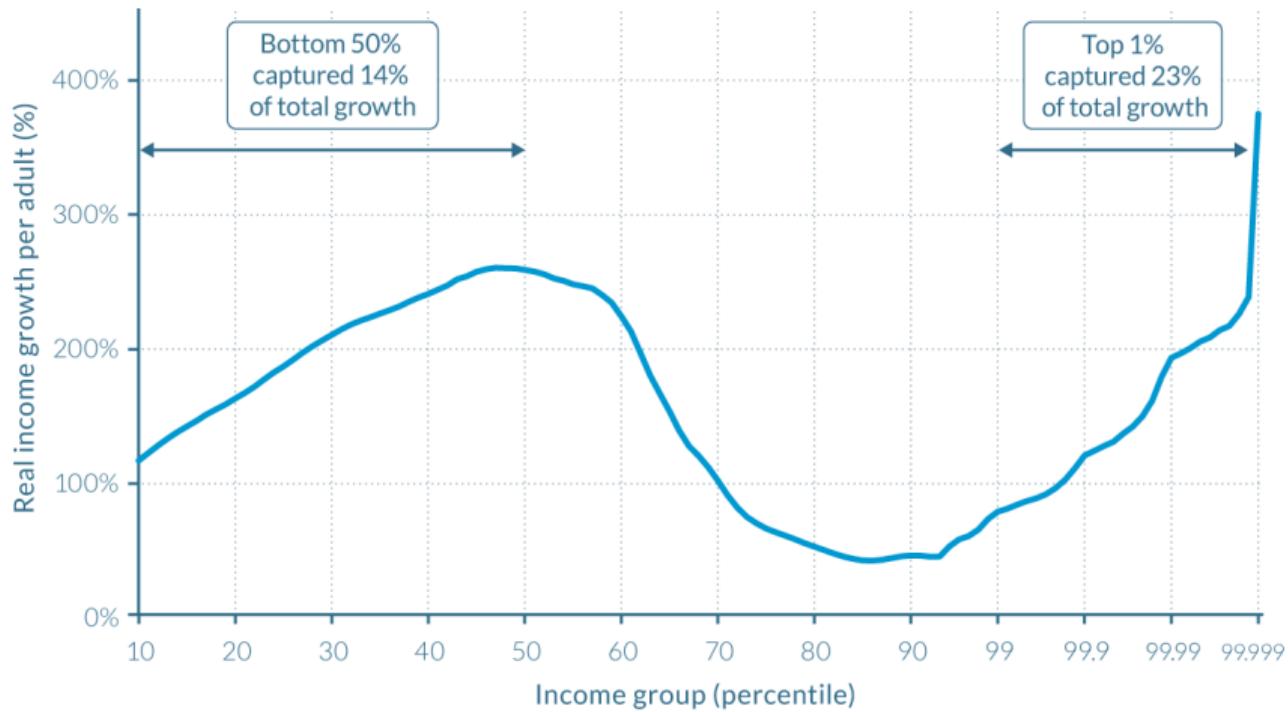
## Thyroid Cancer Incidence Rate (age adjusted)

plotted against glyphosate applied to U.S. corn & soy ( $R = 0.988$ ,  $p \leq 7.612e-09$ )  
along with %GE corn & soy crops  $R = 0.9377$ ,  $p \leq 2.152e-05$   
sources: USDA:NASS; SEER



\*Figure 10, Swanson et al. Journal of Organic Systems 2014; 9(2):6-37.

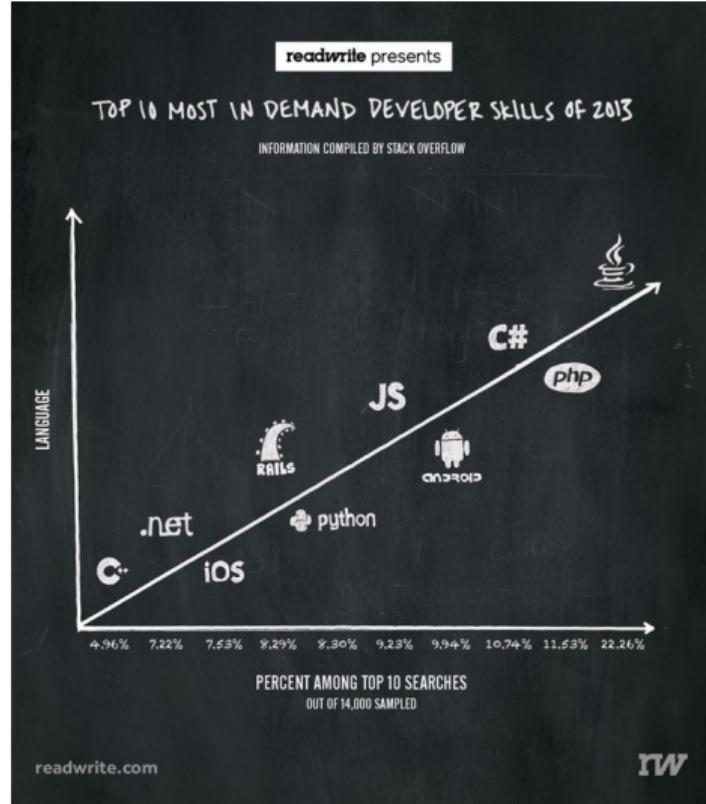
# Manipular las gráficas



Source: WID.world (2017). See [wir2018.wid.world/methodology.html](http://wir2018.wid.world/methodology.html) for data series and notes.

On the horizontal axis, the world population is divided into a hundred groups of equal population size and sorted in ascending order from left to right, according to each group's income level. The Top 1% group is divided into ten groups, the richest of these groups is also divided into ten groups, and the very top group is again divided into ten groups of equal population size. The vertical axis shows the total income growth of an average individual in each group between 1980 and 2016. For percentile group p99p99.1 (the poorest 10% among the world's richest 1%), growth was 77% between 1980 and 2016. The Top 1% captured 23% of total growth over this period. Income estimates account for differences in the cost of living between countries. Values are net of inflation.

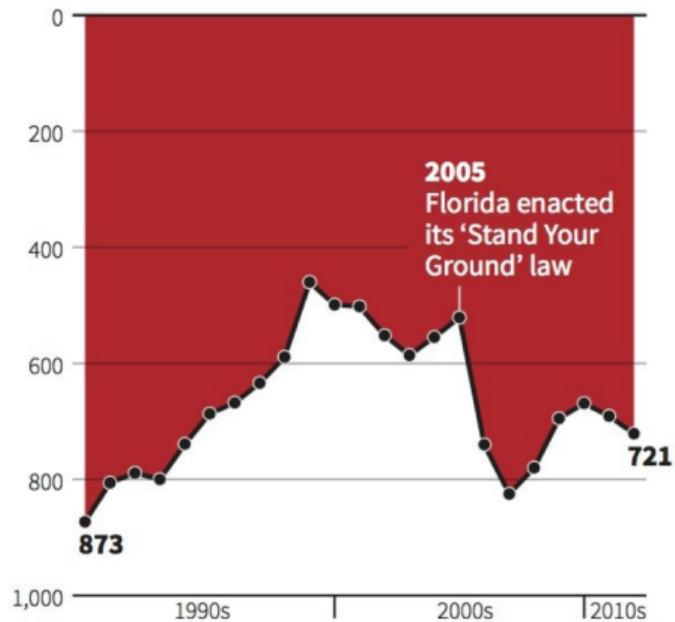
# Manipular las gráficas



# Manipular las gráficas

## Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

1 Literatura en Economía

2 Transparencia y reproducibilidad en Economía

3 Manipular para confundir

- Manipular las gráficas
- **Manipular las muestras**
- Manipular cifras

4 Mensaje Final y Próxima Clase

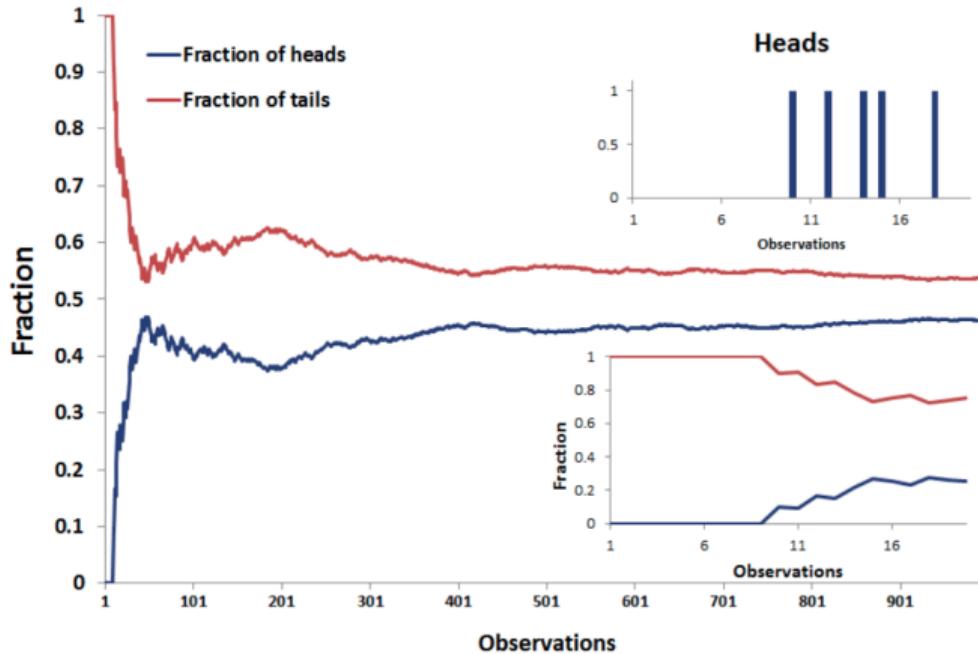
# Podemos identificar el problema

Cierta ciudad cuenta con dos hospitales. En el hospital más grande, nacen alrededor de 45 bebés cada día, y en el hospital más pequeño, nacen alrededor de 15 bebés cada día. Como saben, alrededor del 50% de todos los bebés son varones. Sin embargo, el porcentaje exacto varía de un día a otro. A veces puede ser superior al 50%, a veces inferior.

Durante un período de 1 año, cada hospital registró los días en los que más del 60% de los bebés nacidos eran varones. ¿Qué hospital crees que registró más días de este tipo?

- 1 El hospital más grande
- 2 El hospital más pequeño
- 3 Más o menos iguales (es decir, dentro del 5% entre sí)

# Muestras Pequeñas



Source: <https://towardsdatascience.com/lessons-from-how-to-lie-with-statistics-57060c0d2f19>

# The Literary Digest

NEW YORK

OCTOBER 31, 1936

## *Topics of the day*

**LANDON, 1,293,669; ROOSEVELT, 972,897**

Final Returns in The Digest's Poll of Ten Million Voters

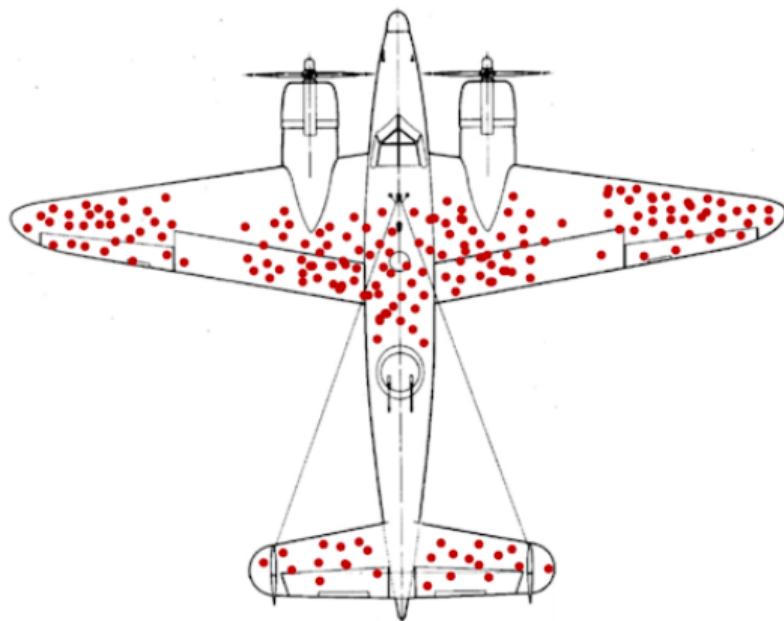
Well, the great battle of the ballots in the Poll of ten million voters, scattered throughout the forty-eight States of the

litan National Committee purchased THE LITERARY DIGEST?" And all types and varieties, including: "Have the Jews purchased

returned and let the people of the Nation draw their conclusions as to our accuracy. So far, we have been right in every Poll. Will we be right in the current Poll? That, as Mrs. Roosevelt said concerning the President's reelection, is in the 'lap of the gods.'

"We never make any claims before election but we respectfully refer you to the opinion of one of the most quoted citizens

## Muestras sesgadas



1 Literatura en Economía

2 Transparencia y reproducibilidad en Economía

3 Manipular para confundir

- Manipular las gráficas
- Manipular las muestras
- Manipular cifras

4 Mensaje Final y Próxima Clase

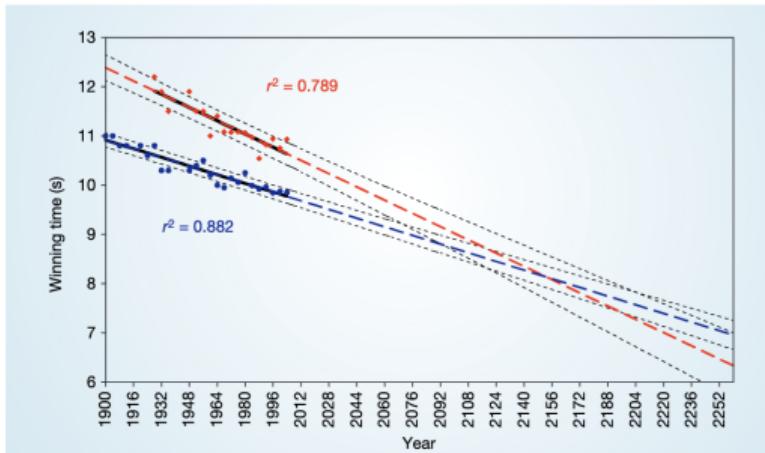
## Momentous sprint at the 2156 Olympics?

Women sprinters are closing the gap on men and may one day overtake them.

The 2004 Olympic women's 100-metre sprint champion, Yuliya Nesterenko, is assured of fame and fortune. But we show here that — if current trends continue — it is the winner of the event in the 2156 Olympics whose name will be etched in sporting history forever, because this may be the first occasion on which the race is won in a faster time than the men's event.

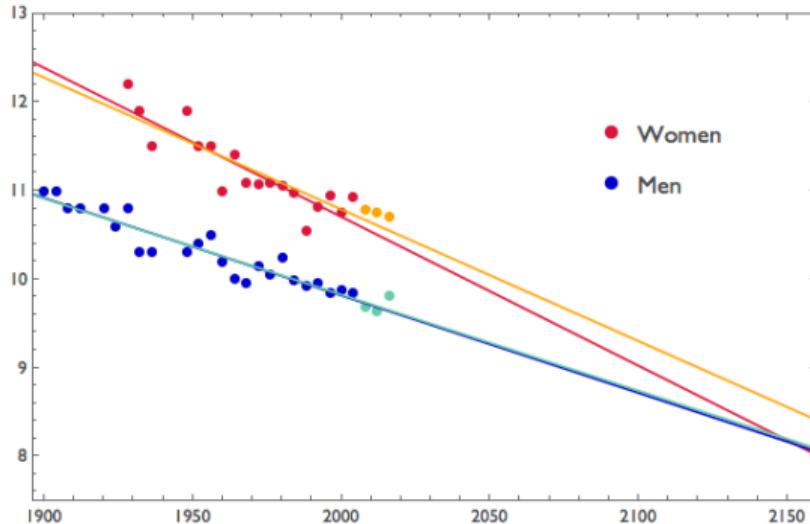
The Athens Olympic Games could be viewed as another giant experiment in human athletic achievement. Are women narrowing the gap with men, or falling further behind? Some argue that the gains made by women in running events between the 1930s and the 1980s are decreasing as the women's achievements plateau<sup>1</sup>. Others contend that there is no evidence that athletes, male or female, are reaching the limits of their potential<sup>1,2</sup>.

In a limited test, we plot the winning times of the men's and women's Olympic finals over the past 100 years (ref. 3; for data set, see supplementary information) against the compe-



**Figure 1** The winning Olympic 100-metre sprint times for men (blue points) and women (red points), with superimposed best-fit linear regression lines (solid black lines) and coefficients of determination. The regression lines are extrapolated (broken blue and red lines for men and women, respectively) and 95% confidence intervals (dotted black lines) based on the available points are superimposed. The projections intersect just before the 2156 Olympics, when the winning women's 100-metre sprint time of 8.079 s will be faster than the men's at 8.098 s.

# Extrapolaciones fuera de muestra



# Cifras absolutas para manipular importancia

## Food Stamp Fraud at All-Time High: Is It Time to End the Program?

f share this

e email



Dec 27, 2016 // 3:08pm ♥ As seen on **Fox & Friends**

**REACTION: Actress Carrie Fisher Dies at Age 60 After Heart Attack**

**Should Japanese PM Shinzo Abe Apologize for Pearl Harbor Attack?**

**Cop Whose House Burned Down Gets Christmas Surprise From Randy Travis**

Food stamp fraud is at an all-time high, with cases this year including a state lawmaker and even a millionaire.

According to the USDA, \$70 million of taxpayer money was wasted in 2016 due to food stamp fraud.

1 Literatura en Economía

2 Transparencia y reproducibilidad en Economía

3 Manipular para confundir

- Manipular las gráficas
- Manipular las muestras
- Manipular cifras

4 Mensaje Final y Próxima Clase

# Mensaje Final y Próxima Clase

- ▶ Ser capaz de leer críticamente es una habilidad clave para los economistas
  - ▶ Importante conocer las manipulaciones/sesgos (importante para no caer en ellas, **NO** para que ustedes las usen!!)
  - ▶ Pensar en la importancia de transparencia en la investigación,
    - ▶ Poner a disposición códigos y datos (si es posible), GitLab, Github, SVN, Docker ... son buenas herramientas
- ▶ Semana que viene:
  - ▶ Trabajo en clase tarea de Literatura