

Combining Randomized Control Trials and Observational Data in the Age of Big Data

Jasjeet S. Sekhon
UC Berkeley

The Opportunity

- Explosion of data sources: administrative, electronic medical records, online behavior
- Population data is becoming more common and precise
- How can it be used?
- Policy makers/firms:
“let’s just use the big data to evaluate programs”
- Tension between identifying causal effects vs. prediction

The Problem

- Most Randomized Controlled Trials (RCTs) are small, especially a problem with medical experiments
- RCTs are usually not conducted on the population of interest
- We can now combine information from both RCTs and population data to estimate treatment effects in the population
- Goal: precision medicine, precise targeting of treatments

The Problem with Observational Data

- Biases can be large: a negative correlation may hide a large positive effect—e.g., campaign expenditures in the U.S.
- Causality is not prediction. A very good predictive model may tell one nothing about causality

ford

Web

Images

Maps

Shopping

News

More ▾

Search tools

About 868,000,000 results (0.30 seconds)

Ads related to **ford**

[Ford.com - Ford Focus Official Site](#)

[www.ford.com/Focus](#) ▾

Be Everywhere @ Once w/ Up to 40MPG **Ford** Focus. Learn More @ **Ford.com**.

[Build and Price](#)

Build & Price the 2014 Ford Focus.
A Small Car That's Big On Features!

[Photo Gallery](#)

See Interior & Exterior Photos
of the 2014 Focus @ Ford.com.

[2013 Toyota Camry - BuyAToyota.com](#)

[www.buyatoyota.com/Camry](#) ▾

0% for 60 mos PLUS \$1,000 Trade-in Cash on a new Camry. Learn more.

[Ford – New Cars, Trucks, SUVs, Hybrids & Crossovers | Ford Vehicles](#)

[www.ford.com](#) ▾

The Official **Ford** Site to research, learn and shop for all new **Ford** Vehicles. View photos, videos, specs, compare competitors, build and price, search inventory ...

[All Vehicles - 2013 F-150 - Mustang - 2014 Ford Focus](#)

Ford Motor Company

2,620,994 followers

Ford Motor Company is a multinational automobile manufacturer headquartered in Dearborn, Michigan, founded by Henry Ford on June 16, 1903. Wik

CEO: [Alan Mulally](#)

Headquarters: Dearborn, Michigan

Founder: [Henry Ford](#)

Founded: June 16, 1903

Awards: Car and Driver's "Most Reliable" (2011-2012)

ebay

Web

Images

Maps

Shopping

News

More ▼

Search tools

About 1,730,000,000 results (0.19 seconds)

[eBay: Electronics, Cars, Fashion, Collectibles, Coupons and More ...](#)

www.ebay.com/ 

Buy and sell electronics, cars, fashion apparel, collectibles, sporting goods, digital cameras, baby items, coupons, and everything else on **eBay**, the world's ...

[Motors](#)

Cars Trucks - Parts & Accessories -
Collector Cars - Motorcycles

[Women](#)

Tops & Blouses - Coats & Jackets -
Swimwear - Shoes - Sweaters

[Daily Deals](#)

Deals are updated daily, so check
back for the deepest discounts ...

[More results from ebay.com »](#)

[Cars Trucks](#)

Salvage - Truck - Rat rod - eBay
Motors - Project - No Reserve

[Men](#)

Casual Shirts - Jeans - Coats &
Jackets - Shorts - Dress Shirts

[Cell Phone, Accessories](#)

Cell Phones & Accessories. Cell
Phones & Smartphones · Cell ...

[News for ebay](#)

eBay

Corporation

eBay Inc. is an American
consumer-to-consumer
headquartered in

Customer service
([Consumer](#))

Stock price: [EBAY](#)
Oct 4, 4:00 PM EDT

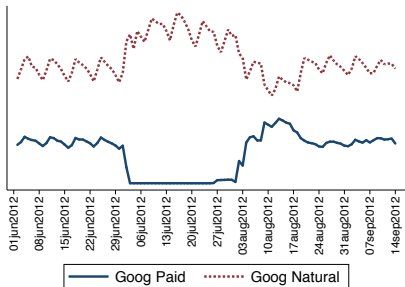
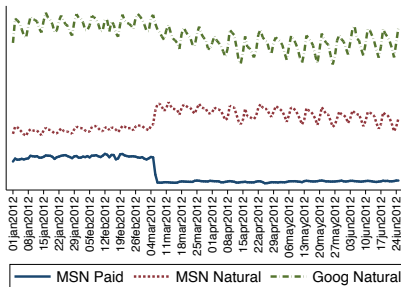
Founder: [Pierre Omidyar](#)

Founded: September 1995

Headquarters: San Jose, California

CEO: [John Donahoe](#)

An Experiment



From: Blake, Nosko, and Tadelis (2013)

Pulmonary Artery Catheterization (PAC)

- PAC is an invasive cardiac monitoring device for critical ill patients (ICU)—e.g., myocardial infarction (ischaemic heart disease)
- PAC-man trial $n=1,013$
- Registry data: 1.5million ICU admissions. 1,052 PAC cases and 31,447 potential controls

Joint Estimation Method

- If the observational data is not useful, one will not bias the estimates. They will be based on the RCT
- If the observational data contains useful information, it will be positively weighted
- One can estimate either sample or population parameters

Combined Estimates

