

R AND RCPP

David Zarruk Valencia

October 11, 2017

WHY USE R?

1. Open Source:

- ▶ No license - free
- ▶ Packages for every single task
- ▶ Tons of documentation

"Am I a developer or just a good googler?"

The screenshot shows the Stack Overflow homepage with a search bar at the top containing the query "[r]". Below the search bar, there are four search results boxes, each with a red border around it. The first result is for the tag "[julia-lang]" with 3,064 questions tagged. The second result is for "[matlab]" with 72,613 questions tagged. The third result is for "[stata]" with 2,499 questions tagged. The fourth result is for "[r]" with 189,351 questions tagged.

Tag	Count
[julia-lang]	3,064
[matlab]	72,613
[stata]	2,499
[r]	189,351

WHY DOES OPEN SOURCE WORK?

► Signaling device:



231,521 REPUTATION

27 421 530

Dirk Eddelbuettel top 0.03% overall

<http://dirk.eddelbuettel.com/code/>

- See my [blog](#) for some updates on what I've been up to.
- Sometimes I tweet using the [@eddelbuettel](#) tag.
- And I also update my [Google+ page](#).

3,137 answers 5 questions ~18.0m people reached

📍 Chicago, IL, United States

👤 eddelbuettel

✉️ eddelbuettel

🌐 dirk.eddelbuettel.com

⌚ Member for 7 years, 11 months

👁️ 31,116 profile views

🕒 Last seen 30 mins ago

Communities (18)

Stack Overflow	231.5k
Cross Validated	6.8k
Quantitative Finance	4.3k
Super User	1.2k
TeX - LaTeX	588

Top Tags (1,177)

r	●	SCORE 17,952 POSTS 2,534 POSTS % 81	
rcpp	●	SCORE 2,000 POSTS 621	c++ ● SCORE 1,700 POSTS 499
plot	●	SCORE 938 POSTS 74	c ● SCORE 694 POSTS 152 linux ● SCORE 651 POSTS 179

[View network profile →](#)

[View all tags →](#)

WHY USE R?

2. Leading tool for statistics and data processing

- ▶ Coordination game
- ▶ Estimation: regression, RD, non-parametric, semi-parametric, etc.
- ▶ Time series: forecasting, filters, ARIMA, etc.
- ▶ Optimization: global, local, etc.
- ▶ Graphing packages: ggplot2, dplyr
- ▶ Ongoing research:



[Explore this journal >](#)

Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs

Sebastian Calonico, Matias D. Cattaneo, Rocio Titiunik

First published: November 2014 [Full publication history](#)



[Explore this journal >](#)

Inference on Counterfactual Distributions

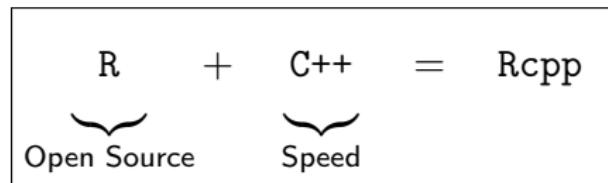
Victor Chernozhukov, Iván Fernández-Val, Blaise Melly

First published: 13 November 2013 [Full publication history](#)

3. Additional tools: Machine learning, Data scraping , GIS tools

R AND RCPP

- ▶ Only problem of R...
- ▶ Very slow at some tasks
- ▶ Eg: nested loops
- ▶ Solution:



LIFE-CYCLE MODEL

- ▶ Households solve:

$$V(t, e, x) = \max_{\{c, x'\}} \frac{c^{1-\sigma}}{1-\sigma} + \beta \mathbb{E} V(t+1, e', x') \quad s.t.$$

$$c + x' \leq (1+r)x + ew$$

$$\mathbb{P}(e'|e) = \Gamma(e)$$

$$x' \geq 0$$

$$t \in \{1, \dots, T\}$$

COMPUTING THE MODEL

1. Choose grids for assets $X = \{x_1, \dots, x_{n_x}\}$ and shocks $E = \{e_1, \dots, e_{n_e}\}$.
2. Backwards induction:

2.1 For $t = T$ and every $x_i \in X$ and $e_j \in E$, solve the static problem:

$$V(t, e_j, x_i) = \max_{\{c\}} u(c) \quad s.t. \quad c \leq (1+r)x_i + e_j w$$

2.2 For $t = T-1, \dots, 1$, use $V(t+1, e_j, x_i)$ to solve:

$$\begin{aligned} V(t, e_j, x_i) &= \max_{\{c, x' \in X\}} u(c) + \beta \mathbb{E} V(t+1, e', x') \quad s.t. \\ &c + x' \leq (1+r)x_i + e_j w \\ &\mathbb{P}(e' \in E | e_j) = \Gamma(e_j) \end{aligned}$$

CODE STRUCTURE

```
for(age = T:-1:1)
    for(ix = 1:nx)
        for(ie = 1:ne)

            VV = -10^3;
            for(ixp = 1:nx)

                expected = 0.0;
                if(age < T)
                    for(iep = 1:ne)
                        expected = expected + P[ie, iep]*V[age+1, ixp, iep];
                    end
                end

                cons = (1 + r)*xgrid[ix] + egrid[ie]*w - xgrid[ixp];
                utility = (cons^(1-ssigma))/(1-ssigma) + bbeta*expected;

                if(cons <= 0)
                    utility = -10^5;
                end

                if(utility >= VV)
                    VV = utility;
                end

            end

            V[age, ix, ie] = VV;

        end
    end
end
```