

Adding Sparkle to Social Coding: An Empirical Study of Repository Badges in the npm Ecosystem

Asher Trockman, Shurui Zhou, Christian Kästner, Bogdan Vasilescu

ICSE '18, May 27-June 3, 2018, Gothenburg, Sweden

GitHub Repository Badges

The screenshot displays the GitHub repository page for 'caolan / async'. At the top, the repository name is shown with navigation links for 'Code', 'Issues' (21), 'Pull requests' (6), 'Projects' (0), 'Wiki', and 'Insights'. On the right, there are buttons for 'Watch' (721), 'Star' (23,937), and 'Fork' (2,203). Below this, a description states 'Async utilities for node and the browser' with a link to 'http://caolan.github.io/async/'. There are also tags for 'javascript', 'async', and 'callbacks'. A horizontal bar shows repository statistics: '1,629 commits', '11 branches', '72 releases', '206 contributors', and 'MIT' license. The main content area shows the 'README.md' file with the 'async' logo and a row of badges: 'build passing', 'npm v2.6.0', 'coverage 99%', 'gitter join chat', 'examples 26348', and 'jsDelivr 407k hits/month'. The README text describes Async as a utility module for asynchronous JavaScript, installable via 'npm install --save async'.

caolan / async

Watch 721 Star 23,937 Fork 2,203

Code Issues 21 Pull requests 6 Projects 0 Wiki Insights

Async utilities for node and the browser <http://caolan.github.io/async/>

javascript async callbacks

1,629 commits 11 branches 72 releases 206 contributors MIT

README.md

async

build passing npm v2.6.0 coverage 99% gitter join chat examples 26348 jsDelivr 407k hits/month

Async is a utility module which provides straight-forward, powerful functions for working with [asynchronous JavaScript](#). Although originally designed for use with [Node.js](#) and installable via `npm install --save async`, it can also be used directly in the browser.

Key features: Transparency & signaling

request / request

Watch 417

Star 18,384

Fork 2,196

Code

Issues 578

Pull requests 52

Projects 0

Wiki

Insights

Simplified HTTP request client.

2,199 commits

17 branches

134 releases

270 contributors

Apache-2.0

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

mikeal committed on Sep 27, 2017 2.83.1

Latest commit 253c5e5 on Sep 27, 2017

.github	small change to template wording	a year ago
examples	Adds example for Tor proxy	2 years ago
lib	refactor(lint): replace eslint with standard (#2579)	11 months ago
tests	lint fix, PR from pre-standard was merged with passing tests	10 months ago
.gitignore	Updating deps.	5 months ago
.travis.yml	Add Node.js v8 to Travis CI	7 months ago
CHANGELOG.md	Update changelog	11 months ago
CONTRIBUTING.md	Update contributing guidelines	2 years ago
LICENSE	Adding license information.	7 years ago

Projects

776

Followers

38

Starred

15

Following

Organizations

Users

Contributions

Repositories

Public activity

Follow

Popular repositories

breakfast-repo

a collection of videos, recordings, and podcast...

208

x86-kernel

a simple x86 kernel, extended with Rust

48

ashleygwilliams.github.io

hi, i'm ashley. nice to meet you.

37

jsconf-2015-deck

deck for jsconf2015 talk, "if you wish to learn e...

32

ratpack

sinatra boilerplate using activerecord, sqlite, a...

32

Repositories contributed to

npm/docs

The place where all the npm docs live.

44

mozilla/publish.webmaker.org

The teach.org publishing service for goggles a...

2

npm/marky-markdown

npm's markdown parser

104

artisan-tattoo/assistant-frontend

ember client for assistant-API

5

npm/npm-camp

a community conference for all things npm

1

Public contributions

Feb

Mar

Apr

May

Jun

Jul

Aug

Sep

Oct

Nov

Dec

Jan

M

W

F

Summary of pull requests, issues opened, and commits. [Learn how we count contributions.](#)

Less More

Contributions in the last year

1,886 total

Jan 24, 2015 – Jan 24, 2016

Longest streak

37 days

October 7 – November 12

Current streak

7 days

January 18 – January 24

STRIDEL
Carnegie Mellon University

Key features: Transparency & signaling

request / request

Watch 417

Star 18,384

Fork 2,196

Code

Issues 578

Pull requests 52

Projects 0

Wiki

Insights

Simplified HTTP request client.

2,199 commits

17 branches

134 releases

270 contributors

Apache-2.0

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

mikeal committed on Sep 27, 2017 2.83.1

Latest commit 253c5e5 on Sep 27, 2017

.github	small change to template wording	a year ago
examples	Adds example for Tor proxy	2 years ago
lib	refactor(lint): replace eslint with standard (#2579)	11 months ago
tests	lint fix, PR from pre-standard was merged with passing tests	10 months ago
.gitignore	Updating deps.	5 months ago
.travis.yml	Add Node.js v8 to Travis CI	7 months ago
CHANGELOG.md	Update changelog	11 months ago
CONTRIBUTING.md	Update contributing guidelines	2 years ago
LICENSE	Adding license information.	7 years ago

Users

Contributions

Repositories

Public activity

Follow

Popular repositories

breakfast-repo

a collection of videos, recordings, and podcast...

208

x86-kernel

a simple x86 kernel, extended with Rust

48

ashleygwilliams.github.io

hi, i'm ashley. nice to meet you.

37

jsconf-2015-deck

deck for jsconf2015 talk, "if you wish to learn e...

32

ratpack

sinatra boilerplate using activerecord, sqlite, a...

32

Repositories contributed to

npm/docs

The place where all the npm docs live.

44

mozilla/publish.webmaker.org

The teach.org publishing service for goggles a...

2

npm/marky-markdown

npm's markdown parser

104

artisan-tattoo/assistant-frontend

ember client for assistant-API

5

npm/npm-camp

a community conference for all things npm

1

Public contributions

Summary of pull requests, issues opened, and commits. [Learn how we count contributions.](#)

Less More

Contributions in the last year

1,886 total

Jan 24, 2015 – Jan 24, 2016

Longest streak

37 days

October 7 – November 12

Current streak

7 days

January 18 – January 24

Projects

776

Followers

38

Starred

15

Following

Organizations

STRIDEL

Carnegie Mellon University

Key features: Transparency & signaling

request / request

Watch

417

Star

18,384

Fork

2,196

<> Code


Issues 578

Pull requests 52

Projects 0

Wiki

Insights

 Simplified HTTP request client.

2,199 commits

17 branches

134 releases

270 contributors

Apache-2.0

Branch: master


New pull request

Create new file










Upload files

Find file

Clone or download

 mikeal committed on Sep 27, 2017 2.83.1

Latest commit 253c5e5 on Sep 27, 2017

 .github	small change to template wording	a year ago
 examples	Adds example for Tor proxy	2 years ago
 lib	refactor(lint): replace eslint with standard (#2579)	11 months ago
 tests	lint fix, PR from pre-standard was merged with passing tests	10 months ago
 .gitignore	Updating deps.	5 months ago
 .travis.yml	Add Node.js v8 to Travis CI	7 months ago
 CHANGELOG.md	Update changelog	11 months ago
 CONTRIBUTING.md	Update contributing guidelines	2 years ago
 LICENSE	Adding license information.	7 years ago

Projects

776
Followers

38
Starred

15
Following

Organizations

Users

Contributions

Repositories

Public activity

Follow

Popular repositories

breakfast-repo

a collection of videos, recordings, and podcast...

208

★

x86-kernel

a simple x86 kernel, extended with Rust

48

★

ashleygwilliams.github.io

hi, i'm ashley. nice to meet you.

37

★

jsconf-2015-deck

deck for jsconf2015 talk, "if you wish to learn e...

32

★

ratpack

sinatra boilerplate using activerecord, sqlite, a...

32

★

Repositories contributed to

npm/docs

The place where all the npm docs live.

44

★

mozilla/publish.webmaker.org

The teach.org publishing service for goggles a...

2

★

npm/marky-markdown

npm's markdown parser

104

★

artisan-tattoo/assistant-frontend

ember client for assistant-API

5

★

npm/npm-camp

a community conference for all things npm

1

★

Public contributions

Feb

Mar

Apr

May

Jun

Jul

Aug

Sep

Oct

Nov

Dec

Jan

M

W

F

Summary of pull requests, issues opened, and commits. [Learn how we count contributions.](#)

Less

More

Contributions in the last year

1,886 total

Jan 24, 2015 – Jan 24, 2016

Longest streak

37 days





October 7 – November 12

Current streak

7 days

January 18 – January 24

Badges are ^{Mostly} Reliable Signals

-  of the presence of **tests**
-  of **up-to-date** and secure **dependencies**
-  of the presence of **tests in pull requests**
-  of **popularity**

Mixed methods study



Survey

- 32 maintainers, 57 contributors
- Maintainers:
 - What do you intend to signal?
 - What effects do you expect?
- Contributors:
 - What do badges tell you?

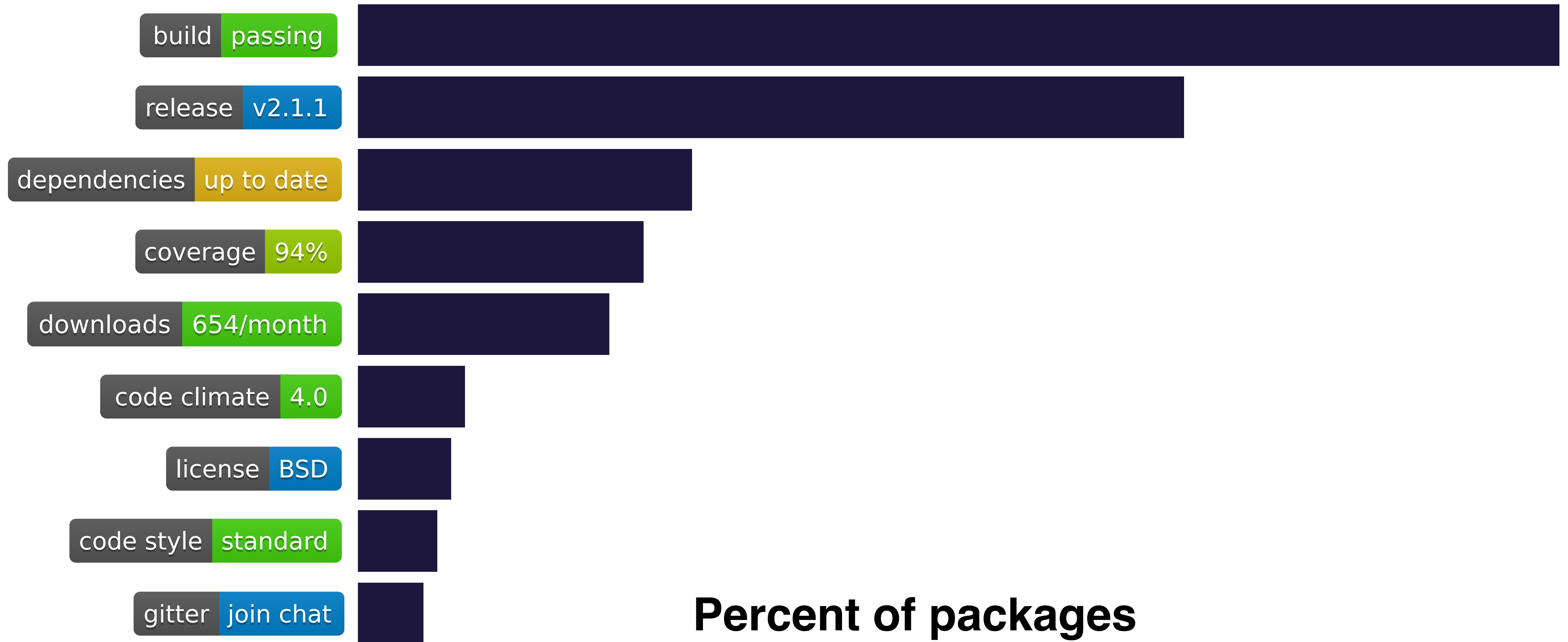
+



Repository Mining

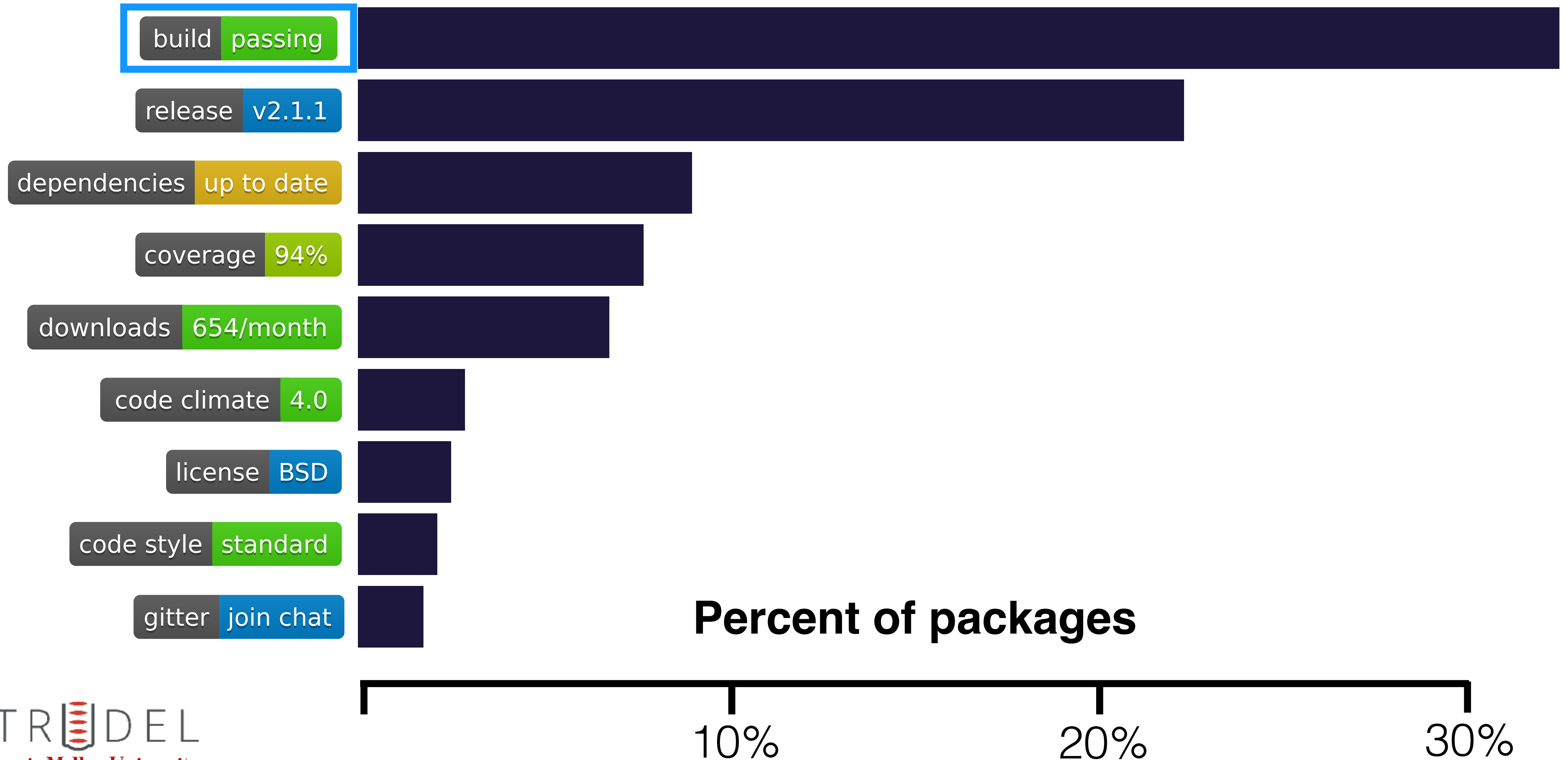
- 294,941 *npm* packages
- Mined badge adoptions/removals from README files
- Measured proxies for code quality, test suite quality, popularity, dependency freshness, ...

Popular Badges in

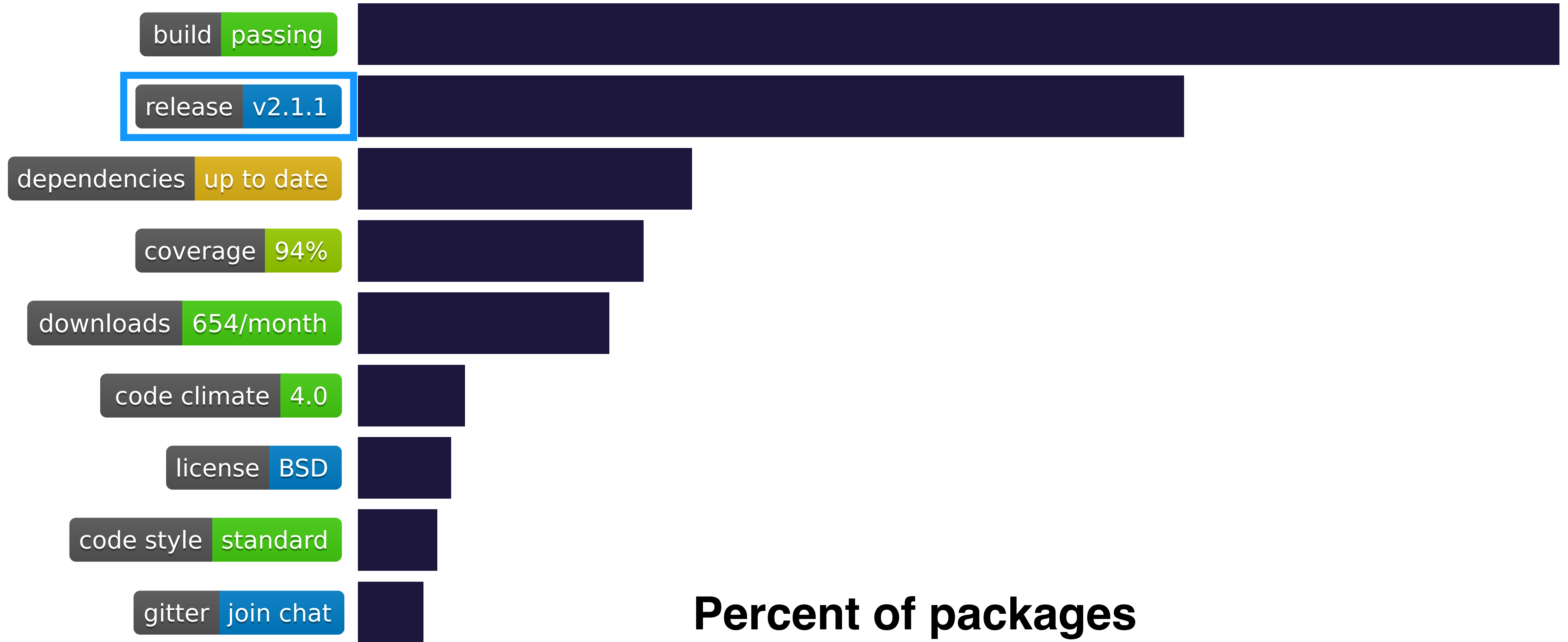


Percent of packages

Popular Badges in

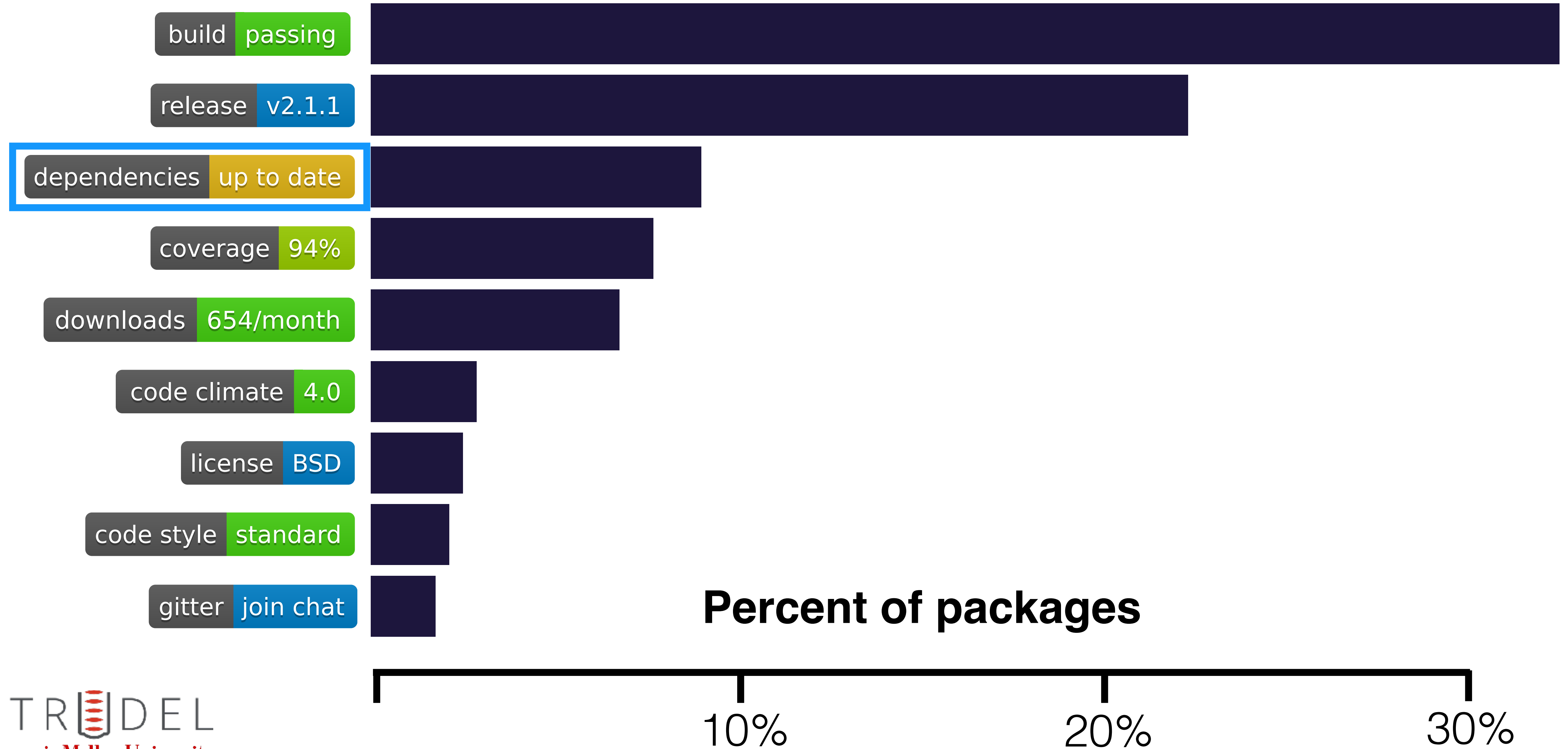


Popular Badges in

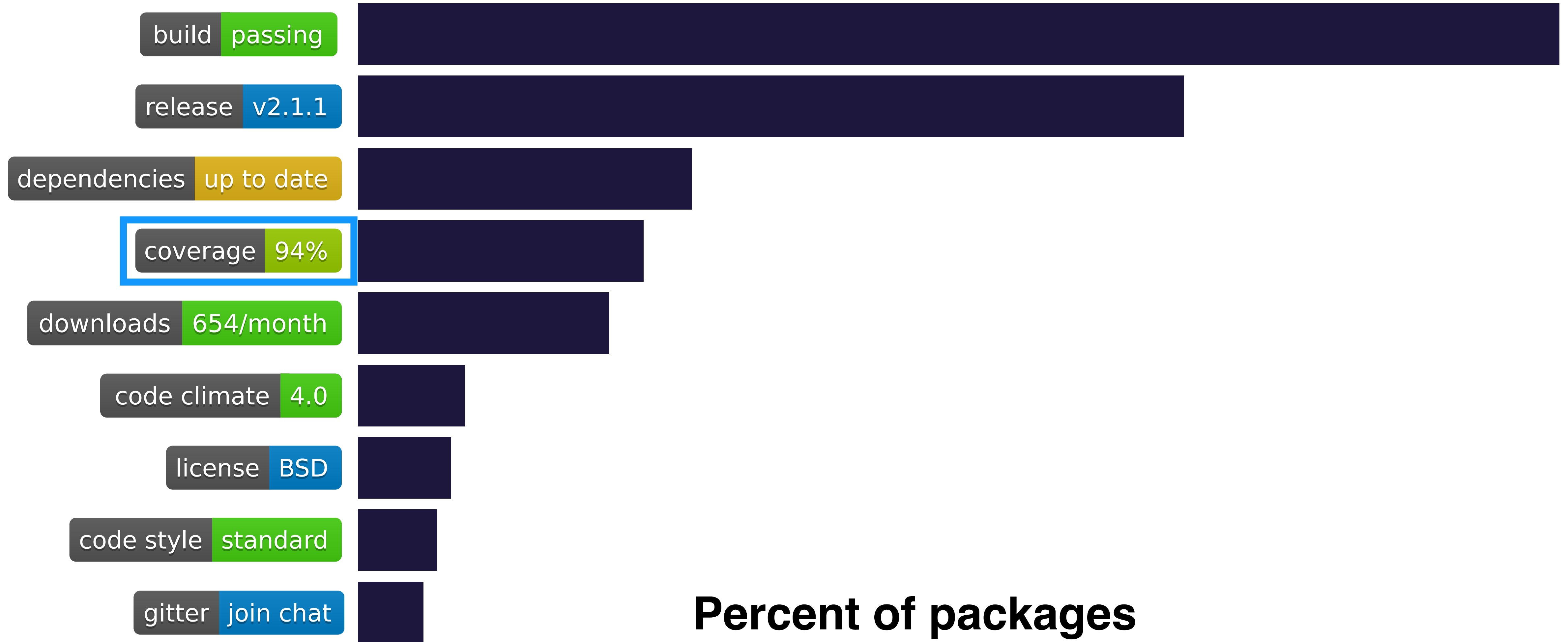


Percent of packages

Popular Badges in

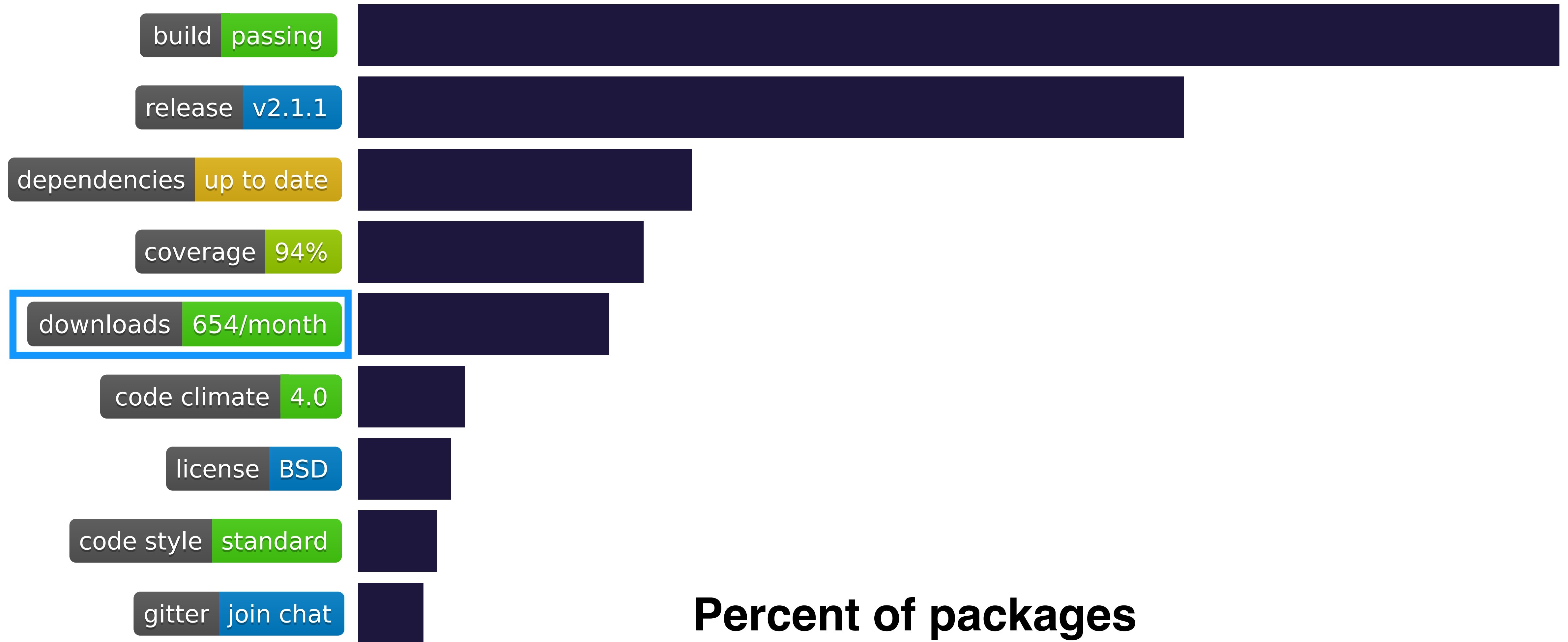


Popular Badges in



Percent of packages

Popular Badges in



What do developers expect from badges?

“welcoming contributions”

“expectations of contribution quality”

- 32 Maintainers
 - What do you intend to signal?
 - What effects do you expect?
- 57 Contributors
 - What do badges tell you?

“dedicated to offering support”

“reduced chances of conflicting versions of dependencies”

“indicator of product quality”

Analysis



Correlation

If all you saw was the badge, how much would that tell you?



Regression Analysis

How much more does the badge tell you, relative to existing signals?



Time Series Analysis

How do things change after adding the badge?

Analysis



Correlation

If all you saw was the badge, how much would that tell you?



Regression Analysis

How much more does the badge tell you, relative to existing signals?



Time Series Analysis

How do things change after adding the badge?

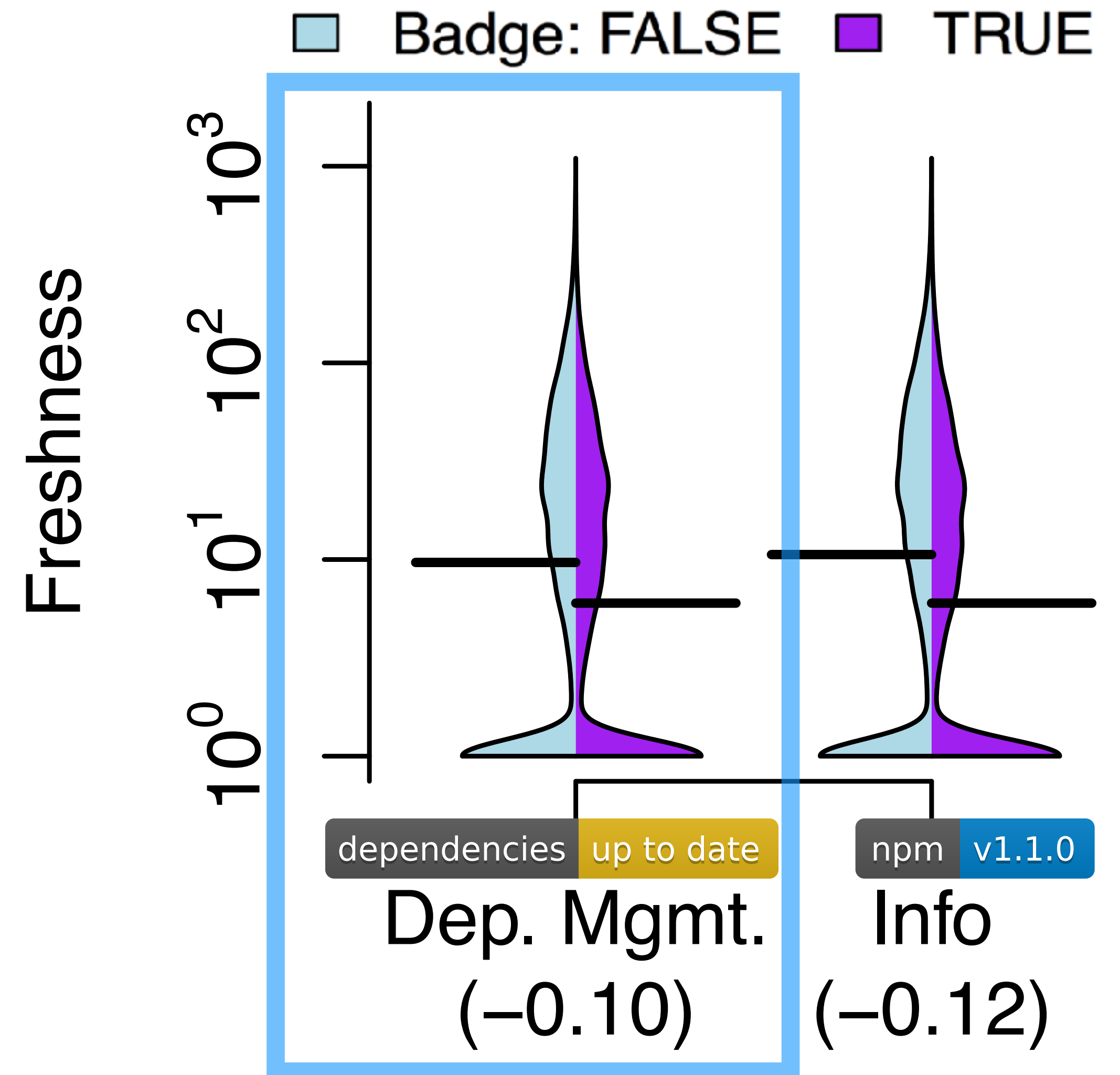
Step 1: Correlation

Signals of fresh dependencies



- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
 - (More up-to-date deps.)

Result: Dep. badges correlate with fresher dependencies



Analysis



Correlation

If all you saw was the badge, how much would that tell you?



Regression Analysis

How much more does the badge tell you, relative to existing signals?



Time Series Analysis

How do things change after adding the badge?

Step 2: Regression Analysis



Signals of fresh dependencies

- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
 - (More up-to-date deps.)

Result: Dep. badges are *the best* signals of fresh dependencies

Basic Model			Full Model		
response: <i>freshness</i> = 0			response: <i>freshness</i> = 0		
17.3% deviance explained			17.4% deviance explained		
	Coeffs (Err.)	LR Chisq		Coeffs (Err.)	LR Chisq
(Interc.)	3.54 (0.03)***			3.50 (0.03)***	
Dep.	−1.78 (0.01)***	32077.8***		−1.79 (0.01)***	32292.8***
RDep.	0.22 (0.01)***	610.3***		0.21 (0.01)***	560.6***
Stars	−0.08 (0.00)***	301.4***		−0.09 (0.00)***	311.2***
Contr.	−0.24 (0.01)***	500.5***		−0.25 (0.01)***	548.7***
lastU	−0.65 (0.01)***	12080.9***		−0.64 (0.01)***	11537.9***
dependencies	up to date			0.24 (0.03)***	116.1***
npm	v1.1.0			0.11 (0.02)***	48.3***
dependencies	up to date	: npm	v1.1.0	−0.05 (0.04)	1.9
hasOther				0.01 (0.01)	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Analysis



Correlation

If all you saw was the badge, how much would that tell you?



Regression Analysis

How much more does the badge tell you, relative to existing signals?

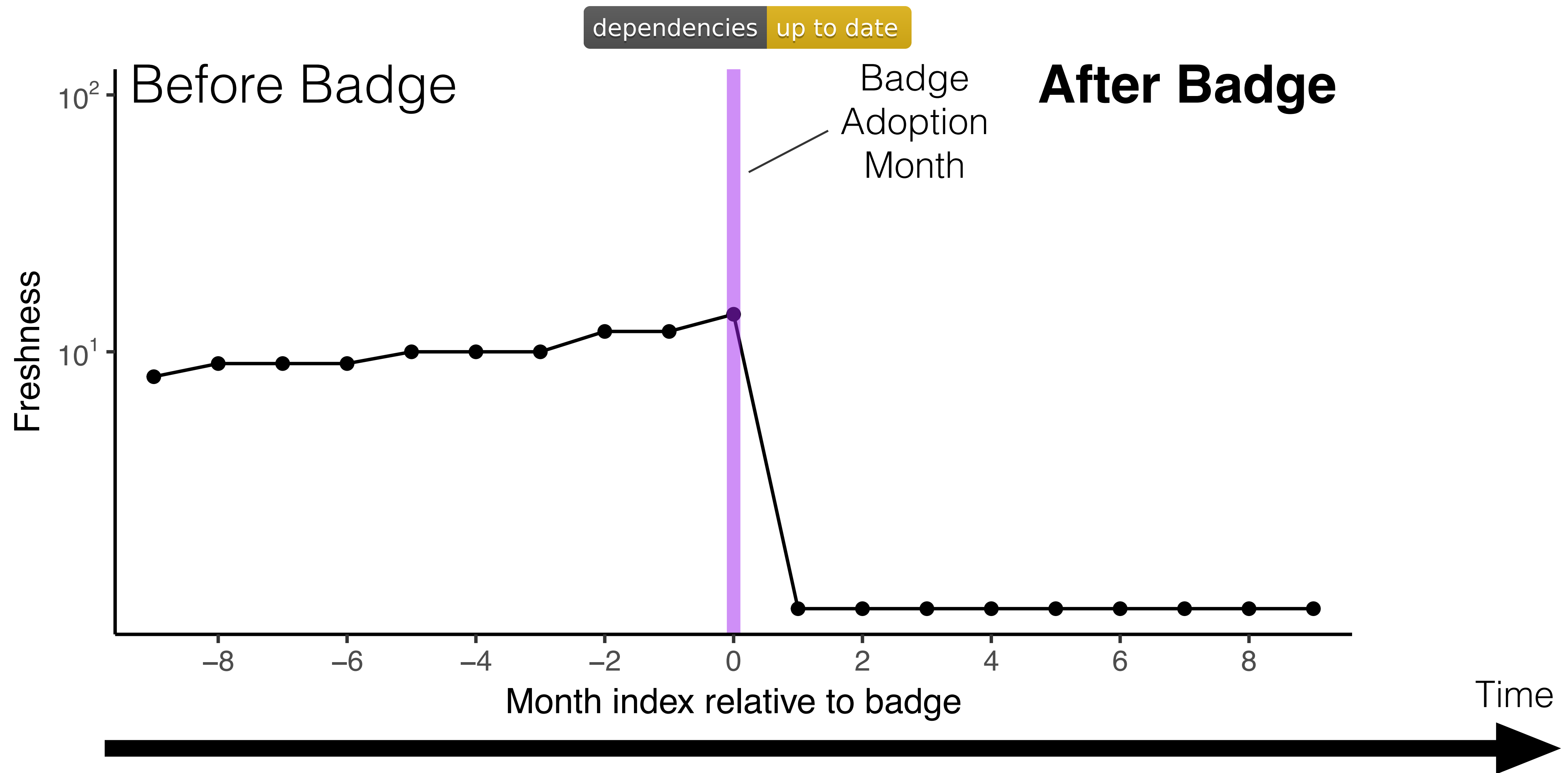


Time Series Analysis

How do things change after adding the badge?

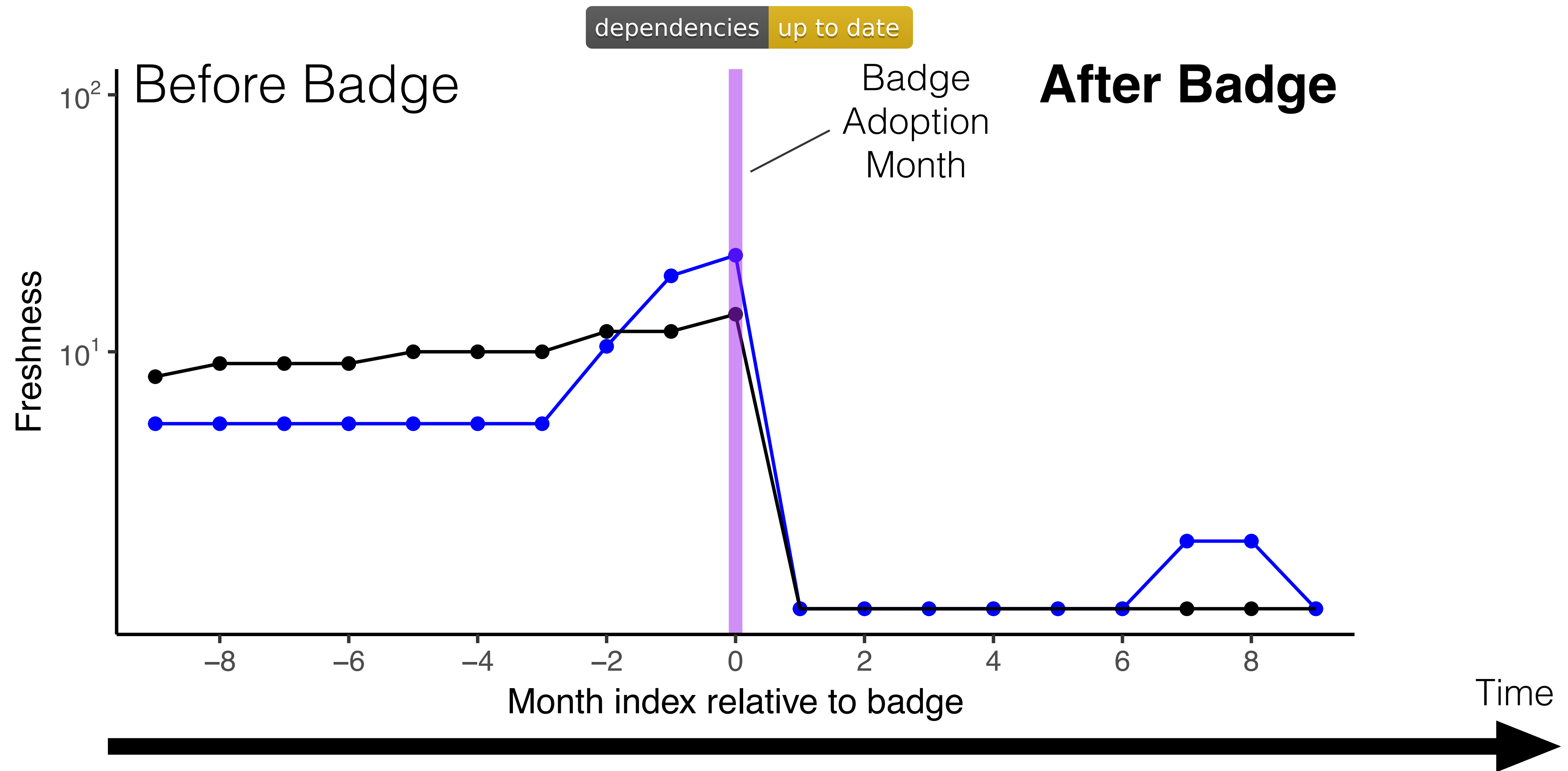
Step 2: Time Series Analysis

Signals of fresh dependencies



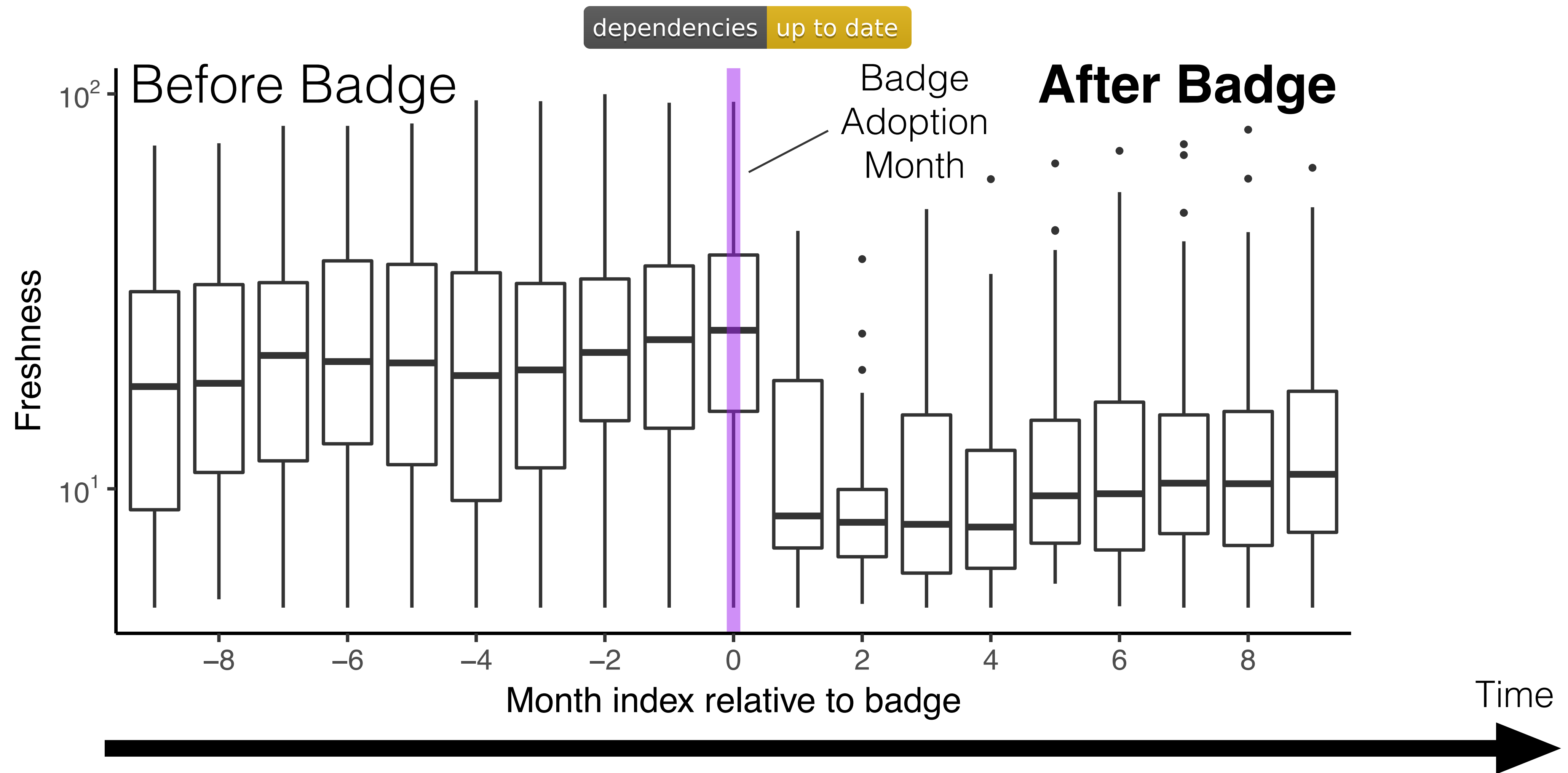
Step 2: Time Series Analysis

Signals of fresh dependencies



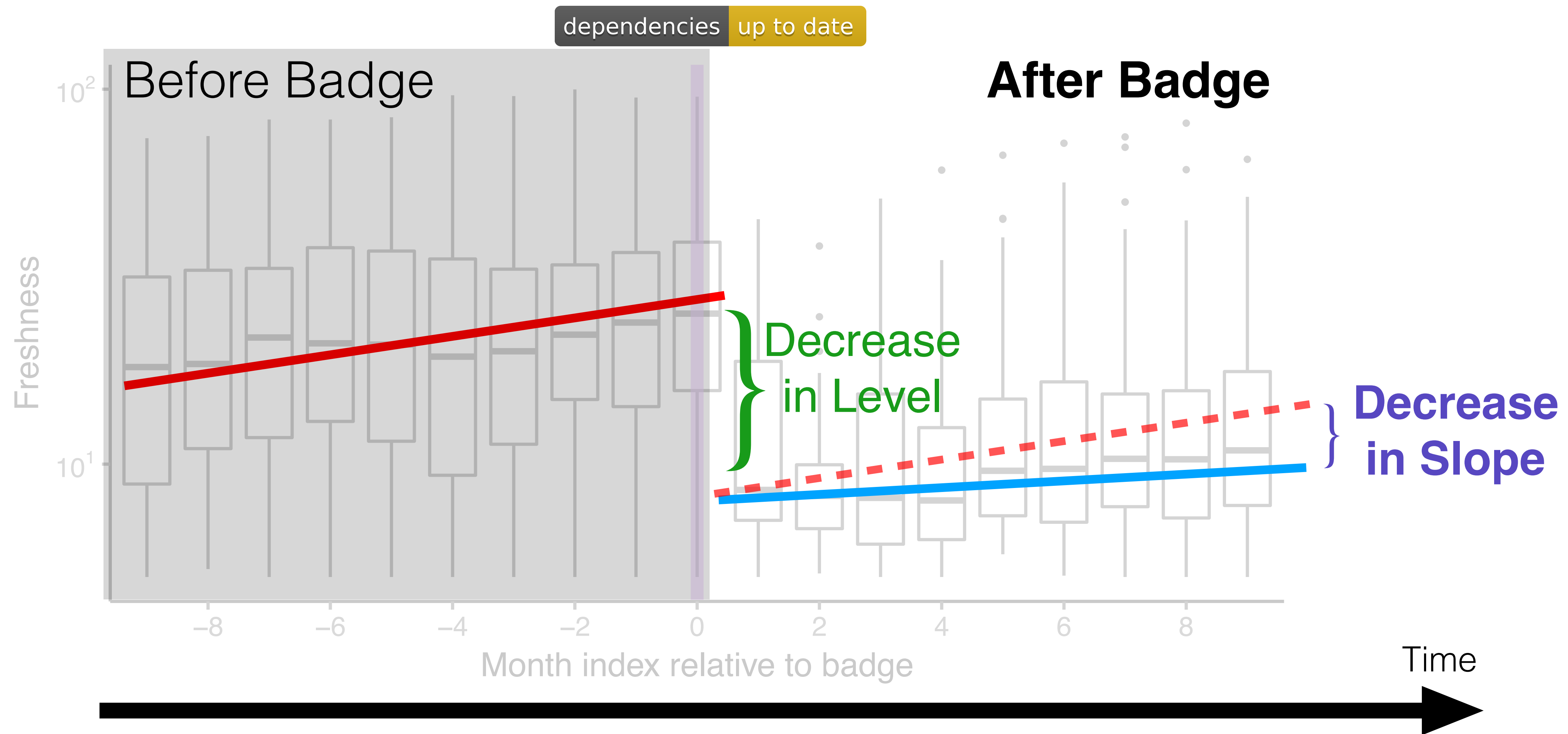
Step 2: Time Series Analysis

Signals of fresh dependencies



Step 2: Time Series Analysis

Signals of fresh dependencies



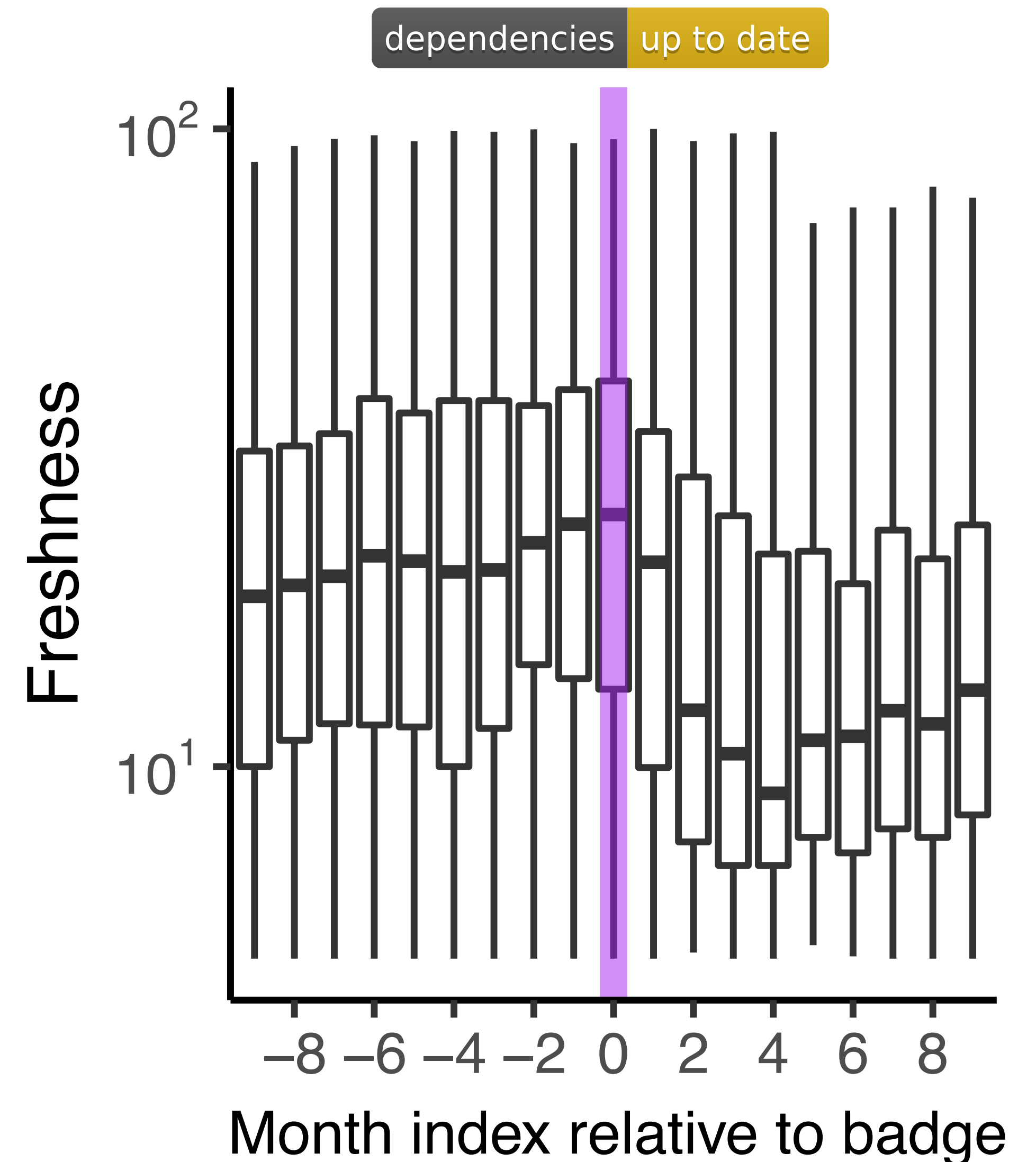
Step 2: Time Series Analysis

Signals of fresh dependencies

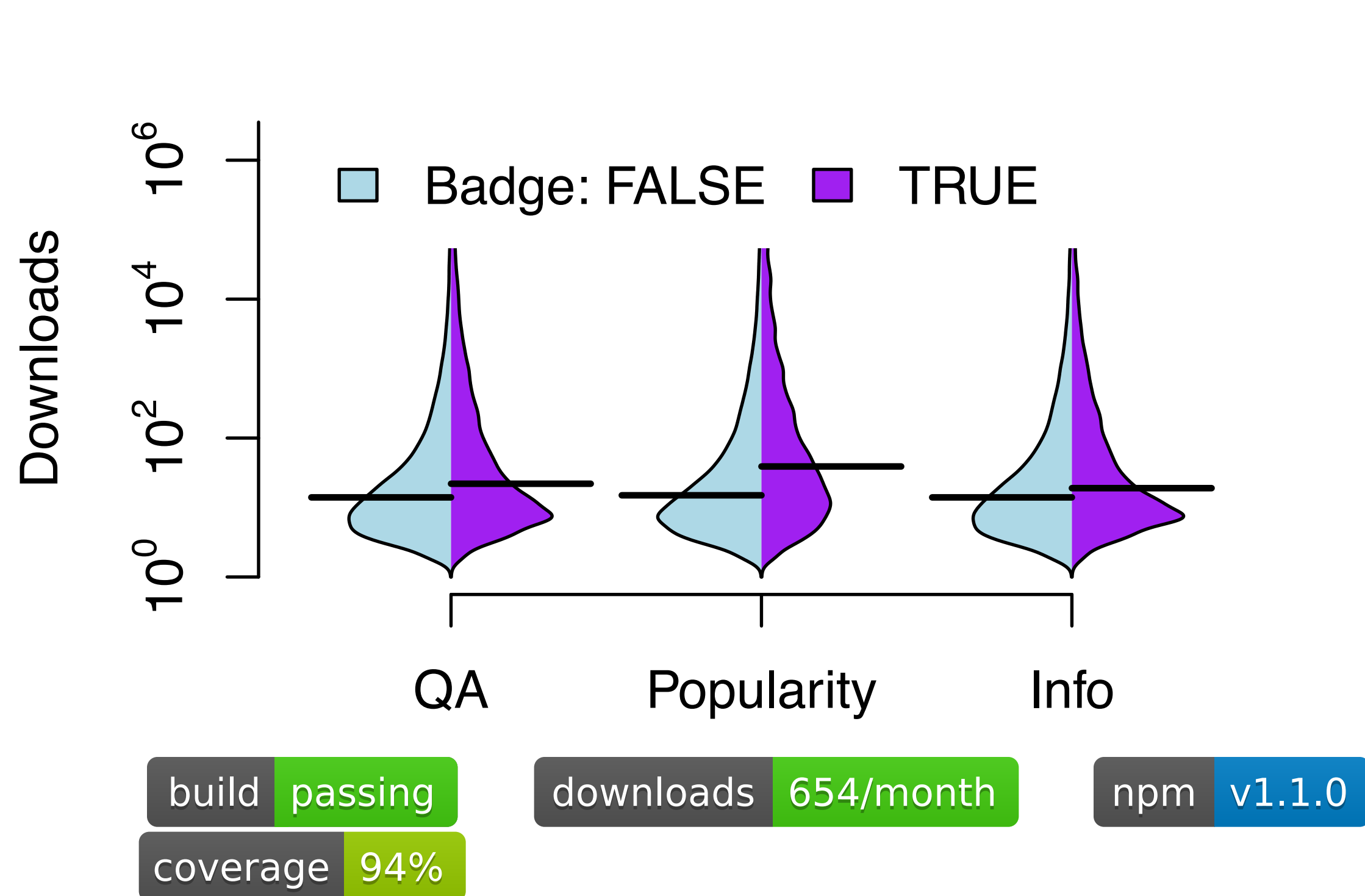


- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
 - (More up-to-date deps.)

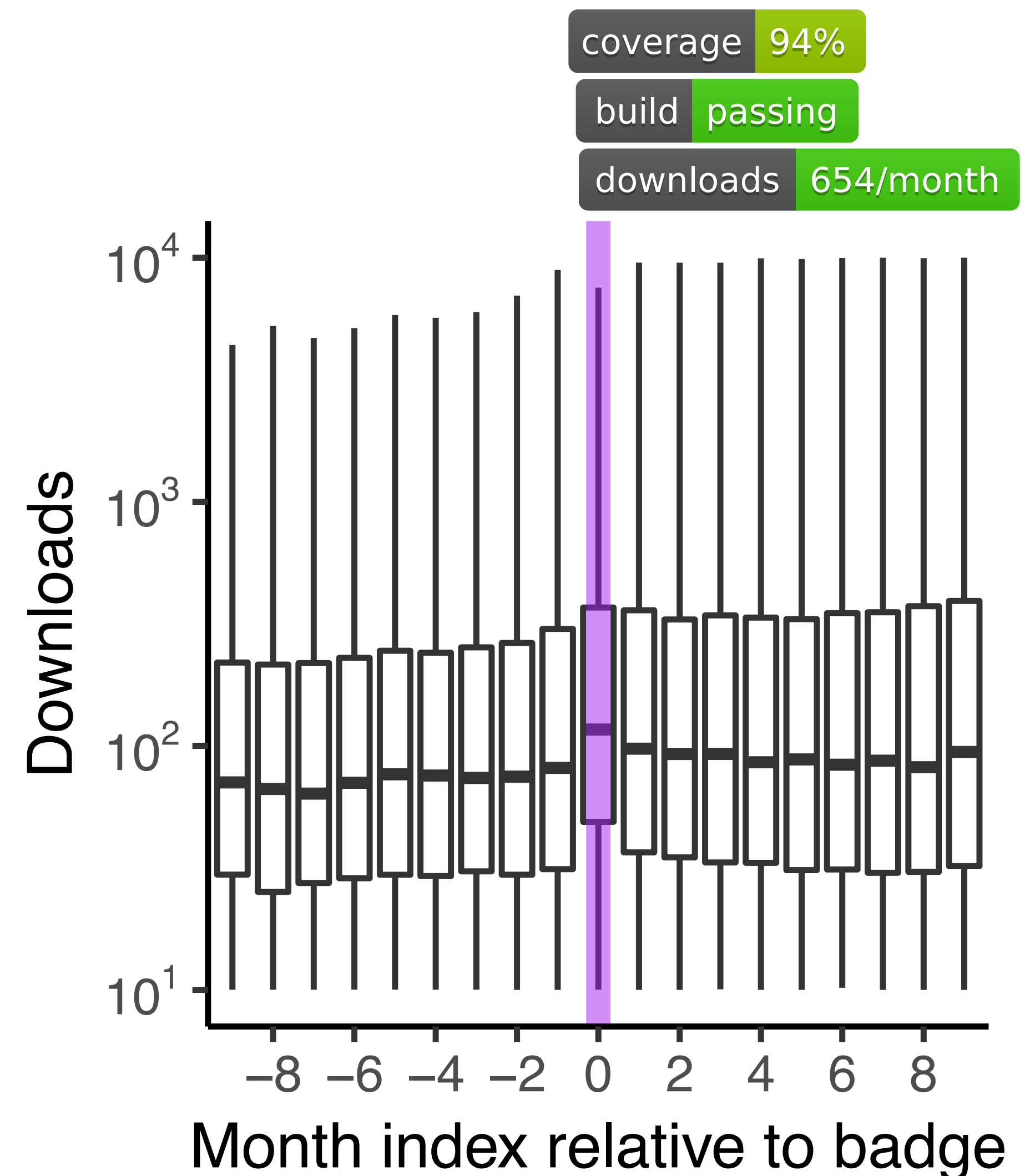
Result: Dep. badges indicate *improved dep. management practices*



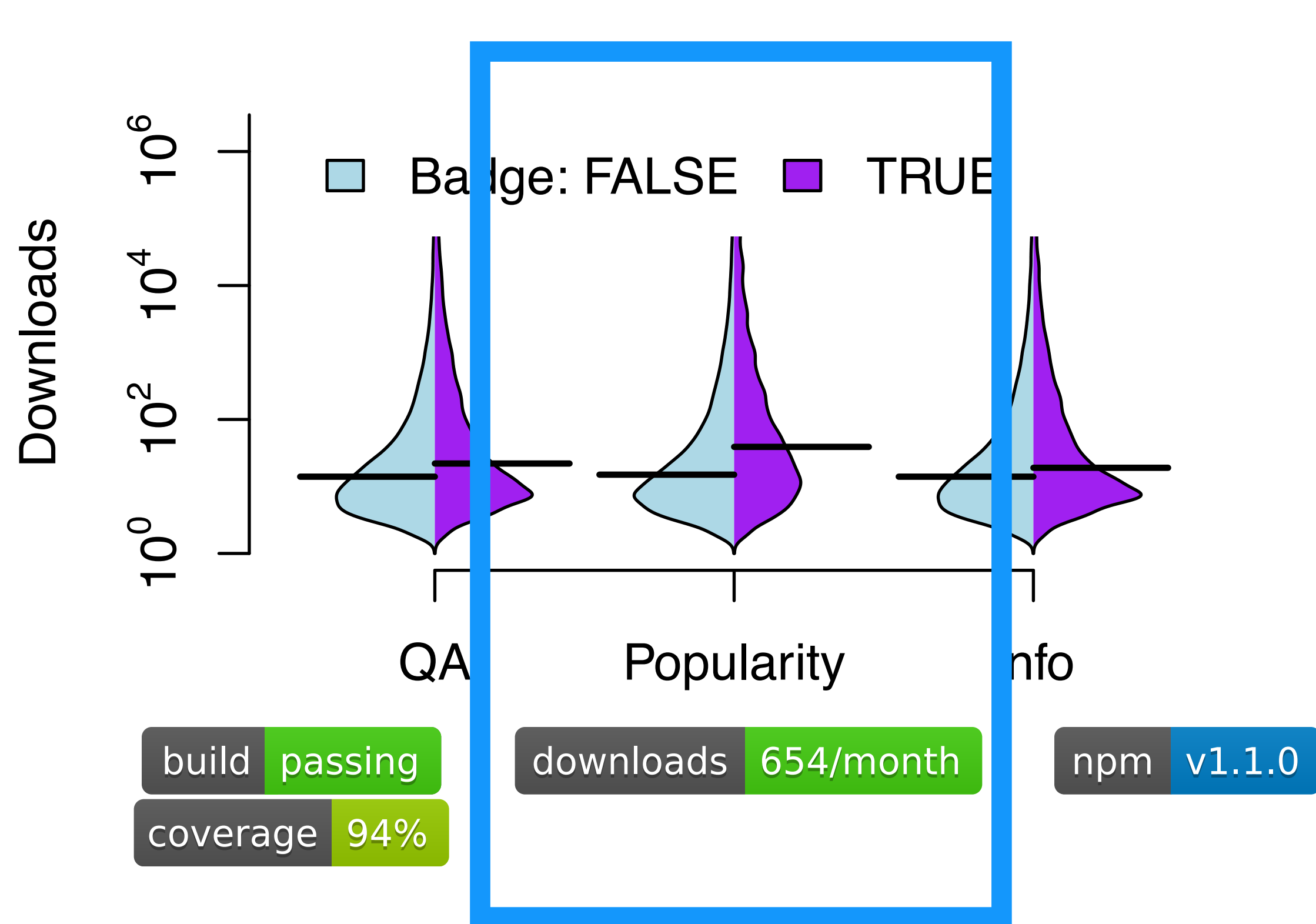
Signals of popularity



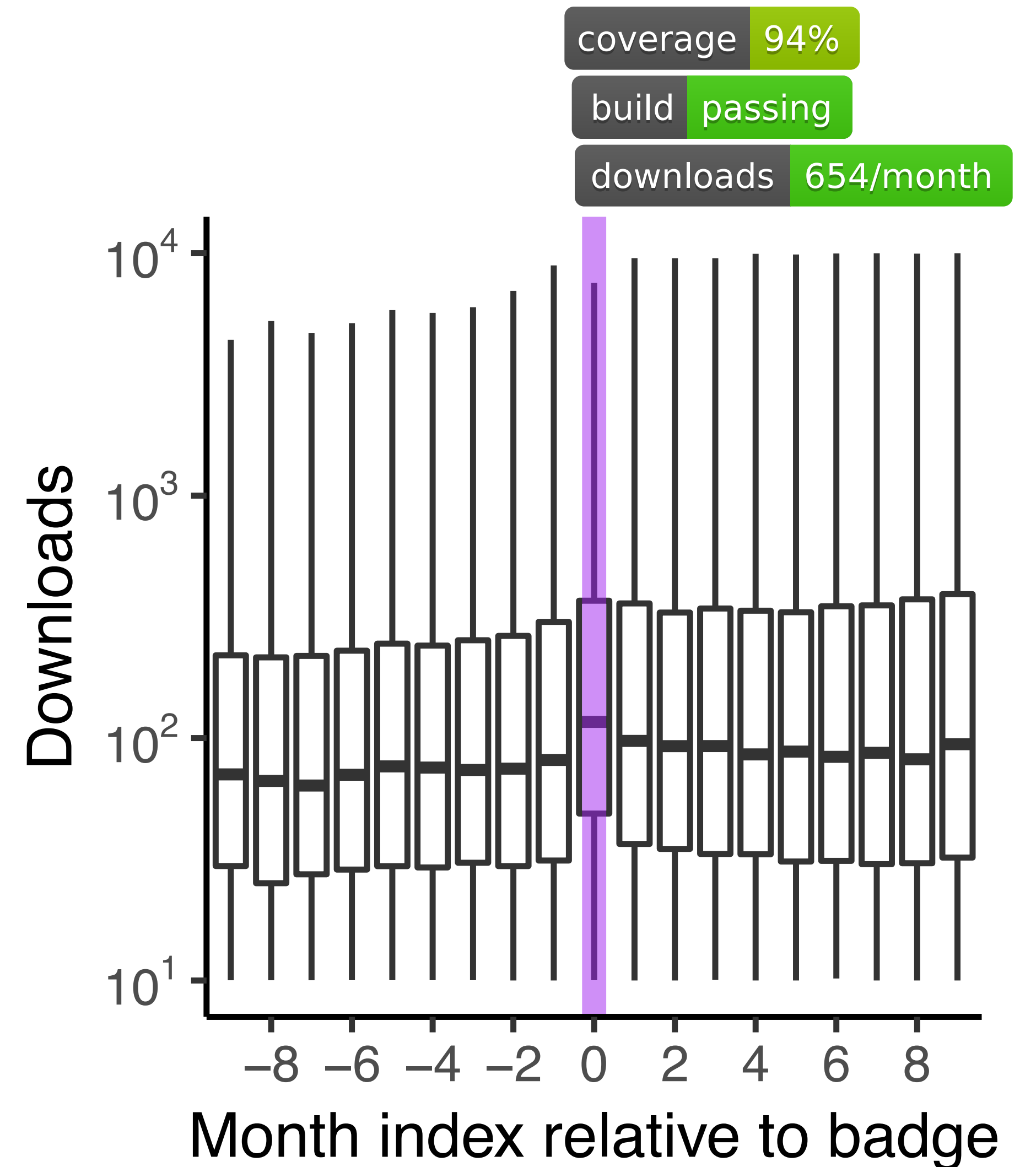
Result: Dep. badges are mostly reliable signals of popularity



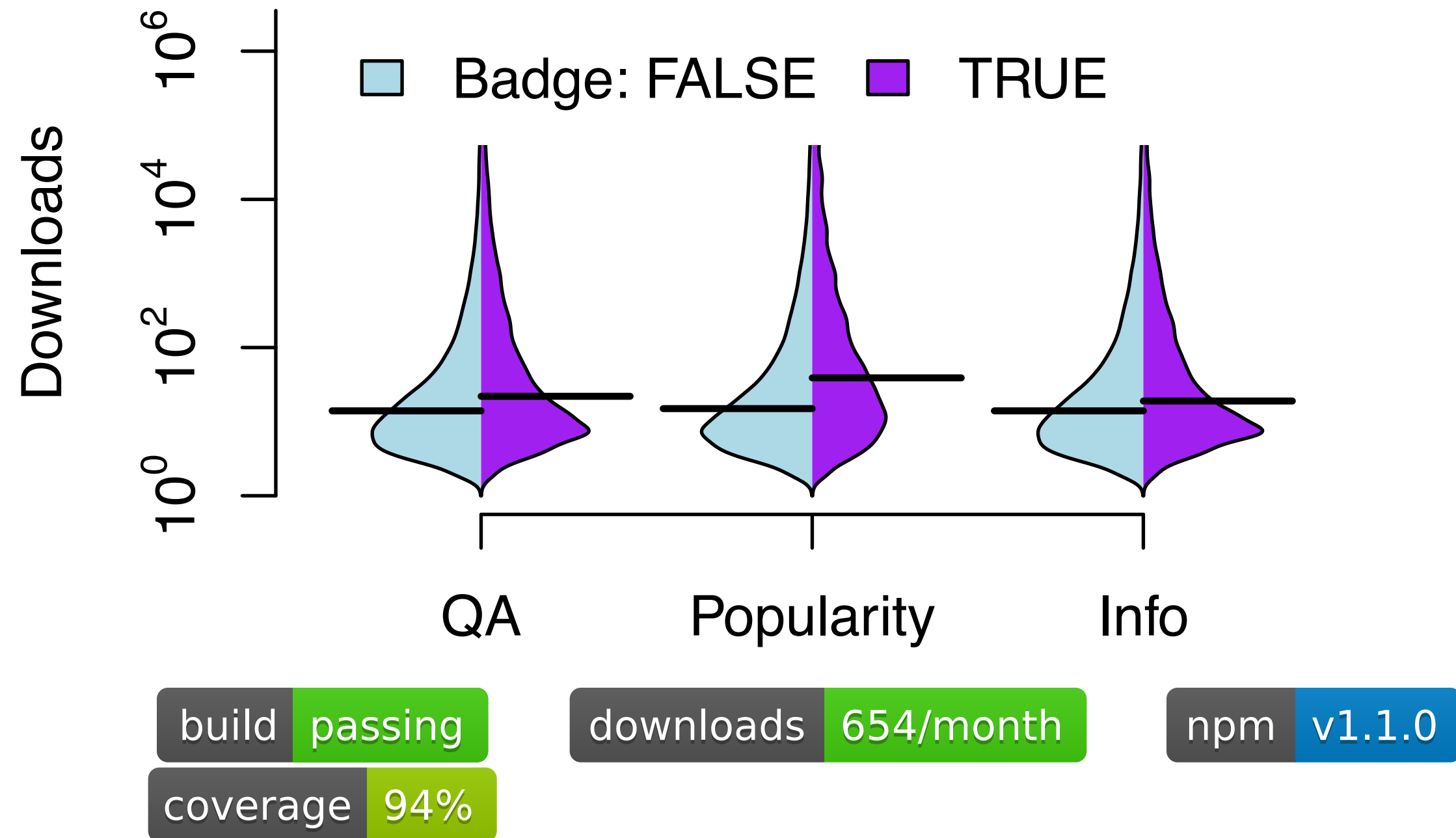
Signals of popularity



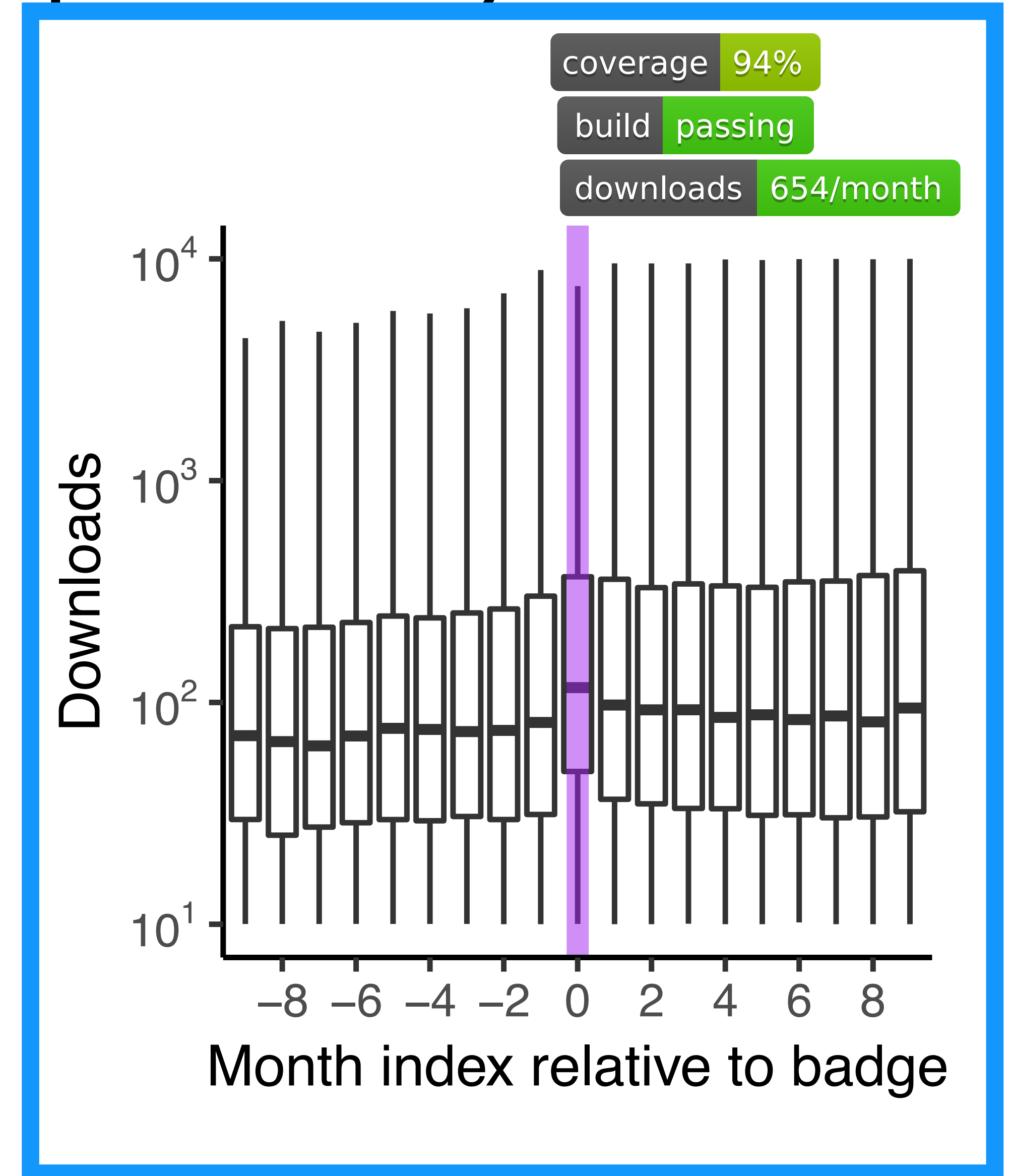
Result: Dep. badges are mostly reliable signals of popularity



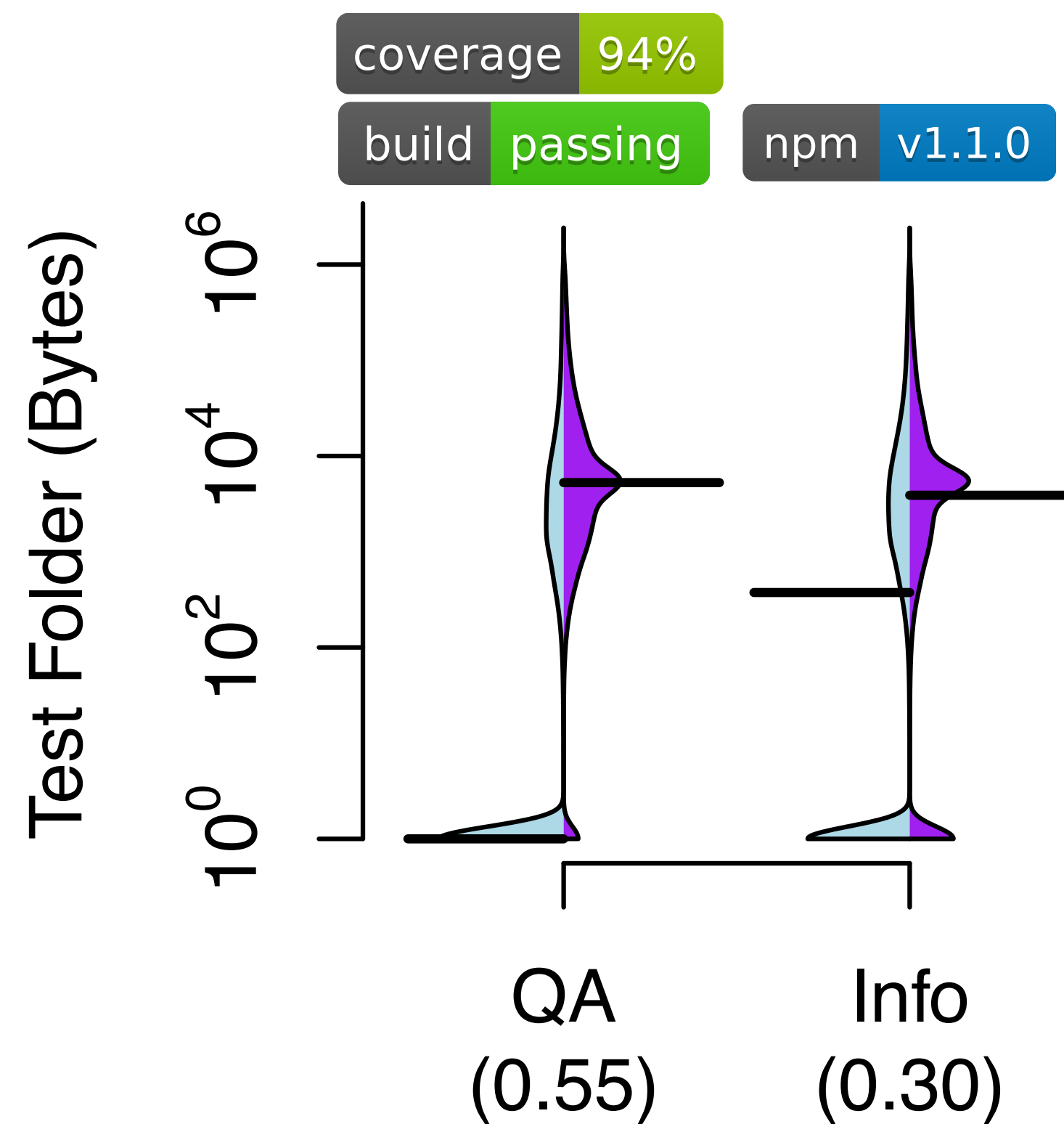
Signals of popularity



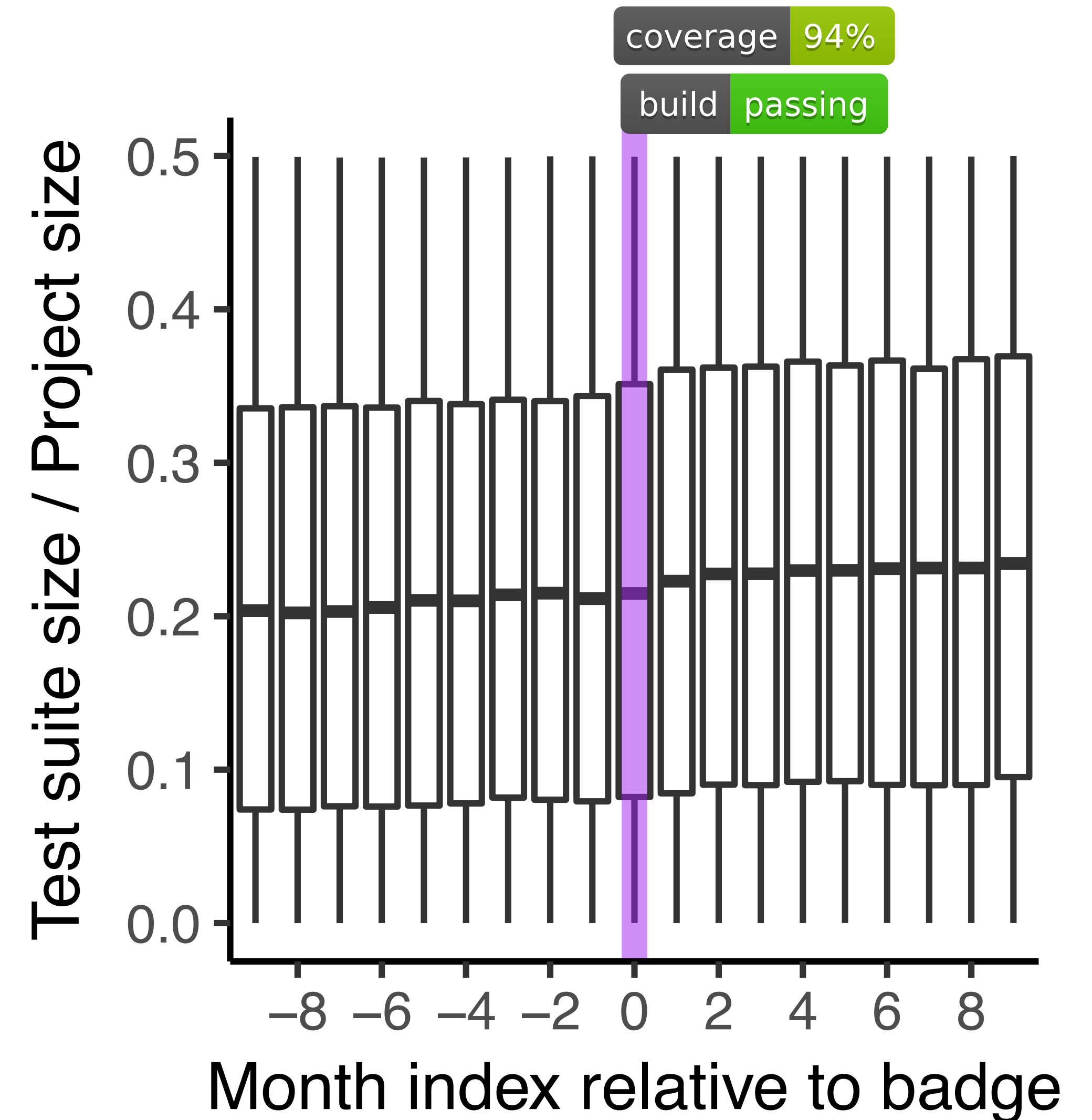
Result: Dep. badges are mostly reliable signals of popularity



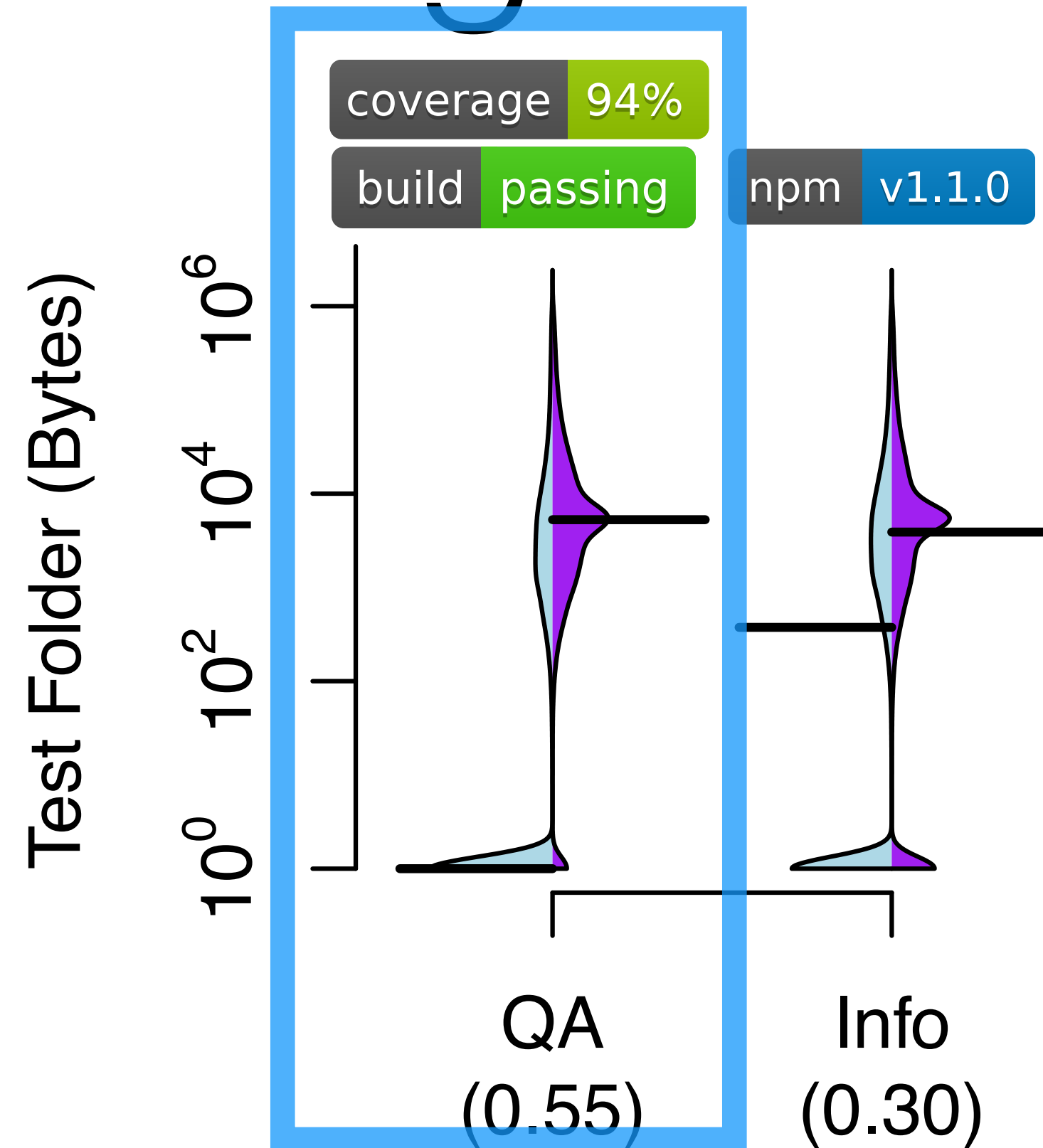
Signals of test suite quality



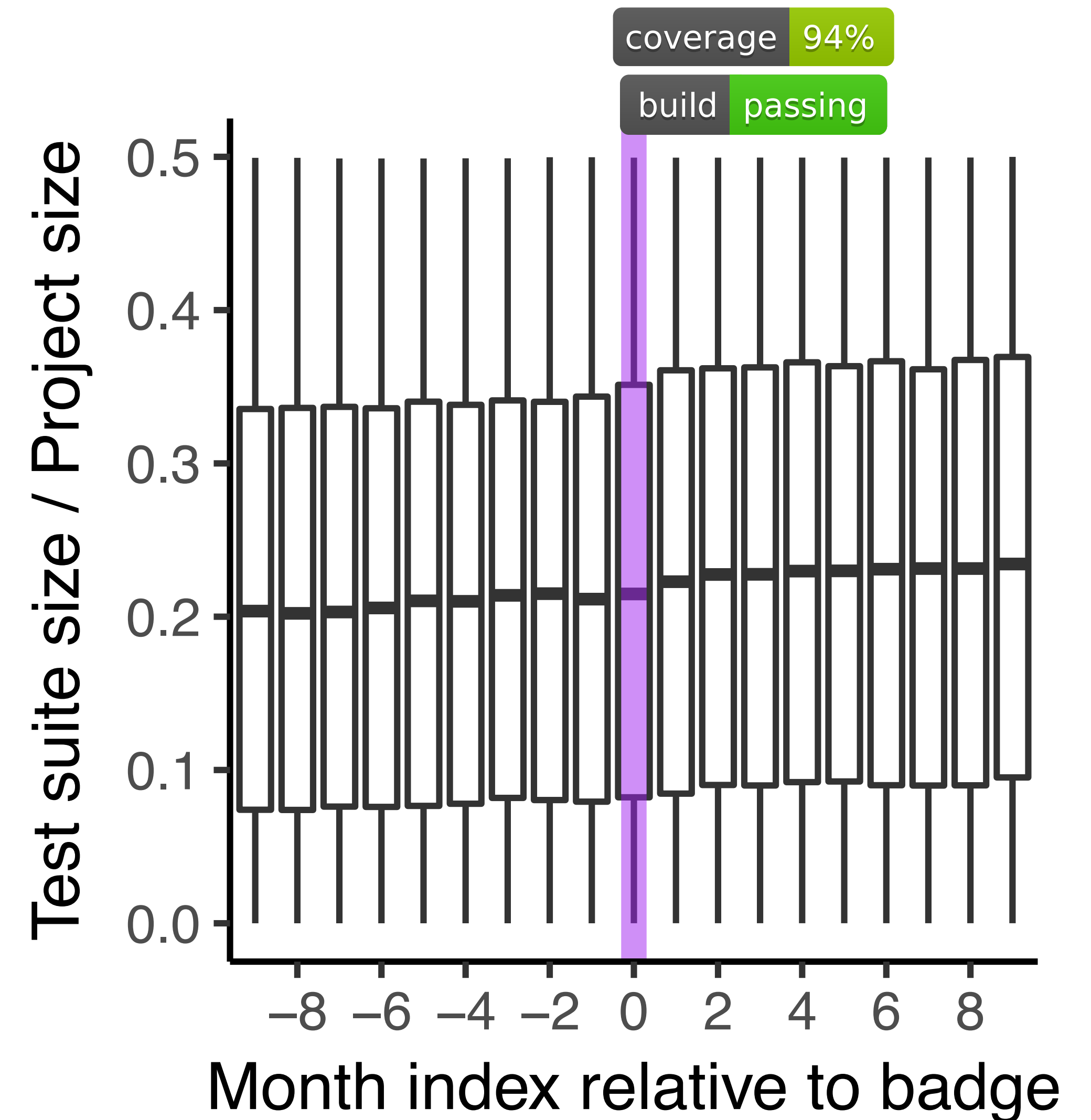
Result: Build status/code coverage badges indicate a test suite



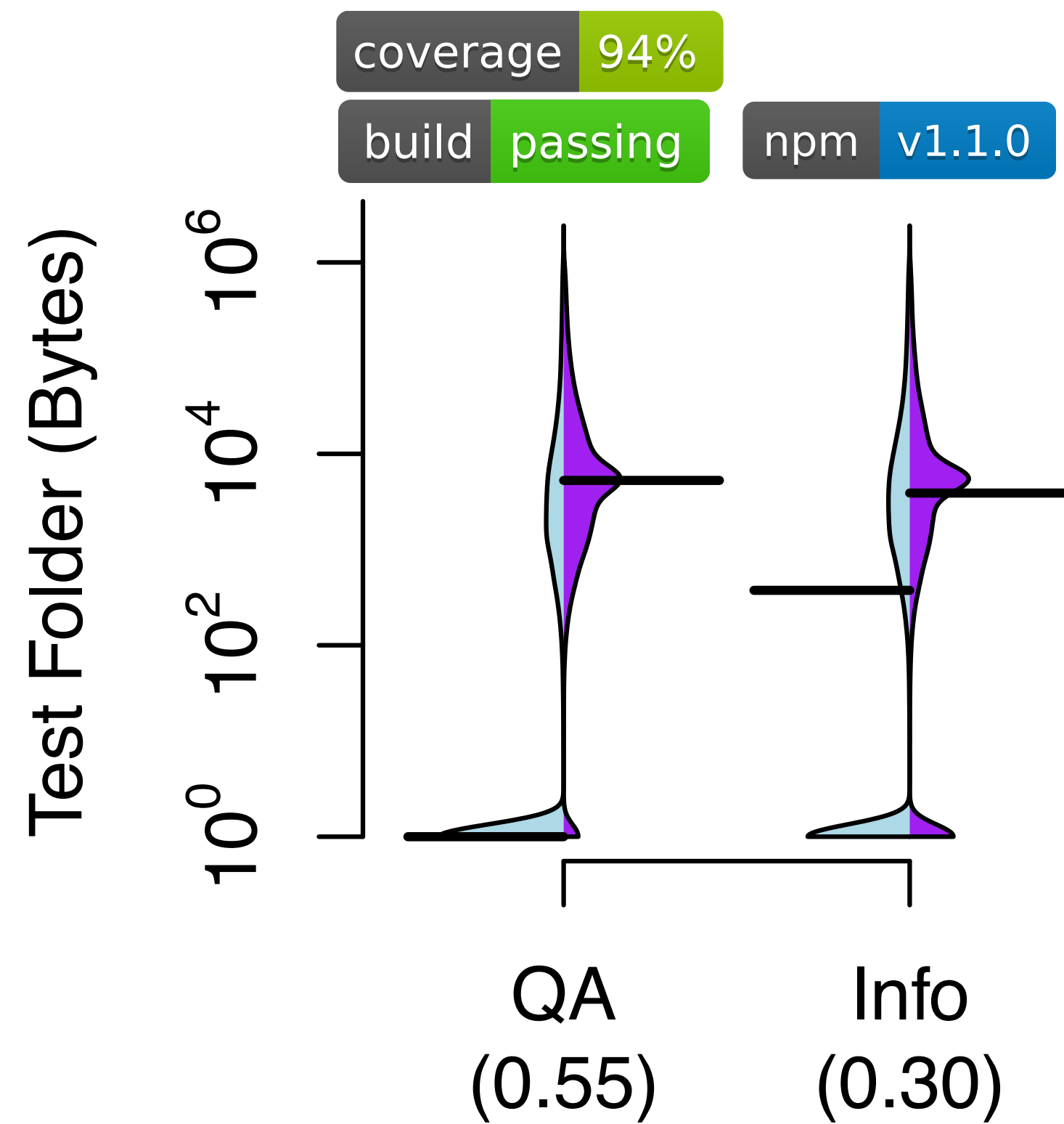
Signals of test suite quality



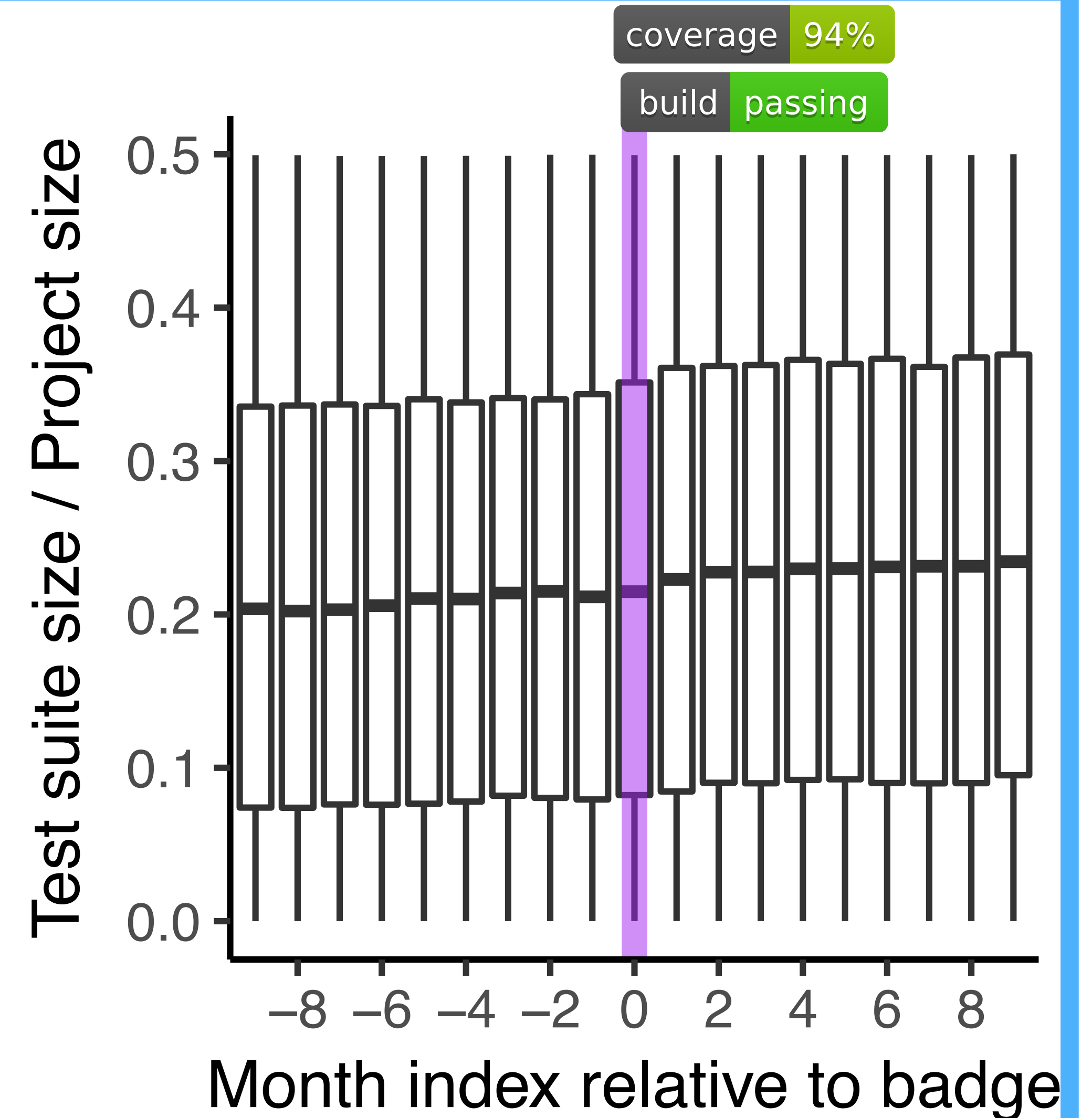
Result: Build status/code coverage badges indicate a test suite



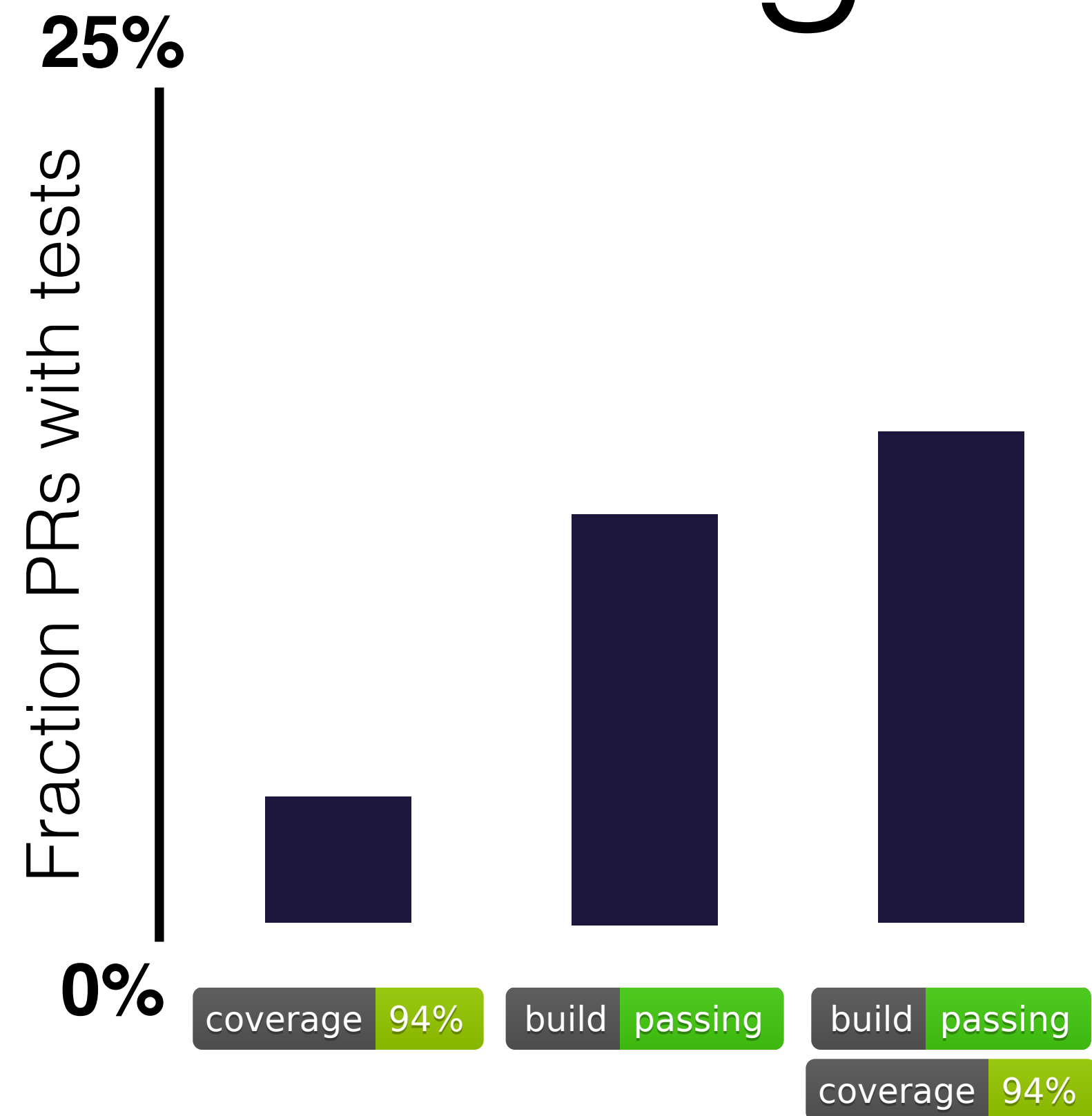
Signals of test suite quality



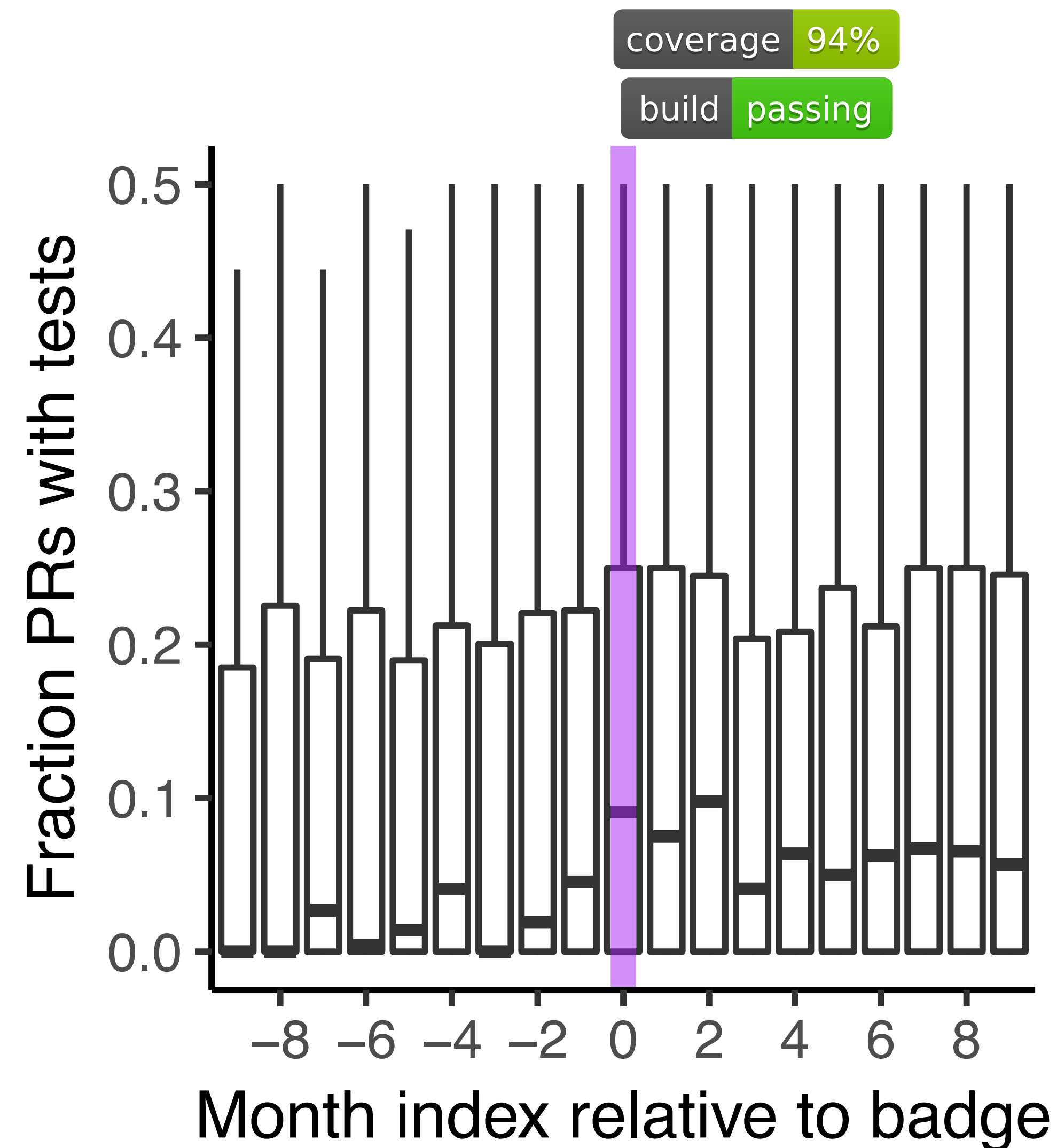
Result: Build status/code coverage badges indicate a test suite



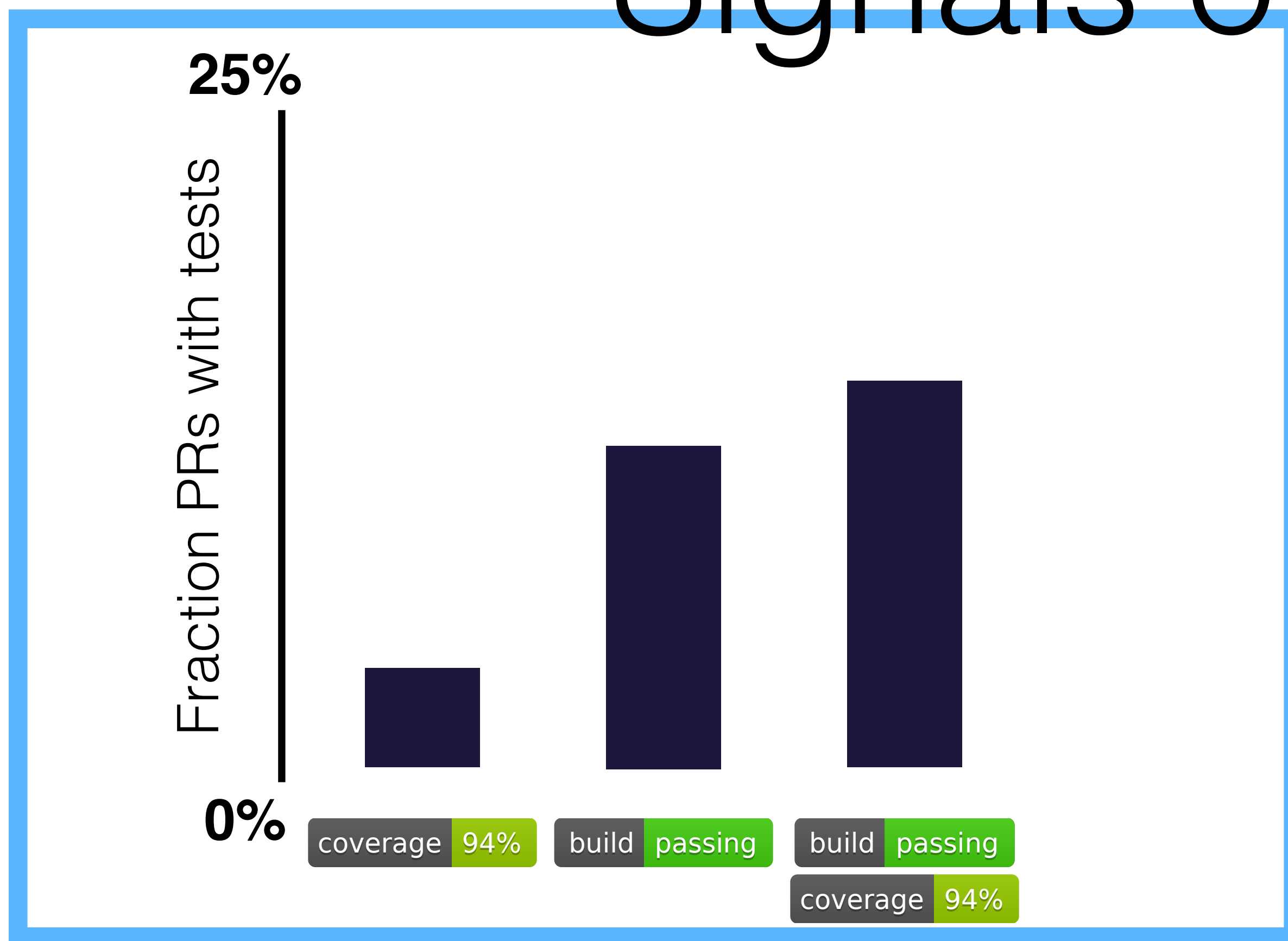
Signals of PR quality



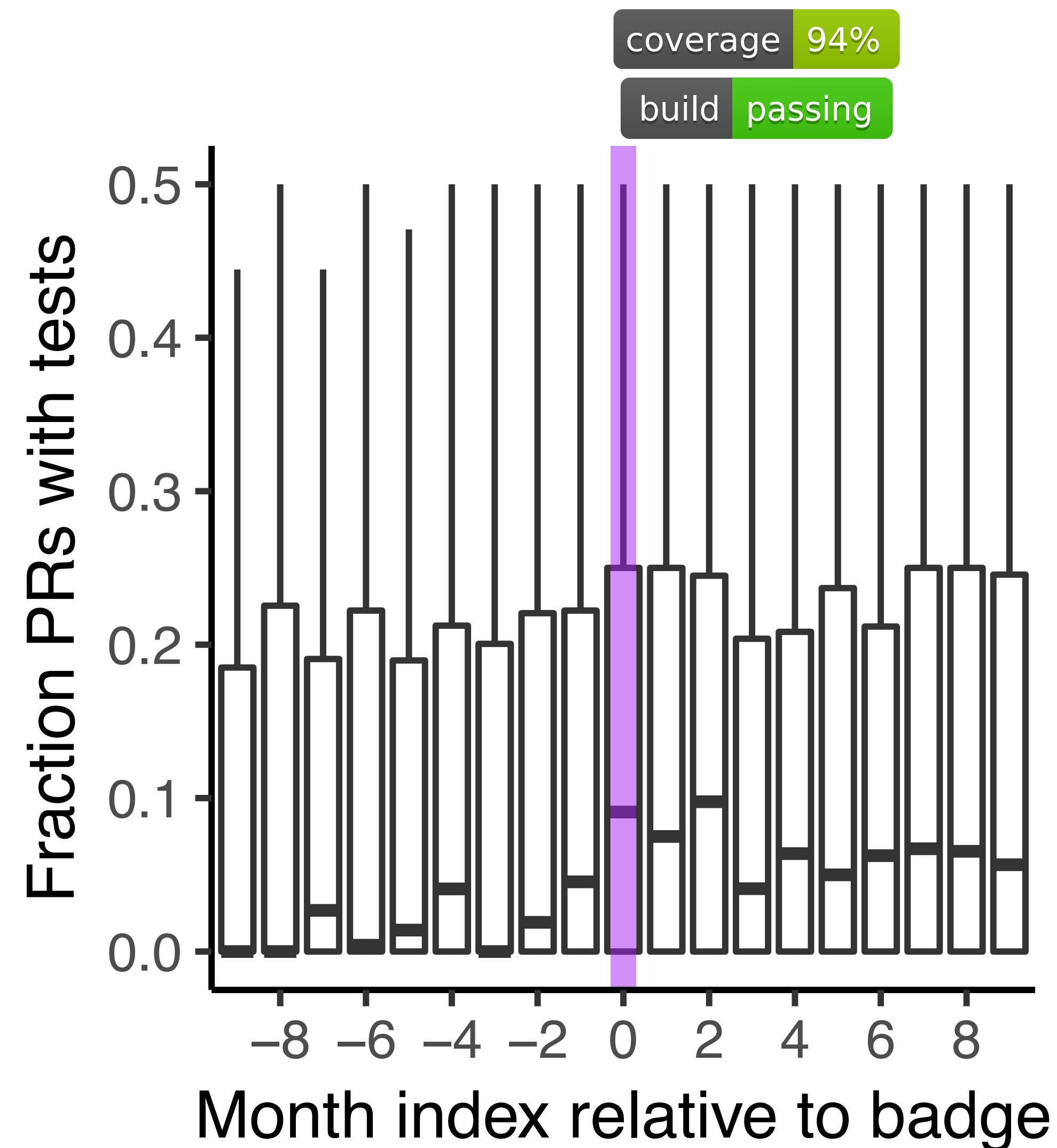
Result: Build status+code coverage badges indicate *more tests in PRs*



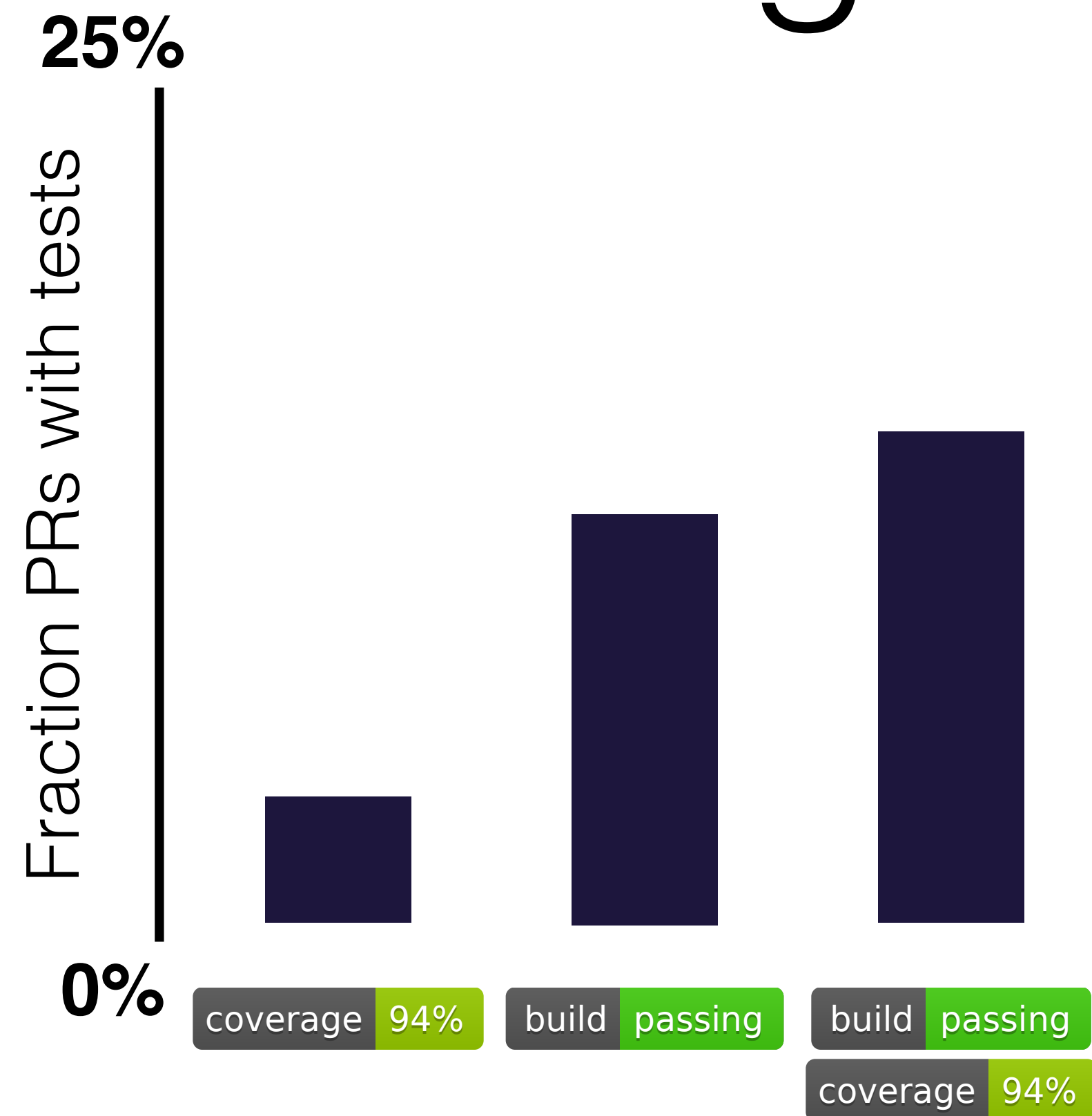
Signals of PR quality



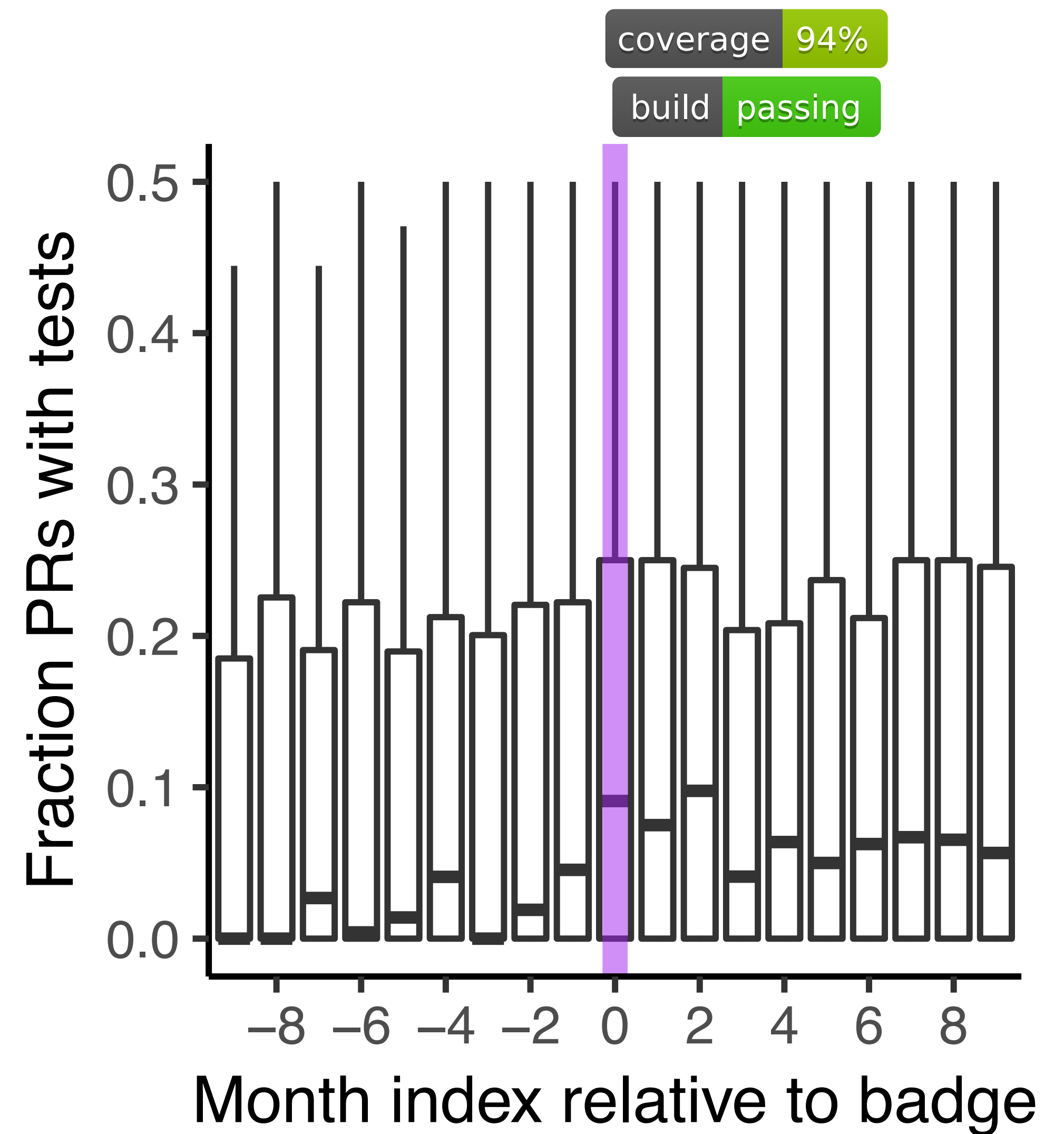
Result: Build status+code coverage badges indicate *more tests in PRs*



Signals of PR quality

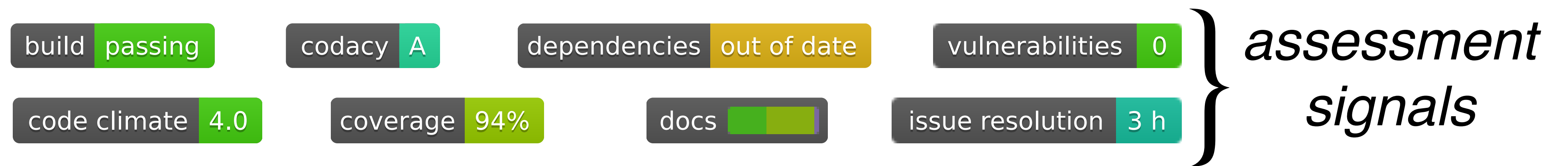


Result: Build status+code coverage badges indicate *more tests in PRs*

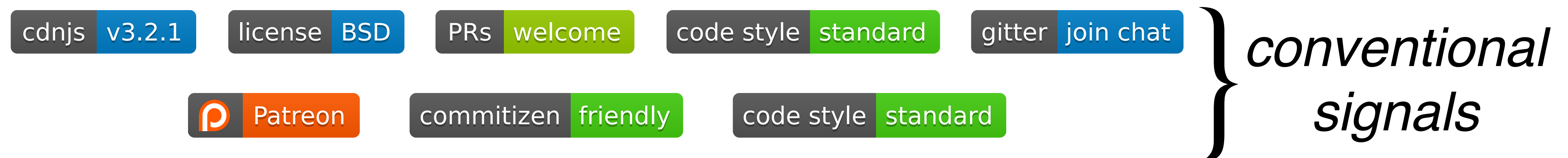


Take-aways

Badges with **underlying analyses**:



are **stronger predictors** than badges that merely state intentions or provide links:

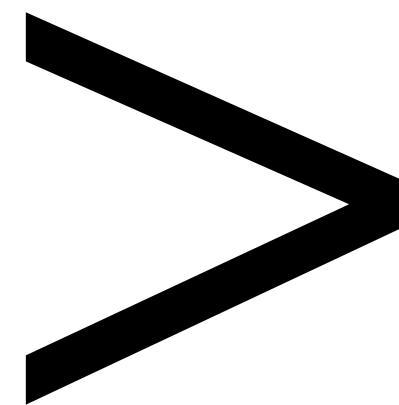


Take-aways

When possible,
design or choose the badge that takes the **most work**:



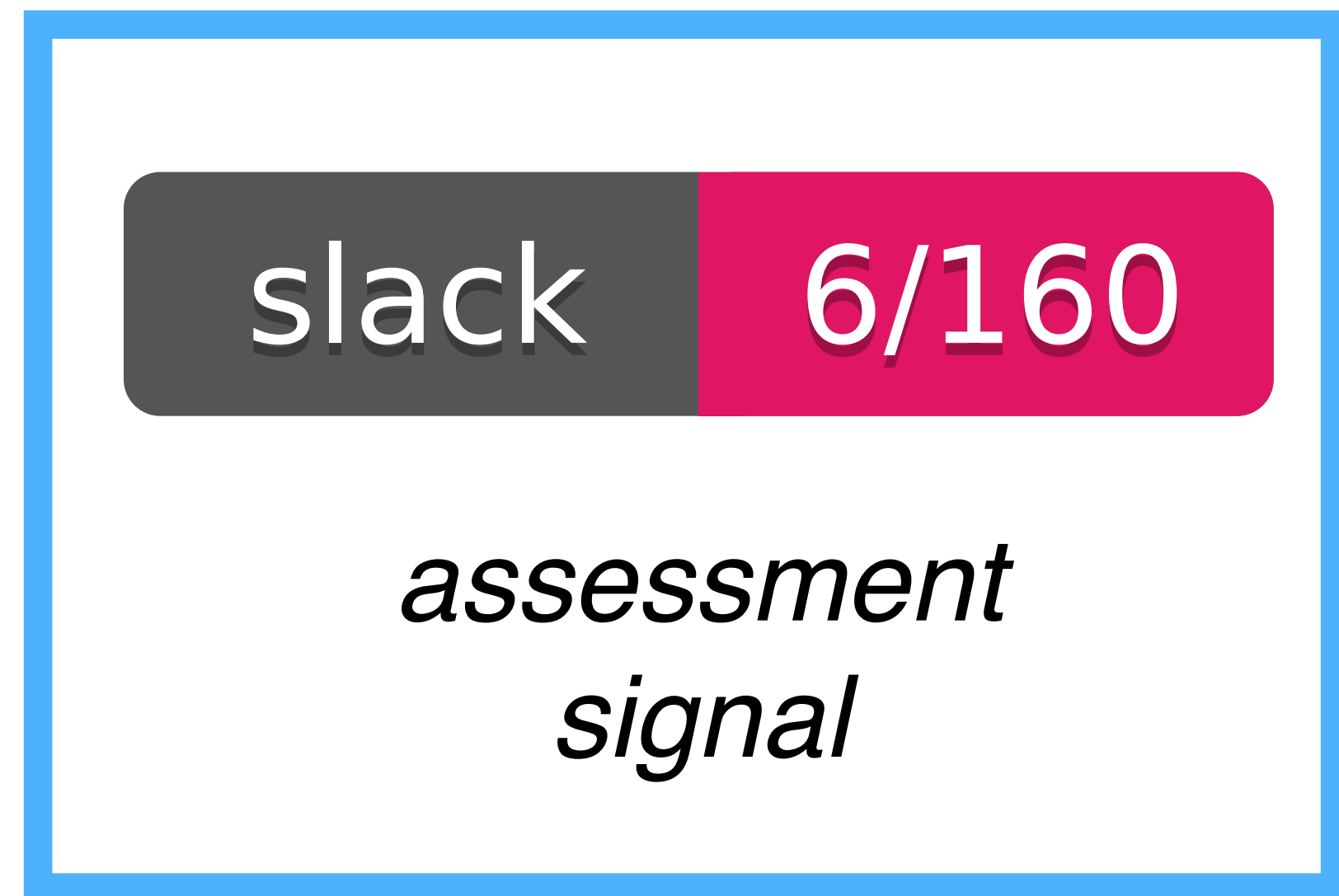
*assessment
signal*



*conventional
signal*

Take-aways

When possible,
design or choose the badge that takes the **most work**:



>

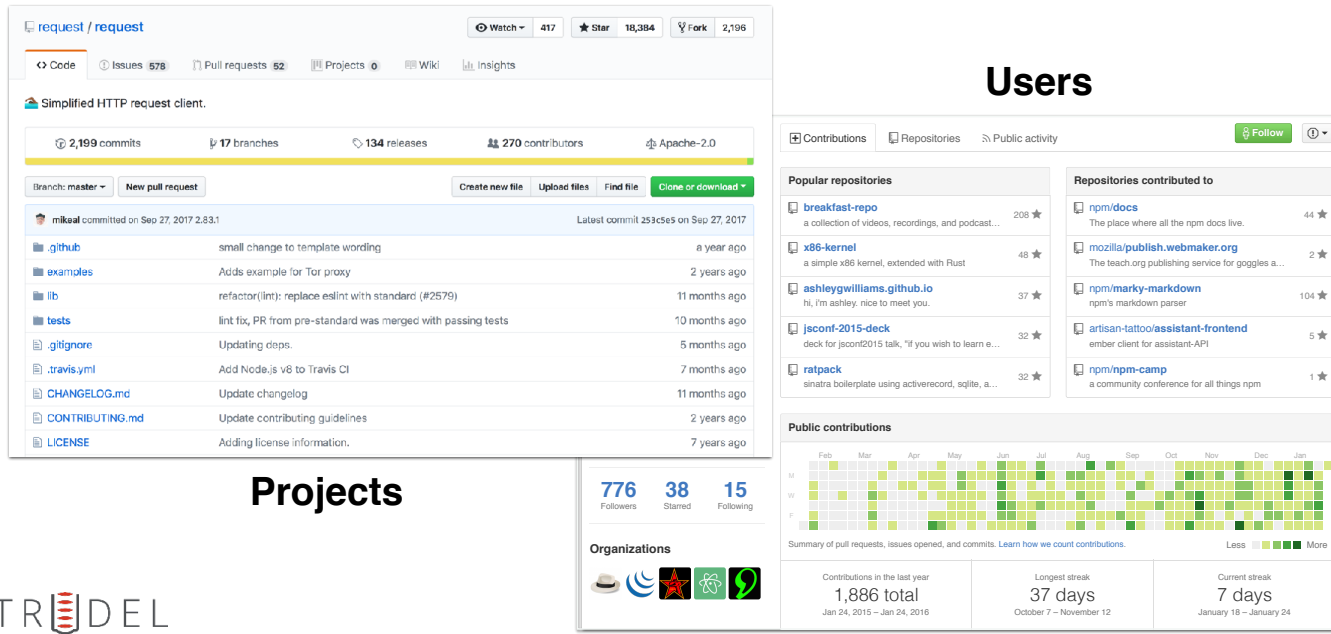


*conventional
signal*

Adding Sparkle to Social Coding:

An Empirical Study of Repository Badges in the *npm* Ecosystem


Key features: Transparency & signaling



Projects


STRIDEL
Carnegie Mellon University

Mixed methods study



Survey

- 32 maintainers, 57 contributors
- Maintainers:
 - What do you intend to signal?
 - What effects do you expect?
- Contributors:
 - What do badges tell you?

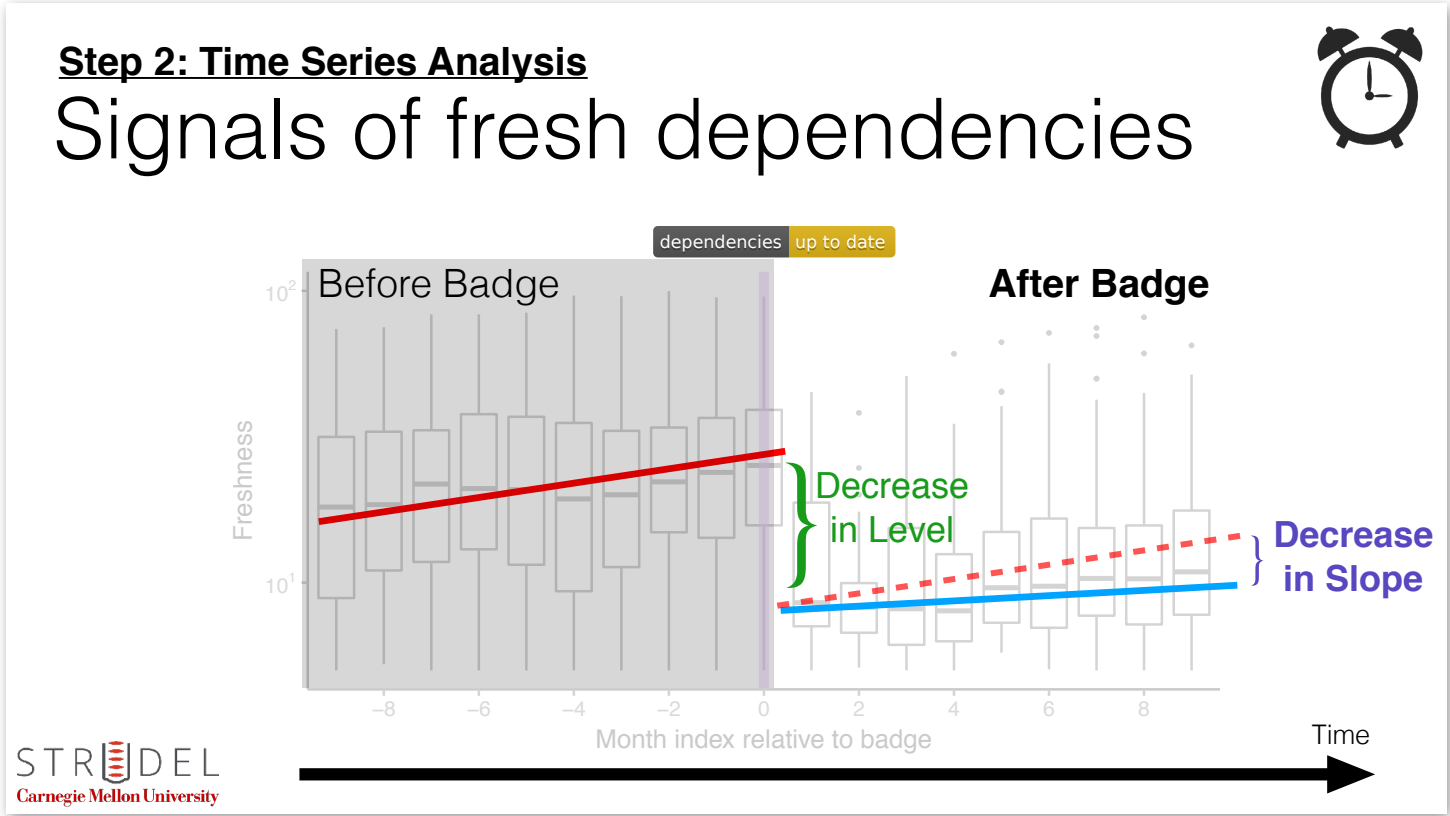


Repository Mining

- 294,941 *npm* packages
- Mined badge adoptions/removals from README files
- Measured proxies for code quality, test suite quality, popularity, dependency freshness, ...

STRIDEL
Carnegie Mellon University

Contact:
asher.trockman@gmail.com
<https://cmustrudel.github.io>



Take-aways

When possible,
design or choose the badge that takes the **most work**:

slack 6/160

assessment signal

>

slack join

conventional signal

STRIDEL
Carnegie Mellon University

badges mostly reliable

<https://cmustrudel.github.io/badges>

