

The Reproducibility of Economics Research: A Case Study

BITSS Annual Meeting 2018, Berkeley

December 10, 2018

Hautahi Kingi

IMPAQ International

Lars Vilhuber

Cornell, American Economic Association

Sylverie Herbert

Cornell

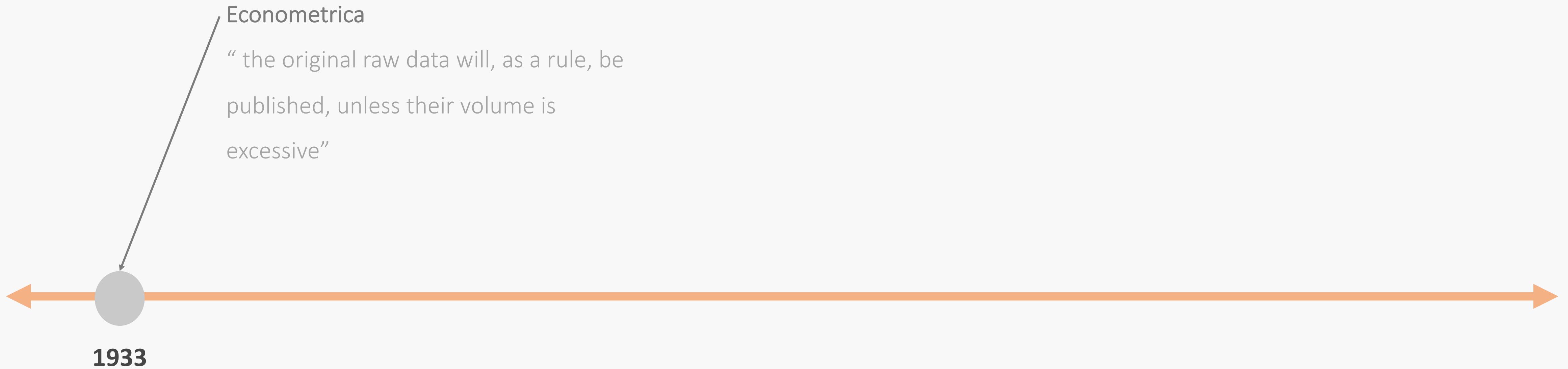
Flavio Stanchi

Cornell

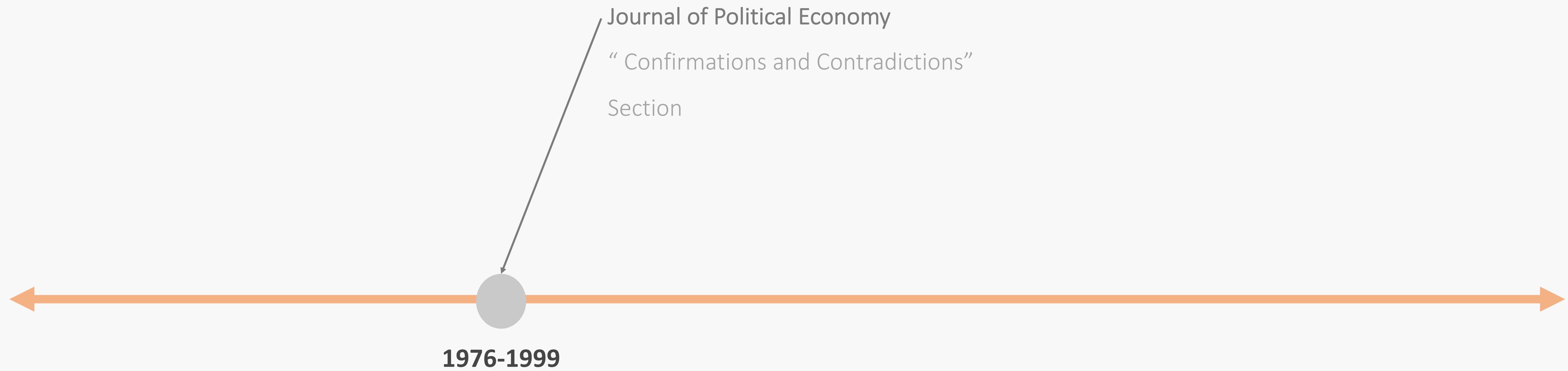


Background

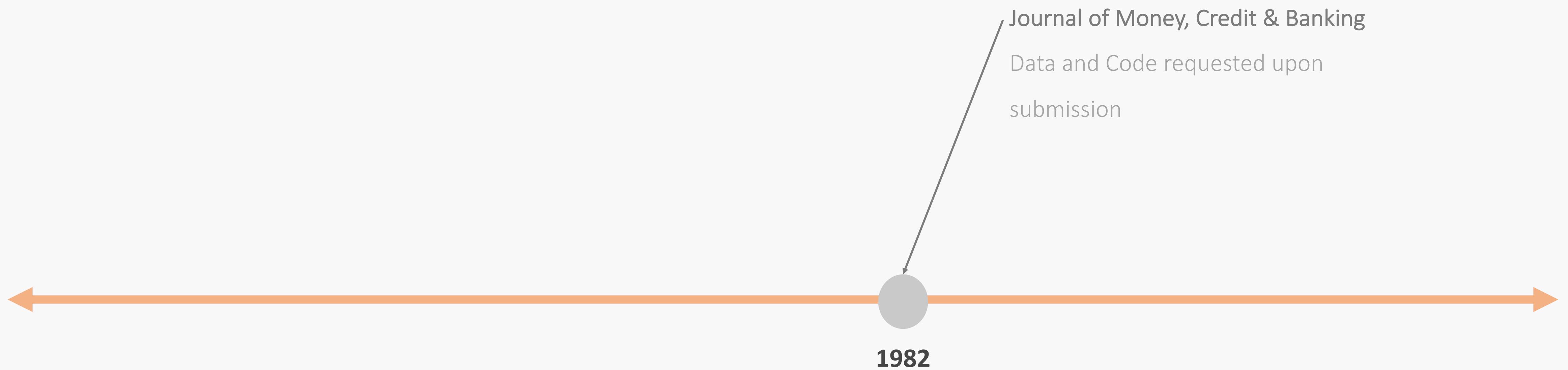
Increasing transparency in economics journals



Increasing transparency in economics journals



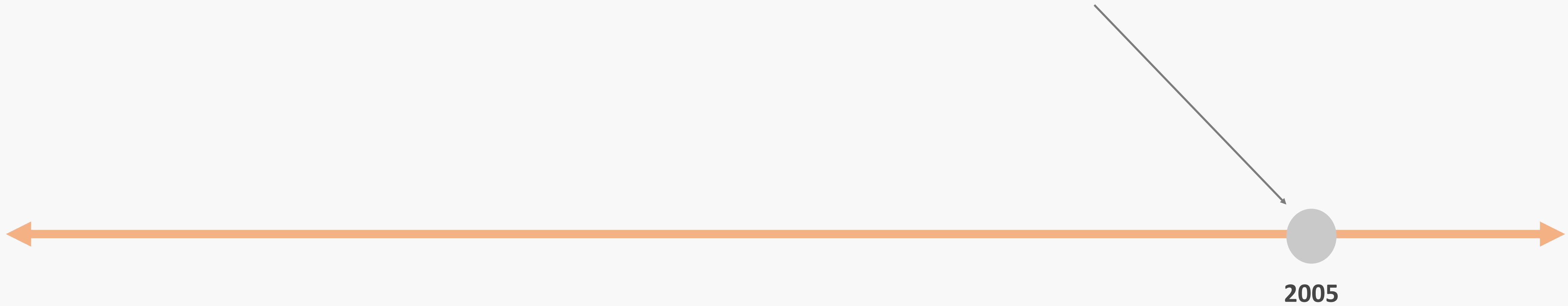
Increasing transparency in economics journals



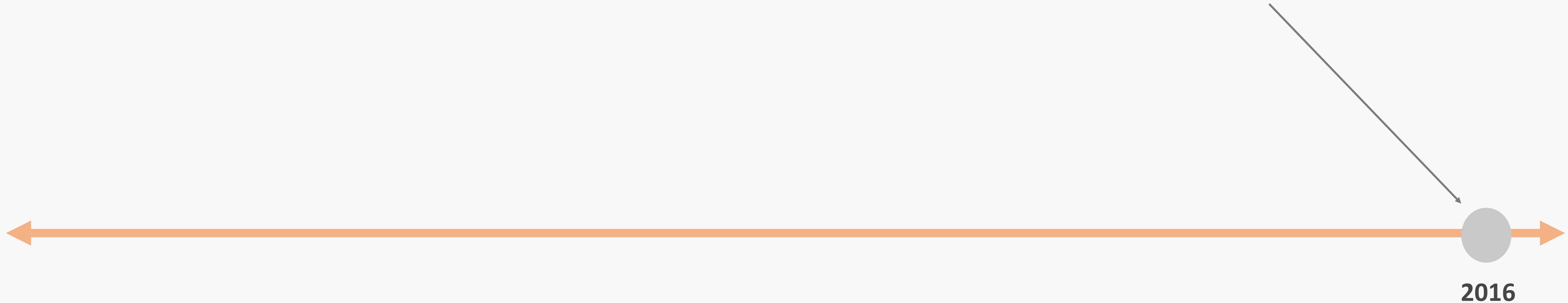
Increasing transparency in economics journals

American Economic Association

Data and code availability policies



Increasing transparency in economics journals



“Authors of accepted papers ... must provide to the Review, prior to publication, the data, programs, and other details of the computations sufficient to permit replication. ... The Editor should be notified at the time of submission if the data used in a paper are proprietary or if, for some other reason, the requirements above cannot be met.”

- AEA Data Policy

How common is replication in economics?

- 0.1% of 126,505 “top 50” published articles 1974-2014 – Mueller-Langer et al. (2018)

How common is replication in economics?

- 0.1% of 126,505 “top 50” published articles 1974-2014 – Mueller-Langer et al. (2018)
- 6.2% of 1,138 “top 10” empirical development articles 2000-2015 - Sukhtankar (2017)

How common is replication in economics?

- 0.1% of 126,505 “top 50” published articles 1974-2014 – Mueller-Langer et al. (2018)
- 6.2% of 1,138 “top 10” empirical development articles 2000-2015 - Sukhtankar (2017)
- 16 journals have published > 3 replications - Duvendack et al. (2017)

How common is replication in economics?

- 0.1% of 126,505 “top 50” published articles 1974-2014 – Mueller-Langer et al. (2018)
- 6.2% of 1,138 “top 10” empirical development articles 2000-2015 - Sukhtankar (2017)
- 16 journals have published > 3 replications - Duvendack et al. (2017)
- But some disagree - Hamermesh (2017)

Replication in economics is “usually not successful”

+ 13%

- Dewald, Thursby &
Anderson (1986)

Replication in economics is “usually not successful”

+ 13%

- Dewald, Thursby & Anderson (1986)

+ 6%

- McCullough et al. (2006)

Replication in economics is “usually not successful”

+ 13%

- Dewald, Thursby & Anderson (1986)

+ 6%

- McCullough et al. (2006)

+ 43%

- Chang & Li (2015)

Replication in economics is “usually not successful”

+ 13%

- Dewald, Thursby & Anderson (1986)

+ 6%

- McCullough et al. (2006)

+ 43%

- Chang & Li (2015)

+ 61%

- Camerer et al. (2016)

Replication continuum (Bollen et al. 2015)



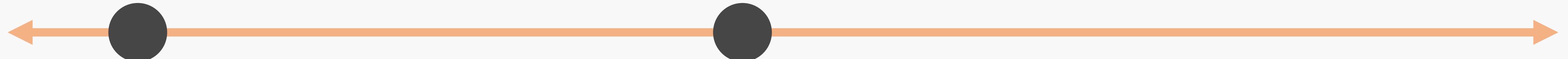
Replication continuum (Bollen et al. 2015)



Reproducibility

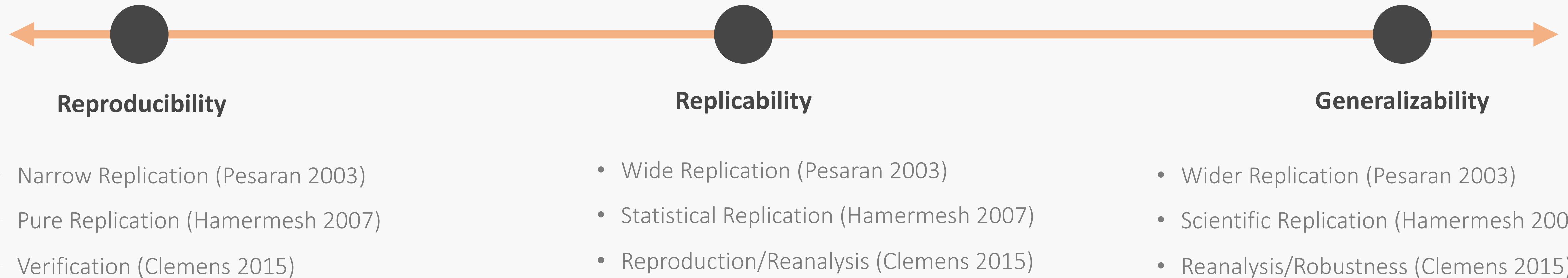
- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

Replication continuum (Bollen et al. 2015)

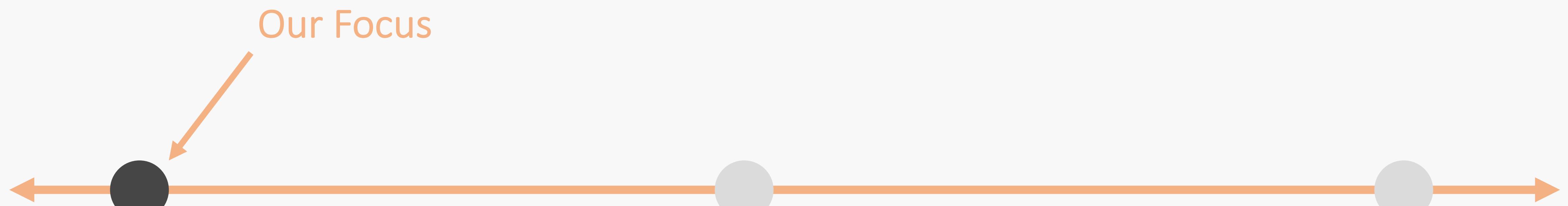


- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)
- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

Replication continuum (Bollen et al. 2015)



Replication continuum (Bollen et al. 2015)



- Narrow Replication (Pesaran 2003)
- Pure Replication (Hamermesh 2007)
- Verification (Clemens 2015)

- Wide Replication (Pesaran 2003)
- Statistical Replication (Hamermesh 2007)
- Reproduction/Reanalysis (Clemens 2015)

- Wider Replication (Pesaran 2003)
- Scientific Replication (Hamermesh 2007)
- Reanalysis/Robustness (Clemens 2015)

“We believe this basic standard should be expected of all published economics research, and hope this expectation is universal among researchers.”

- Christensen & Miguel (2018, JEL)

Method

Undergraduate Summer Research Assistants

2014 - 2018



Undergraduate Summer Research Assistants

2014 - 2018



Economics



Comp Sci.

OpRe



Undergraduate Summer Research Assistants

2014 - 2018



Economics



Comp Sci.

OpRe

Reproduction

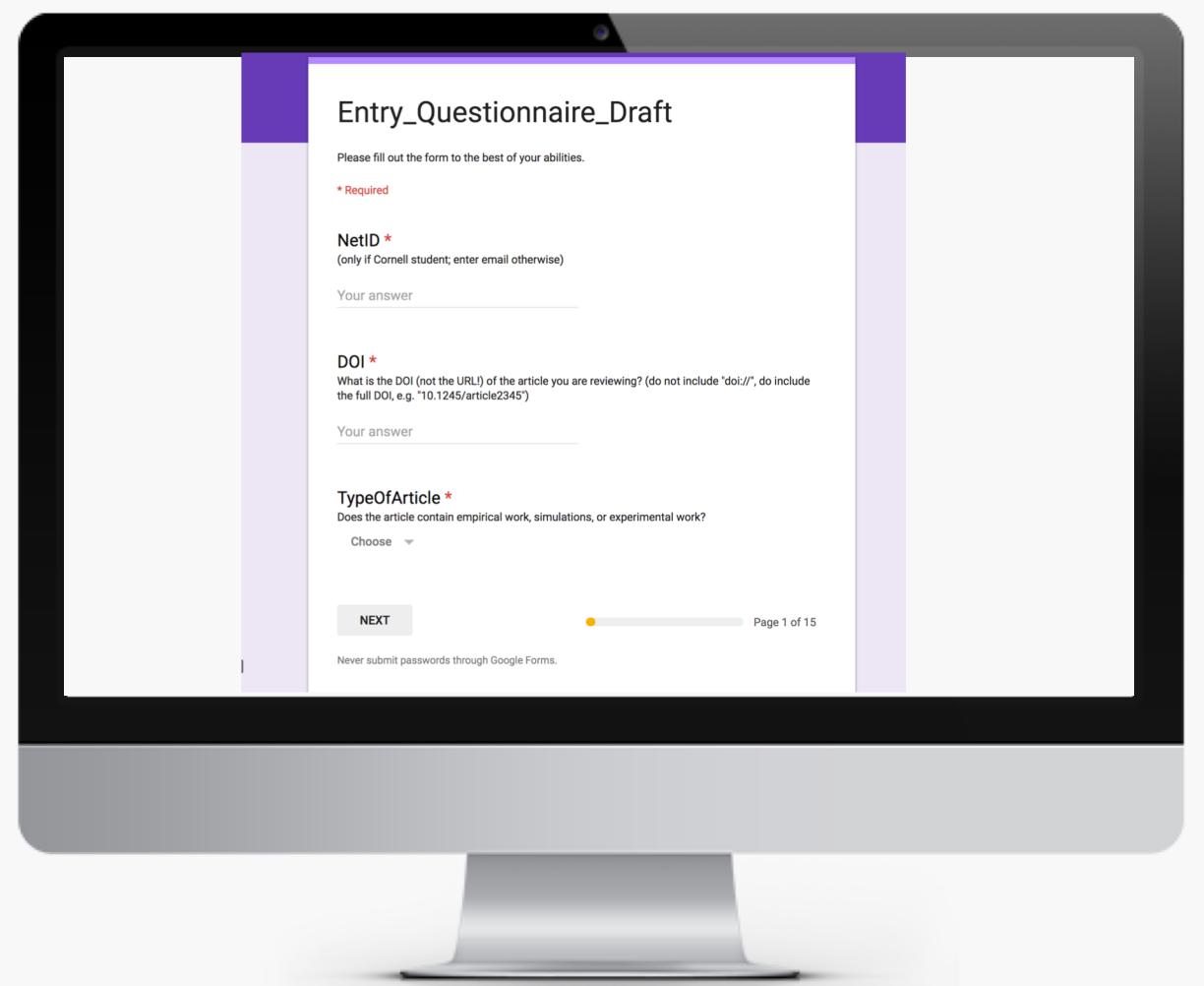


VCS

Cloud Computing

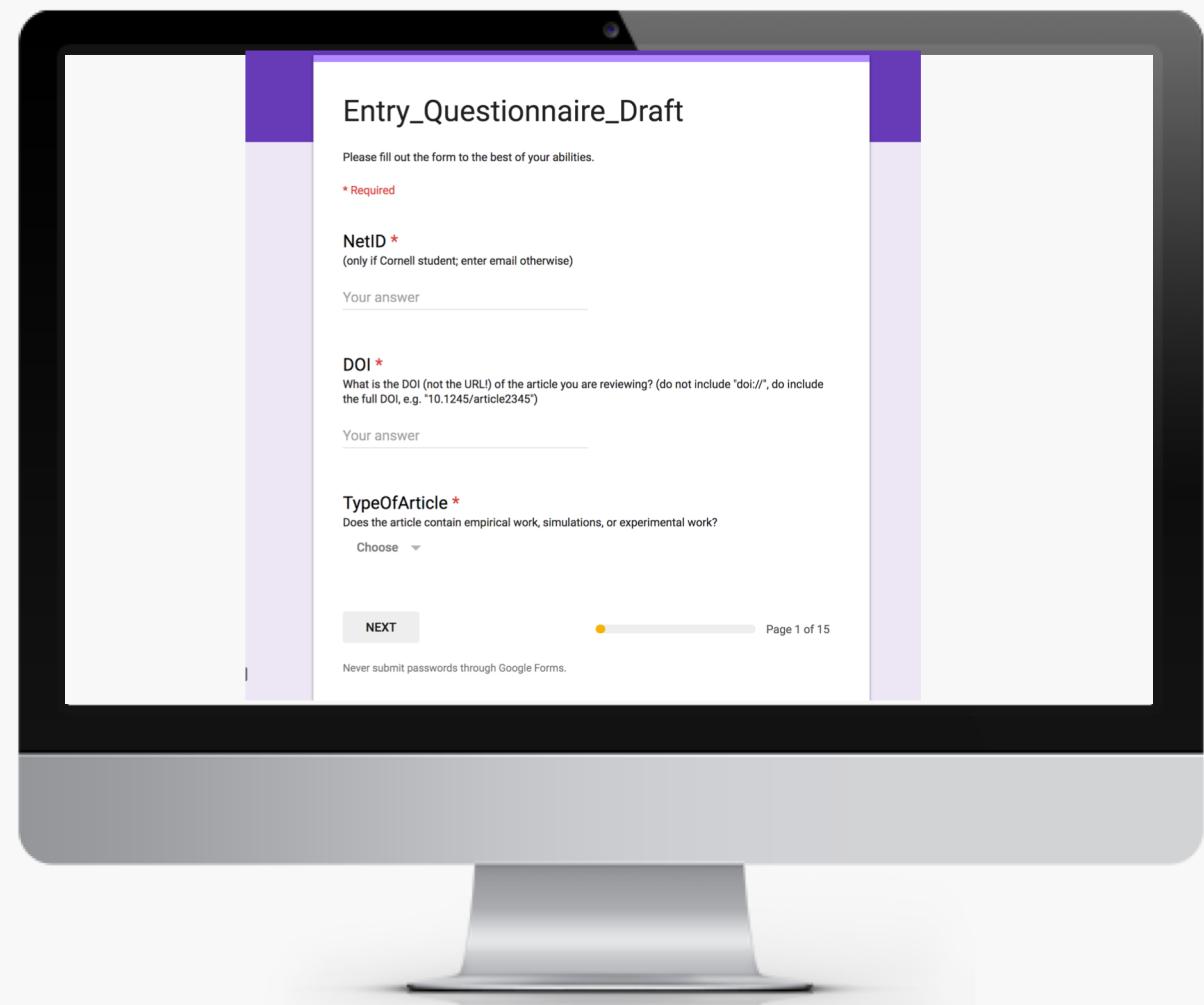


Method



≡ Assessment Form

Method

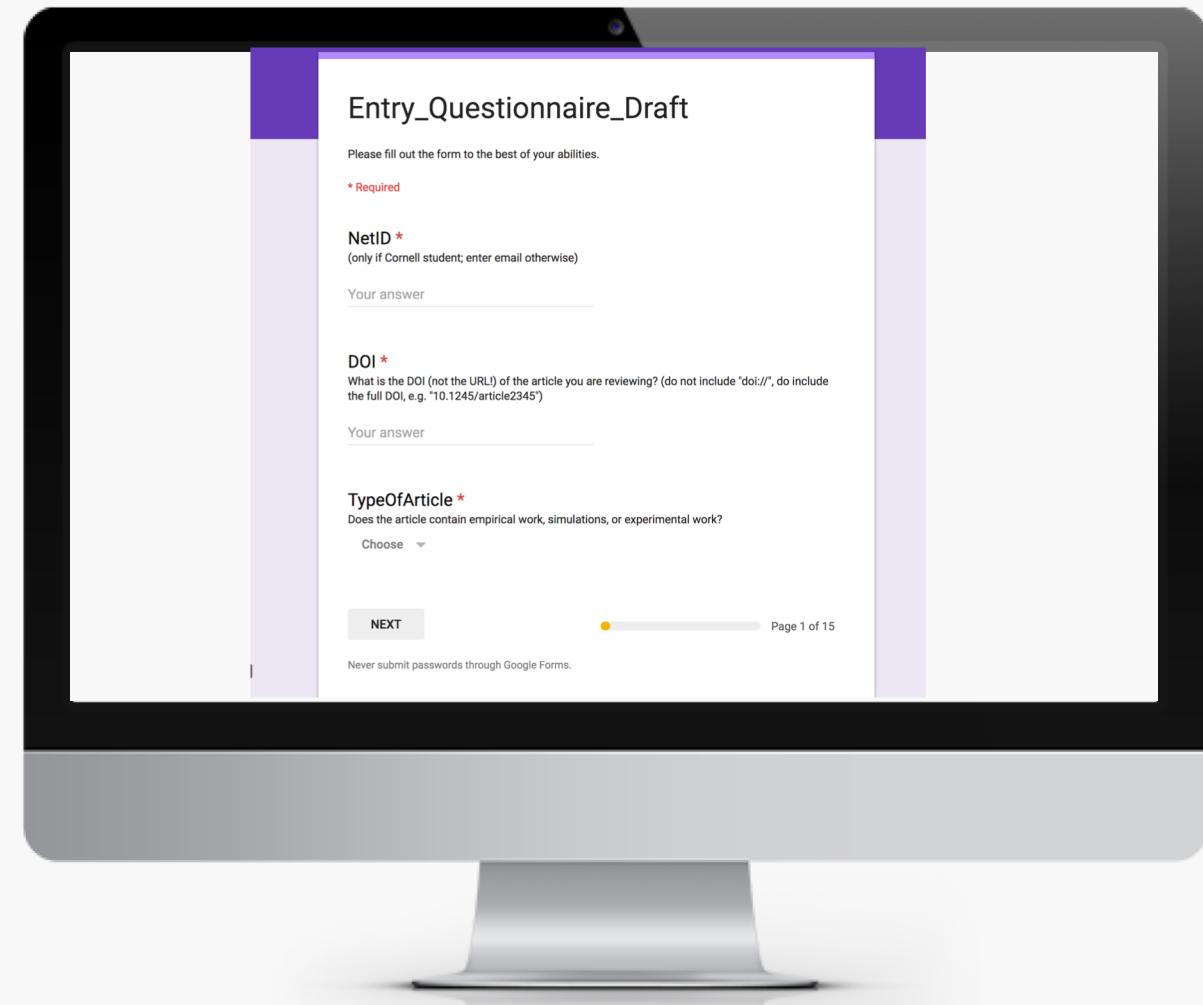


❖ Assessment Form



❖ Reproduction Exercise

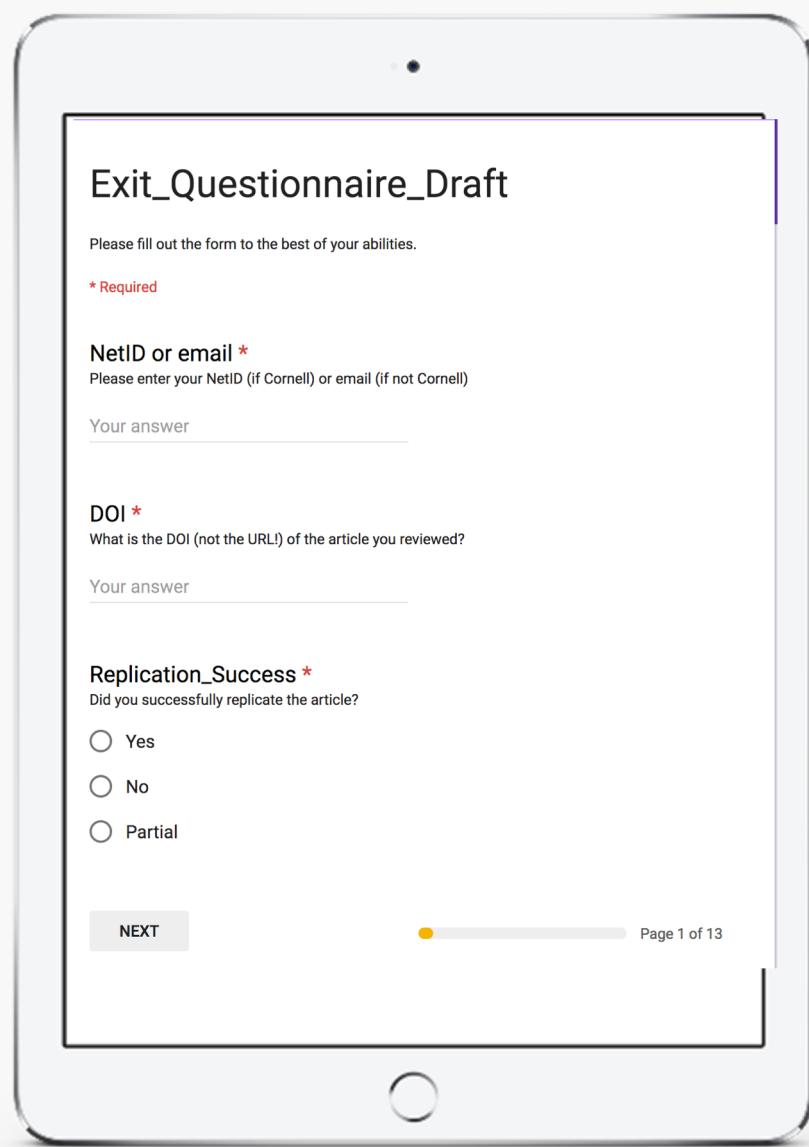
Method



❖ Assessment Form



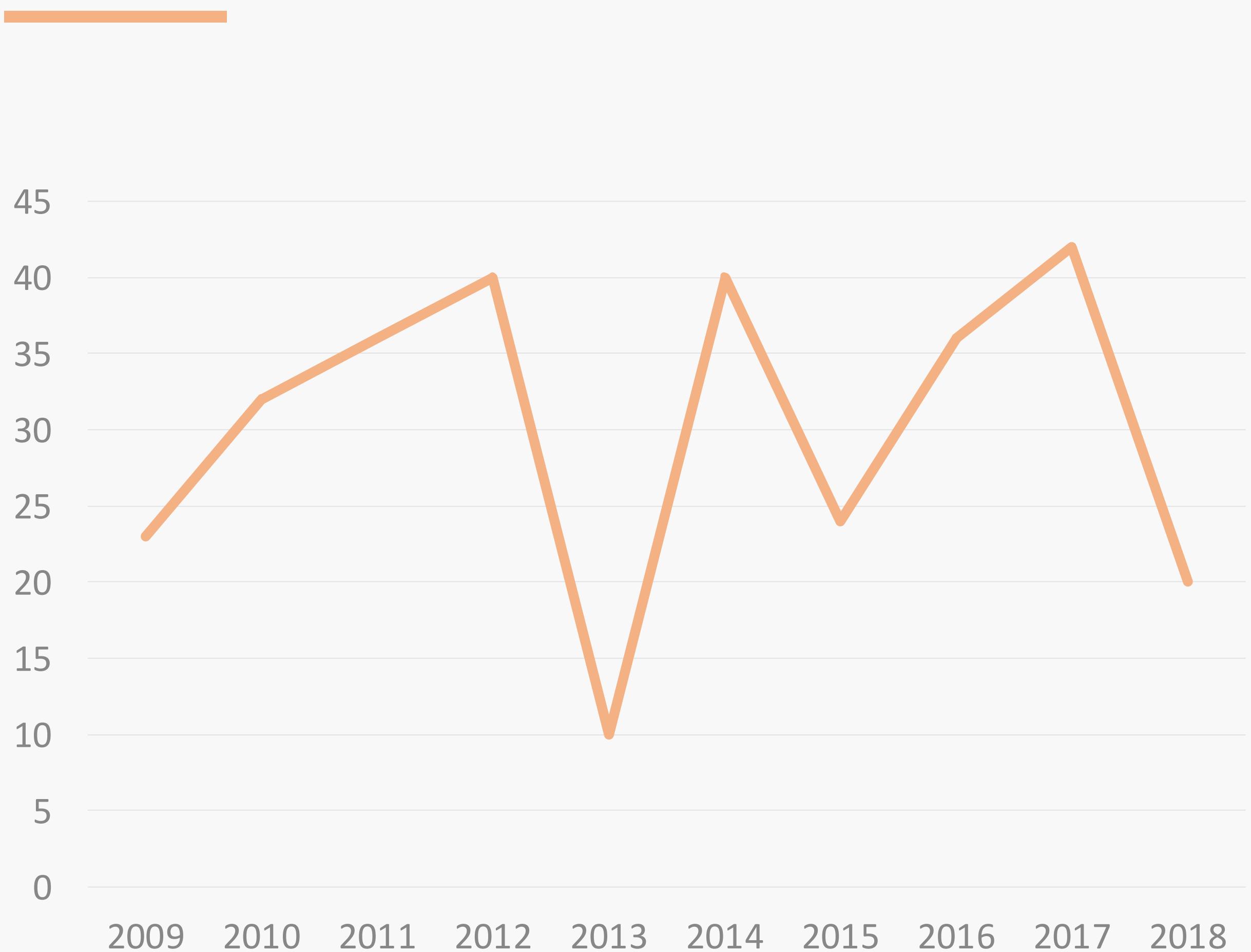
❖ Reproduction Exercise



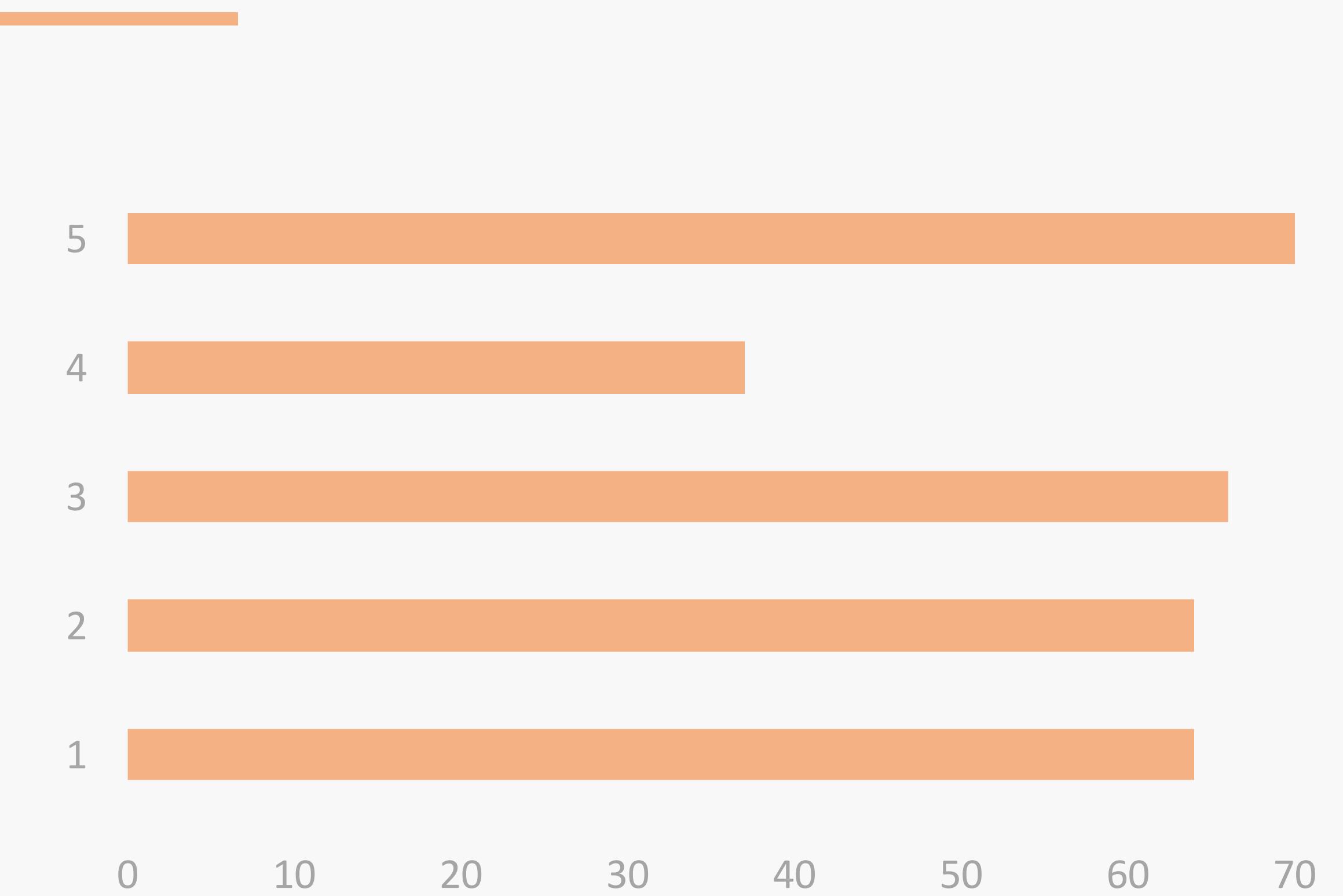
❖ Exit Questionnaire

Article Descriptions

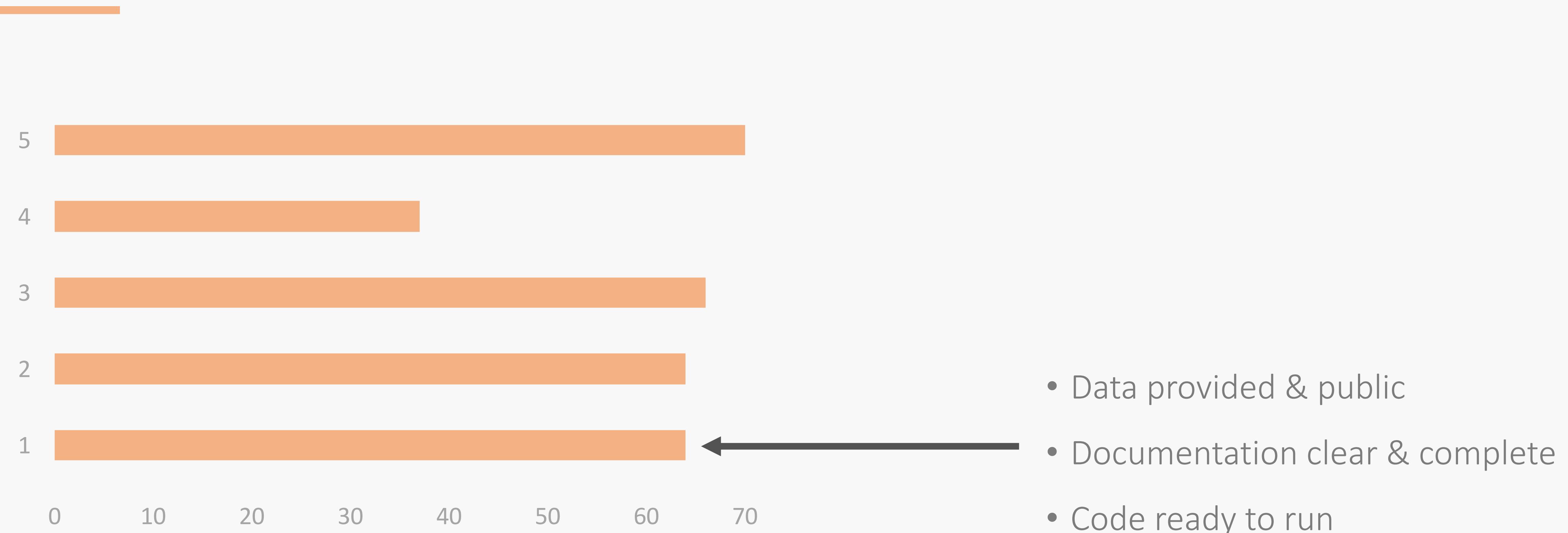
303 articles examined



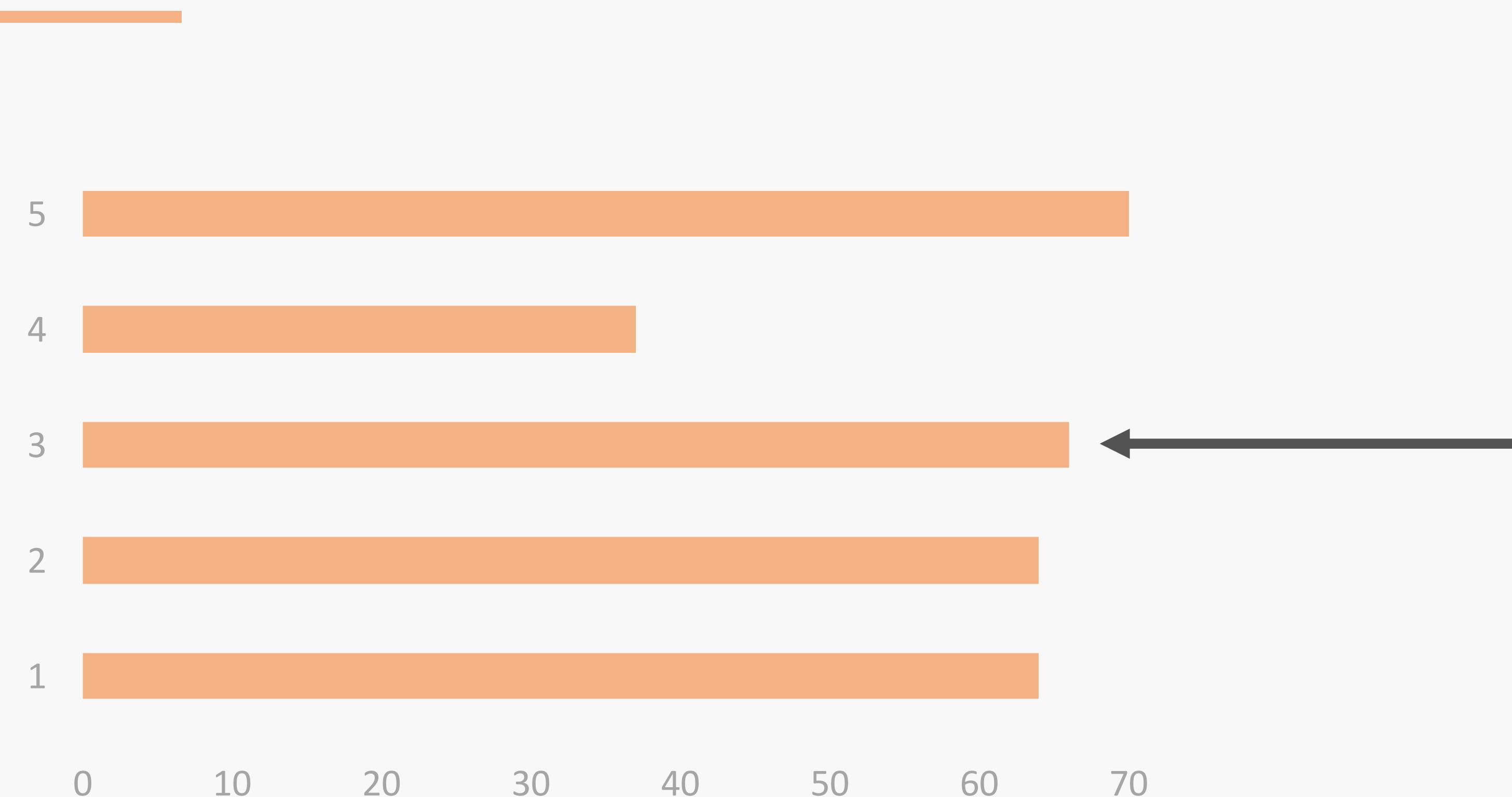
Uniform distribution of difficulty



Uniform distribution of difficulty



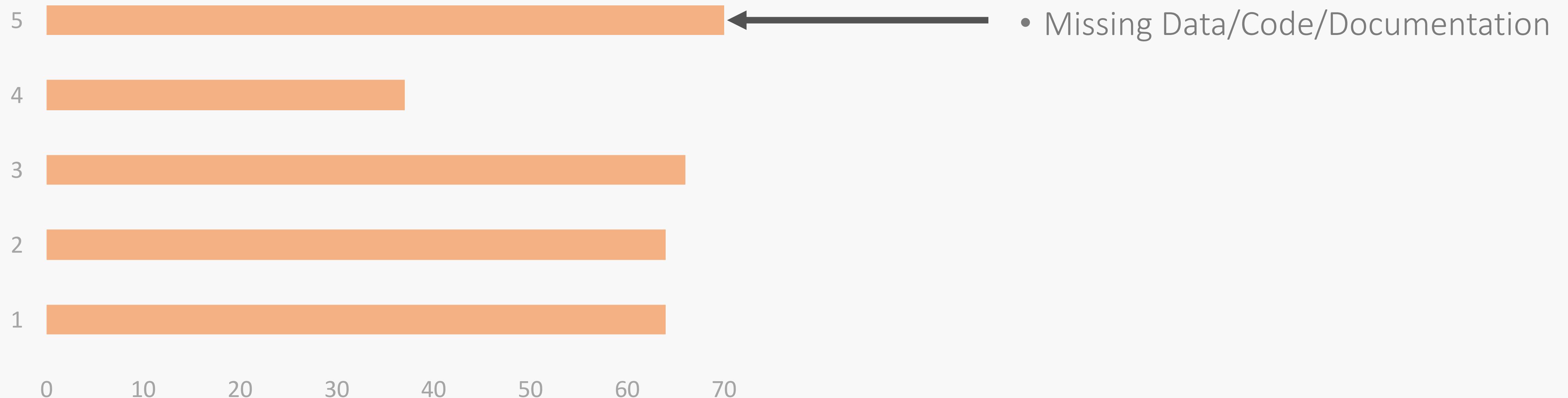
Uniform distribution of difficulty



- Documentation unclear
- Code order ambiguous

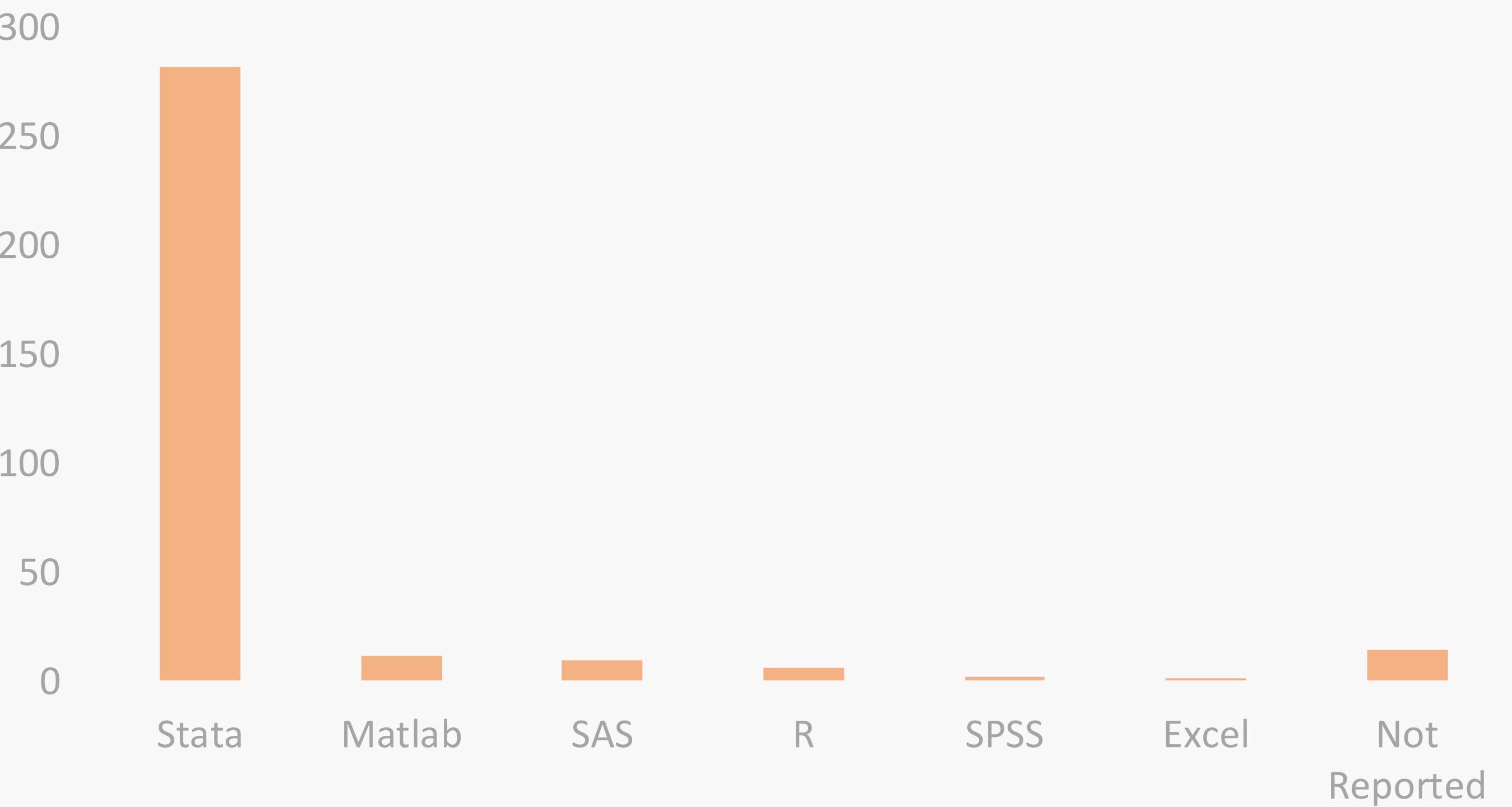
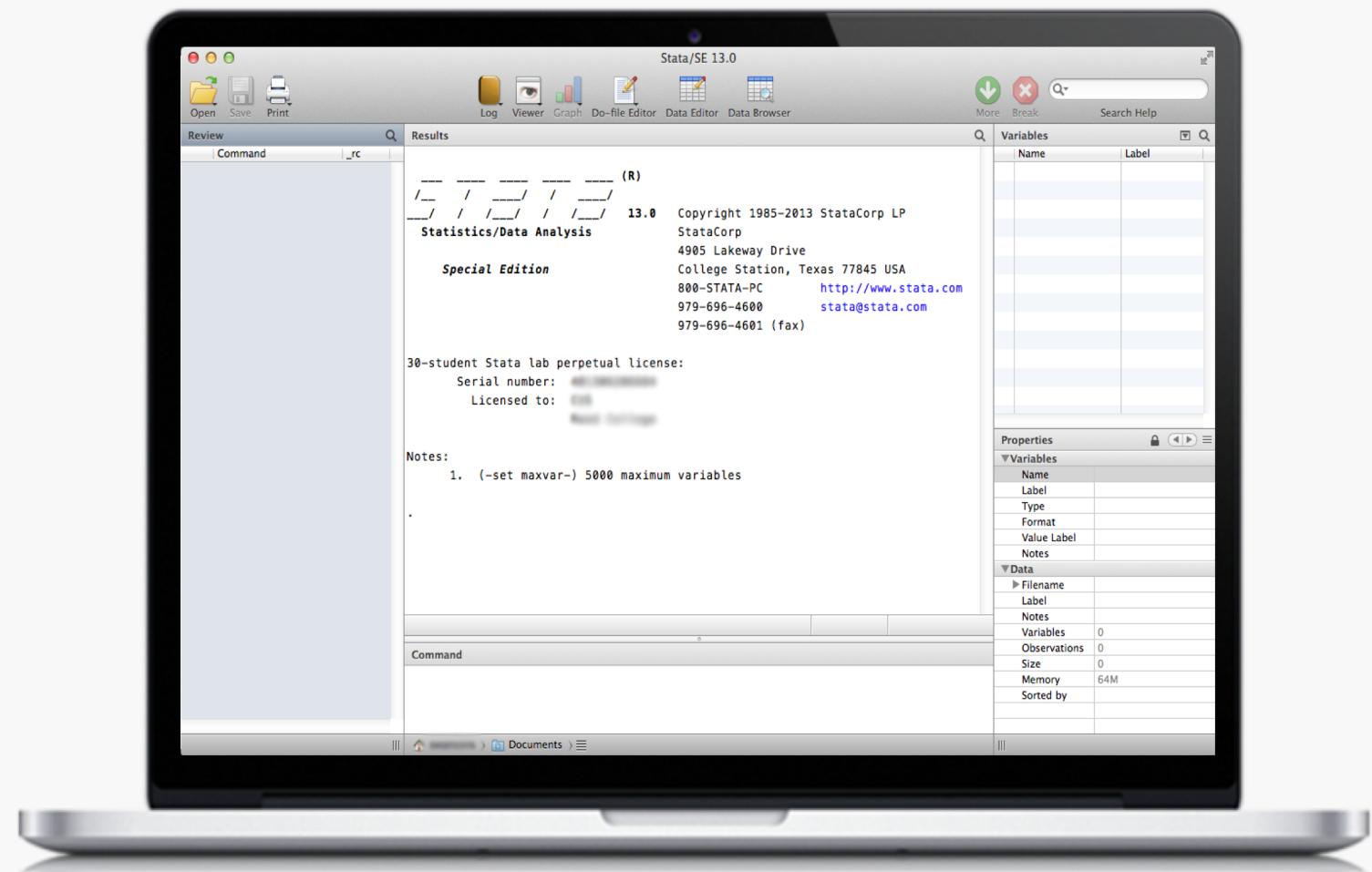
Uniform distribution of difficulty

•

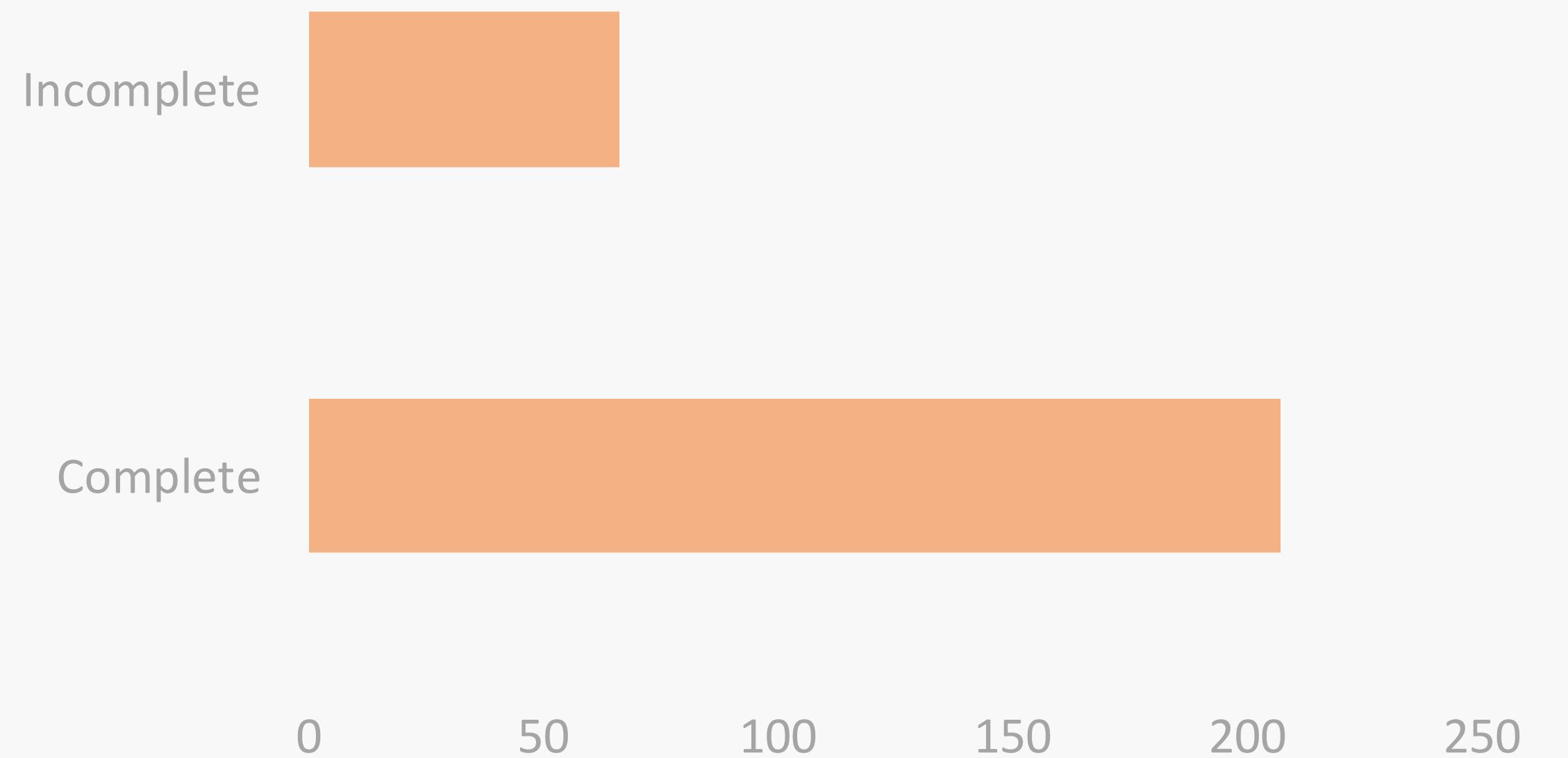


- Missing Data/Code/Documentation

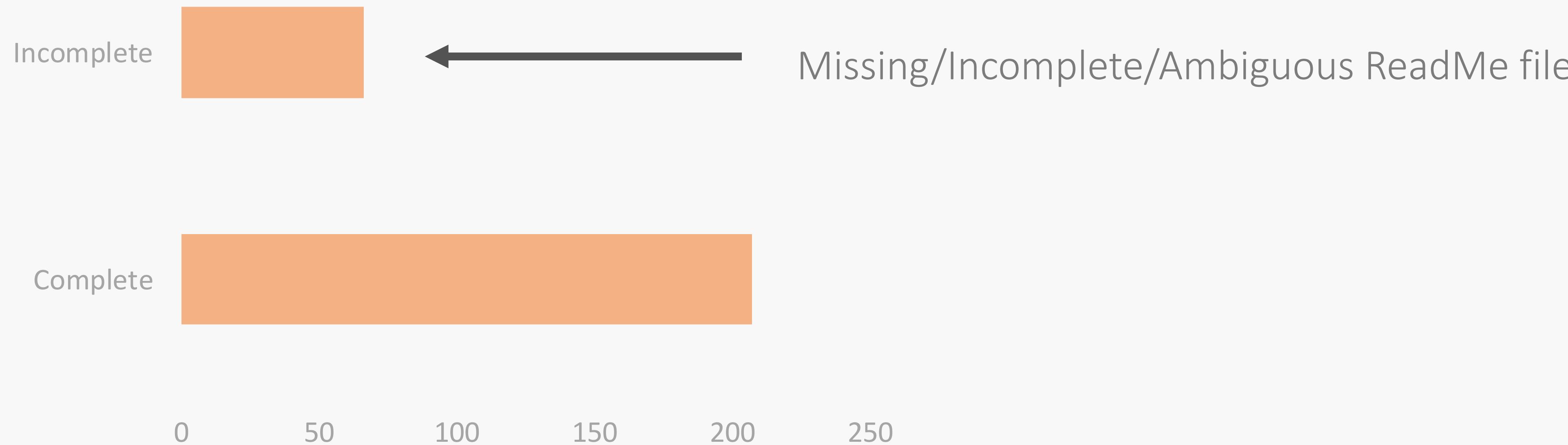
Stata dominates



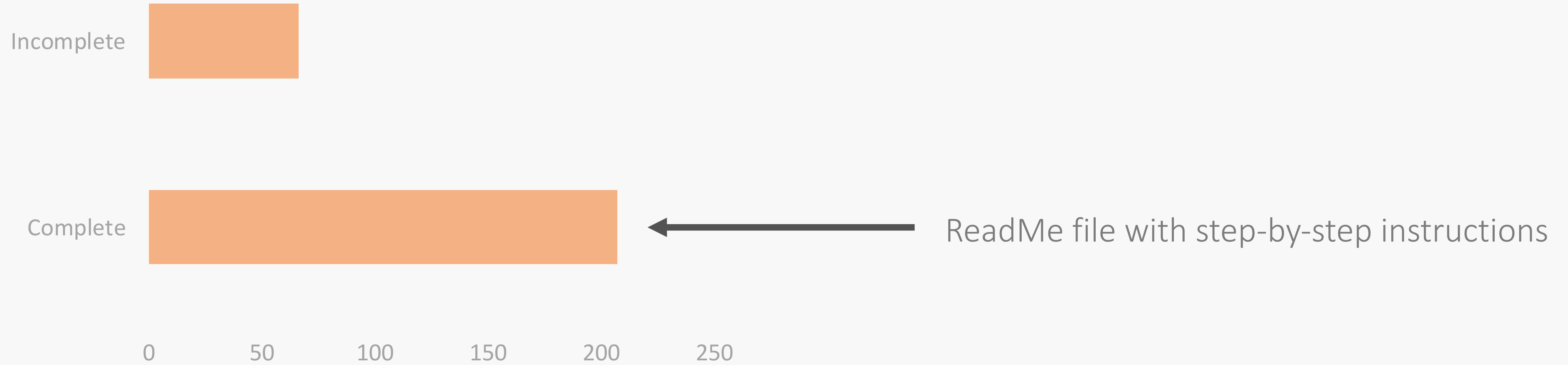
Documentation quality is decent



Documentation quality is decent



Documentation quality is decent



Results

Moderate replication success

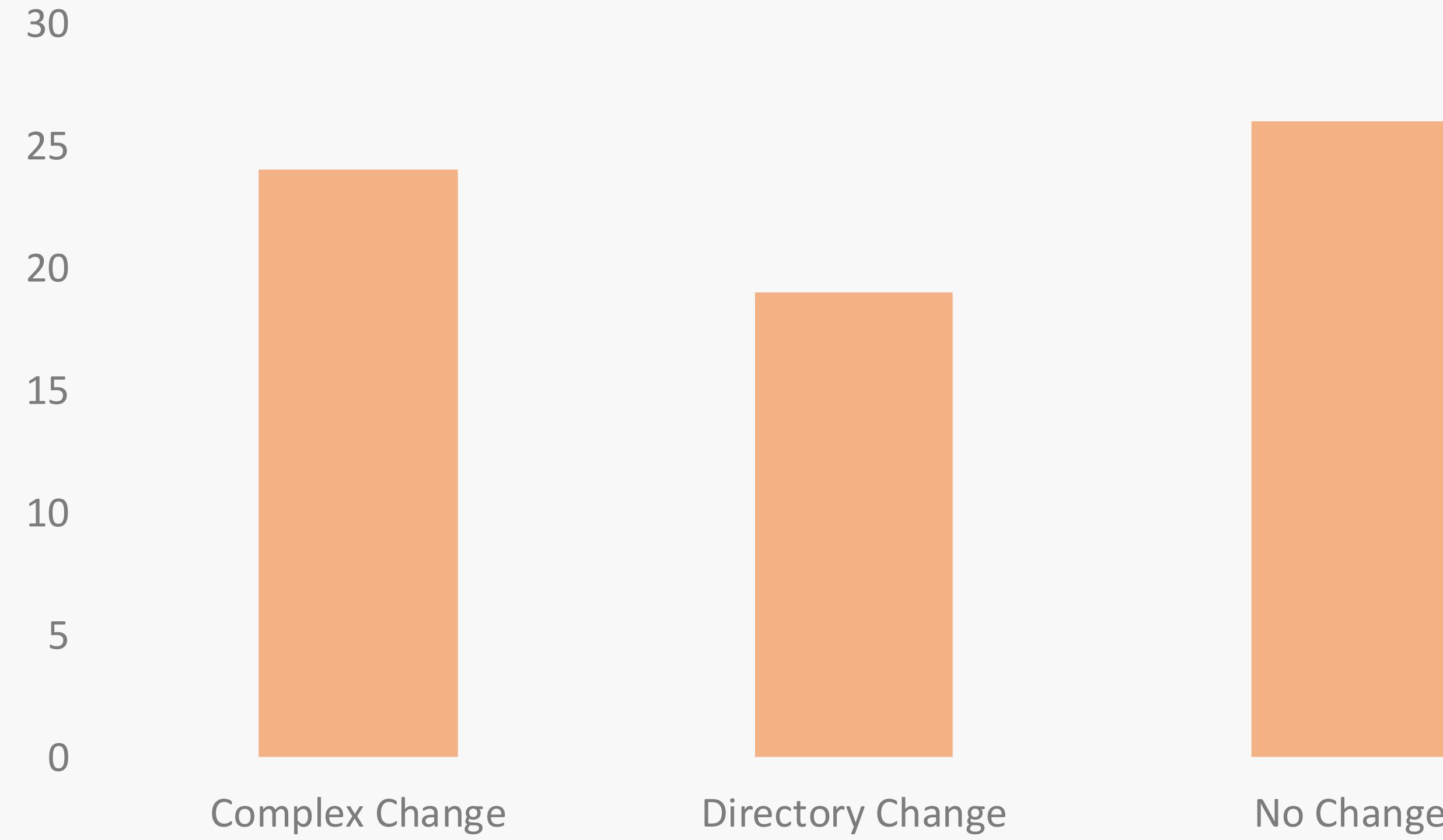


- 25% total success rate
- 33% confidential data
- 43% successful conditional on available data
- 42% partial conditional on available data

Multiple reasons for unsuccessful reproduction

Year	Missing Data	Corrupted Data	Code Error	Software Unavailable	Other
2009		1			
2010					3
2011					
2012	1	1	1		12
2013					
2014	1			1	
2015					
2016	1				
2017	1				1
2018		1			

1/3 of successful reproductions required complex code moderations



Documentation quality is positively associated with reproductive success

	Binary (1 = Successful Reproduction)
Documentation Quality	0.300*** (0.080)
Constant	0.200*** (0.070)

- OLS
- Confirming Findings/Suspicions of:
 - Chang & Li (2015)
 - McCullough et al. (2006)
 - Stark (2018)

Positive but noisy citation bonus for using confidential data

	Annual Citations		
Avg h-index	3.00***		
Top h-index		1.00***	
Low h-index			2.00
Confidential Data	17.00***	13.00*	8.00

- OLS
- High past citations indicate high future citations
- No interaction effect (not shown)

Reproducibility doesn't affect citations

	Log Annual Citations					
Avg h-index	0.20***			0.20***		
Top h-index		0.07***			0.07**	
Low h-index			0.10**			0.09
Successful Reproduction	.70*	0.50	0.50			
Successful/Partial Reproduction				0.40	0.10	-0.10

Recap

- American Economic Journal: Applied Economics
"only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication"
- Reproduction:
 - 25% Success
 - 43% Success (conditional on data)

Thanks