CWTS
Meaningful metrics

Use of the journal impact factor for assessing individual articles need not be wrong

CWTS, Leiden University

6 September 2017

L. Waltman & V.A. Traag

Universiteit Leiden

Impact Factor Debate

Impact Factor Debate

- DORA (2013)

CWTS
Meaningful metrics

Impact Factor Debate

- DORA (2013)
- Seglen (1997):
  "the most cited half of the articles (in a journal) are cited, on average, 10 times as often as the least cited half. Assigning the same score (the JIF) to all articles masks this tremendous difference–which is the exact opposite of what an evaluation is meant to achieve"

CWTS
Meaningful metrics

Impact Factor Debate

- DORA (2013)
- Seglen (1997):
  "the most cited half of the articles (in a journal) are cited, on average, 10 times as often as the least cited half. Assigning the same score (the JIF) to all articles masks this tremendous difference–which is the exact opposite of what an evaluation is meant to achieve"
- Garfield (2006):
  "Typically, when the author's work is examined, the impact factors of the journals involved are substituted for the actual citation count. Thus, the JIF is used to estimate the expected count of individual papers, which is rather dubious considering the known skewness observed for most journals"

CWTS
Meaningful metrics

Impact Factor Debate

Impact Factor Debate

- Abramo et al. (2010):
  "Citations observed at a moment too close to the date of publication will not necessarily offer a proxy of quality that is preferable to impact factor"

Impact Factor Debate

- Abramo et al. (2010):
  "Citations observed at a moment too close to the date of publication will not necessarily offer a proxy of quality that is preferable to impact factor"
- Levitt & Thelwall (2011):
  "Particularly for very recently published articles, an indicator based on the average of the standard indicator of citation and the IF of the journal . . . could form the basis of a useful indicator for peer review panels"

CWTS
Meaningful metrics

Skewness

- Skewness in citations main statistical argument against IF.
- Assumption that skewness in citations represent ‘real’ differences in value.
- However, skewness may also result from citations being a ‘noisy signal’ for value.

CWTS
Meaningful metrics

Scenario 1 (Citations accurate)

Citations in journals are clearly skewed. But why?

Scenario 1 (citations accurate)

1. Citations reasonably accurate proxy of the value of an article.
2. Value of articles in a journal is skewed distributed.
3. Skewness of journal citation distributions results mainly from 2.

Scenario 2 (citations noisy)

1. Citations are a weak proxy of the value of an article.
2. Value of articles in a journal is distributed homogeneous.
3. Skewness of journal citation distributions results mainly from 1.

Scenario 1 (Citations accurate) example

		Journal A		
		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			

Scenario 1 (Citations accurate) example

		Journal A		
		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			**100**

Scenario 1 (Citations accurate) example

Journal A			
		Citations	
		Low	High	**Total**
Value	Low			**20**
	High			**80**
	Total			**100**

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total			**100**

CWTS
Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A				
		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

CWTS
Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			**100**

Scenario 1 (Citations accurate) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low			**80**
	High			**20**
	Total			**100**

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	72	8	**80**
	High	8	18	**20**
	Total			**100**

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	72	8	**80**
	High	8	18	**20**
	Total	**74**	**26**	**100**

CWTS
Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	72	8	**80**
	High	8	18	**20**
	Total	**74**	**26**	**100**

Select papers: how many have a high value?

1. Select highly cited publications:

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low	72	8	**80**
	High	8	18	**20**
	Total	**74**	**26**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{72+18}{74+26} = \frac{90}{100} = 0.9$.

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

Value	Citations: Low	Citations: High	**Total**
Low	18	2	**20**
High	8	72	**80**
Total	**26**	**74**	**100**

Journal B

Value	Citations: Low	Citations: High	**Total**
Low	72	8	**80**
High	8	18	**20**
Total	**74**	**26**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{72+18}{74+26} = \frac{90}{100} = 0.9$.
2. Select publications from high impact journal:

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

Value	Citations: Low	Citations: High	**Total**
Low	18	2	**20**
High	8	72	**80**
Total	**26**	**74**	**100**

Journal B

Value	Citations: Low	Citations: High	**Total**
Low	72	8	**80**
High	8	18	**20**
Total	**74**	**26**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{72+18}{74+26} = \frac{90}{100} = 0.9$.
2. Select publications from high impact journal: $\frac{8+72}{100} = 0.8$.

CWTS Meaningful metrics

Scenario 1 (Citations accurate) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	18	2	**20**
	High	8	72	**80**
	Total	**26**	**74**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low	72	8	**80**
	High	8	18	**20**
	Total	**74**	**26**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{72+18}{74+26} = \frac{90}{100} = 0.9$.
2. Select publications from high impact journal: $\frac{8+72}{100} = 0.8$.

Citations more accurate.

CWTS Meaningful metrics

Scenario 2 (Citations noisy)

Citations in journal are clearly skewed. But why?

Scenario 1 (citations accurate)

1. Citations reasonably accurate proxy of the value of an article.
2. Value of articles in a journal is skewed distributed.
3. Skewness of journal citation distributions results mainly from 2.

Scenario 2 (citations noisy)

1. Citations are a weak proxy of the value of an article.
2. Value of articles in a journal is distributed homogeneous.
3. Skewness of journal citation distributions results mainly from 1.

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

		Journal A		
		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			

Scenario 2 (Citations noisy) example

		Journal A		
		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			**100**

Scenario 2 (Citations noisy) example

Journal A			
		Citations	
		Low	High	**Total**
Value	Low			**20**
	High			**80**
	Total			**100**

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total			**100**

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A				
		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low			
	High			
	Total			**100**

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

Value	Citations: Low	Citations: High	**Total**
Low	14	6	**20**
High	24	56	**80**
Total	**38**	**62**	**100**

Journal B

Value	Citations: Low	Citations: High	**Total**
Low			**80**
High			**20**
Total			**100**

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total			**100**

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total	**62**	**38**	**100**

CWTS
Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

Value	Citations: Low	Citations: High	**Total**
Low	14	6	**20**
High	24	56	**80**
Total	**38**	**62**	**100**

Journal B

Value	Citations: Low	Citations: High	**Total**
Low	56	24	**80**
High	6	14	**20**
Total	**62**	**38**	**100**

Select papers: how many have a high value?

1. Select highly cited publications:

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total	**62**	**38**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{56+14}{62+38} = \frac{70}{100} = 0.7$.

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations		
		Low	High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations		
		Low	High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total	**62**	**38**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{56+14}{62+38} = \frac{70}{100} = 0.7$.
2. Select publications from high impact journal:

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total	**62**	**38**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{56+14}{62+38} = \frac{70}{100} = 0.7$.
2. Select publications from high impact journal: $\frac{24+56}{100} = 0.8$.

CWTS Meaningful metrics

Scenario 2 (Citations noisy) example

Journal A

		Citations: Low	Citations: High	**Total**
Value	Low	14	6	**20**
	High	24	56	**80**
	Total	**38**	**62**	**100**

Journal B

		Citations: Low	Citations: High	**Total**
Value	Low	56	24	**80**
	High	6	14	**20**
	Total	**62**	**38**	**100**

Select papers: how many have a high value?

1. Select highly cited publications: $\frac{56+14}{62+38} = \frac{70}{100} = 0.7$.
2. Select publications from high impact journal: $\frac{24+56}{100} = 0.8$.

Journal more accurate.

CWTS Meaningful metrics

Simulation

Procedure

1. Create articles with some value .

Simulation

Procedure

1. Create articles with some value .
2. Articles are peer reviewed with some error.

CWTS
Meaningful metrics

Simulation

Procedure

1. Create articles with some value .
2. Articles are peer reviewed with some error.

3. Articles are selected for a journal based on peer review.

CWTS
Meaningful metrics

Simulation

Procedure

1. Create articles with some value .
2. Articles are peer reviewed with some error.
3. Articles are selected for a journal based on peer review.
4. Articles attract citations according to their value.

CWTS
Meaningful metrics

Simulation

Procedure

1. Create articles with some value $v_i \sim \text{LogN}(\sigma_v^2)$.
2. Articles are peer reviewed with some error.
3. Articles are selected for a journal based on peer review.
4. Articles attract citations according to their value.

CWTS
Meaningful metrics

Procedure

1. Create articles with some value $v_i \sim \text{LogN}(\sigma_v^2)$.
2. Articles are peer reviewed with some error.
   - Estimated value $e_{ik} = v_i \epsilon_{ik}$.
   - Error is $\epsilon_{ik} \sim \text{LogN}(\sigma_r^2)$.
3. Articles are selected for a journal based on peer review.
4. Articles attract citations according to their value.

Procedure

1. Create articles with some value $v_i \sim \text{LogN}(\sigma_v^2)$.
2. Articles are peer reviewed with some error.
   - Estimated value $e_{ik} = v_i \epsilon_{ik}$.
   - Error is $\epsilon_{ik} \sim \text{LogN}(\sigma_r^2)$.
3. Articles are selected for a journal based on peer review.
   - All articles submitted to highest journal.
   - Only top x articles accepted, others rejected.
   - If rejected, submit to next highest journal, and so on ....
4. Articles attract citations according to their value.

Procedure

1. Create articles with some value $v_i \sim \text{LogN}(\sigma_v^2)$.
2. Articles are peer reviewed with some error.
   - Estimated value $e_{ik} = v_i \epsilon_{ik}$.
   - Error is $\epsilon_{ik} \sim \text{LogN}(\sigma_r^2)$.
3. Articles are selected for a journal based on peer review.
   - All articles submitted to highest journal.
   - Only top x articles accepted, others rejected.
   - If rejected, submit to next highest journal, and so on ....
4. Articles attract citations according to their value.
   - Citations $c_i = v_i \epsilon_i$ where $\epsilon_i \sim \text{LogN}(\sigma_c^2)$.
   - Citations distributed as $c_i \sim \text{LogN}(\sigma_v^2 + \sigma_c^2)$.
   - Consistent with empirical observation that $c_i \sim \text{LogN}(1.3)$.

Simulation

Procedure

1. Create articles with some value $v_i \sim \text{LogN}(\sigma_v^2)$.
2. Articles are peer reviewed with some error.
   - Estimated value $e_{ik} = v_i \epsilon_{ik}$.
   - Error is $\epsilon_{ik} \sim \text{LogN}(\sigma_r^2)$.
3. Articles are selected for a journal based on peer review.
   - All articles submitted to highest journal.
   - Only top x articles accepted, others rejected.
   - If rejected, submit to next highest journal, and so on ....
4. Articles attract citations according to their value.
   - Citations $c_i = v_i \epsilon_i$ where $\epsilon_i \sim \text{LogN}(\sigma_c^2)$.
   - Citations distributed as $c_i \sim \text{LogN}(\sigma_v^2 + \sigma_c^2)$.
   - Consistent with empirical observation that $c_i \sim \text{LogN}(1.3)$.

IF is average of all citations to articles published in journal.

CWTS
Meaningful metrics

Article selection

Two possible ways to try to select high value (top 10%) articles.

1. Select top 10% most highly cited articles.
2. Select top 10% highest impact factor articles.

CWTS
Meaningful metrics

Article selection

Two possible ways to try to select high value (top 10%) articles.

1. Select top 10% most highly cited articles.
2. Select top 10% highest impact factor articles.

Accuracy

Proportion of articles that fall in the top 10% highest value.

CWTS
Meaningful metrics

Article selection

Two possible ways to try to select high value (top 10%) articles.

1. Select top 10% most highly cited articles.
2. Select top 10% highest impact factor articles.

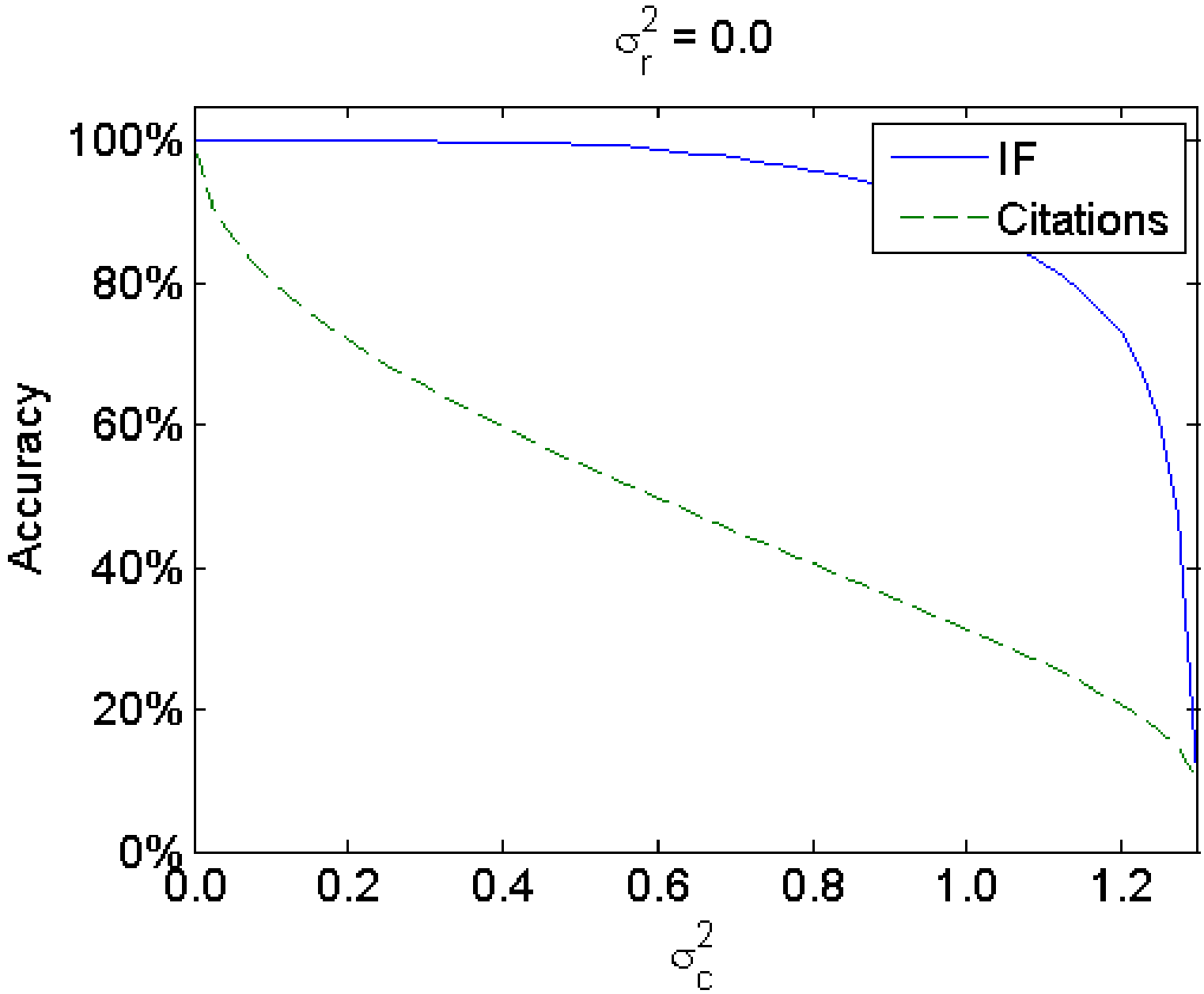
Accuracy

Proportion of articles that fall in the top 10% highest value.

Parameters: 20 journals, 100 articles/journal, 1000 simulations.

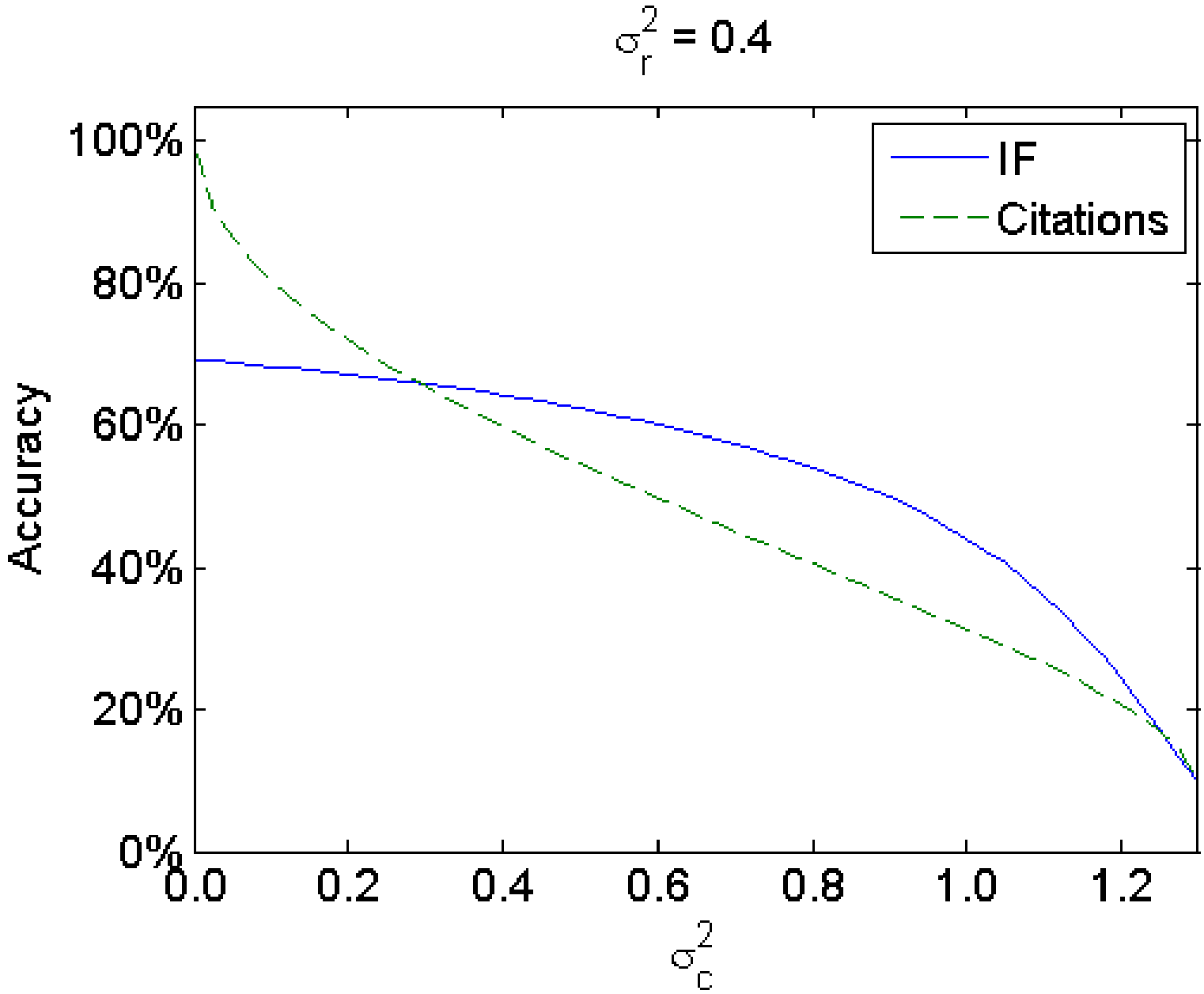
CWTS
Meaningful metrics

Results



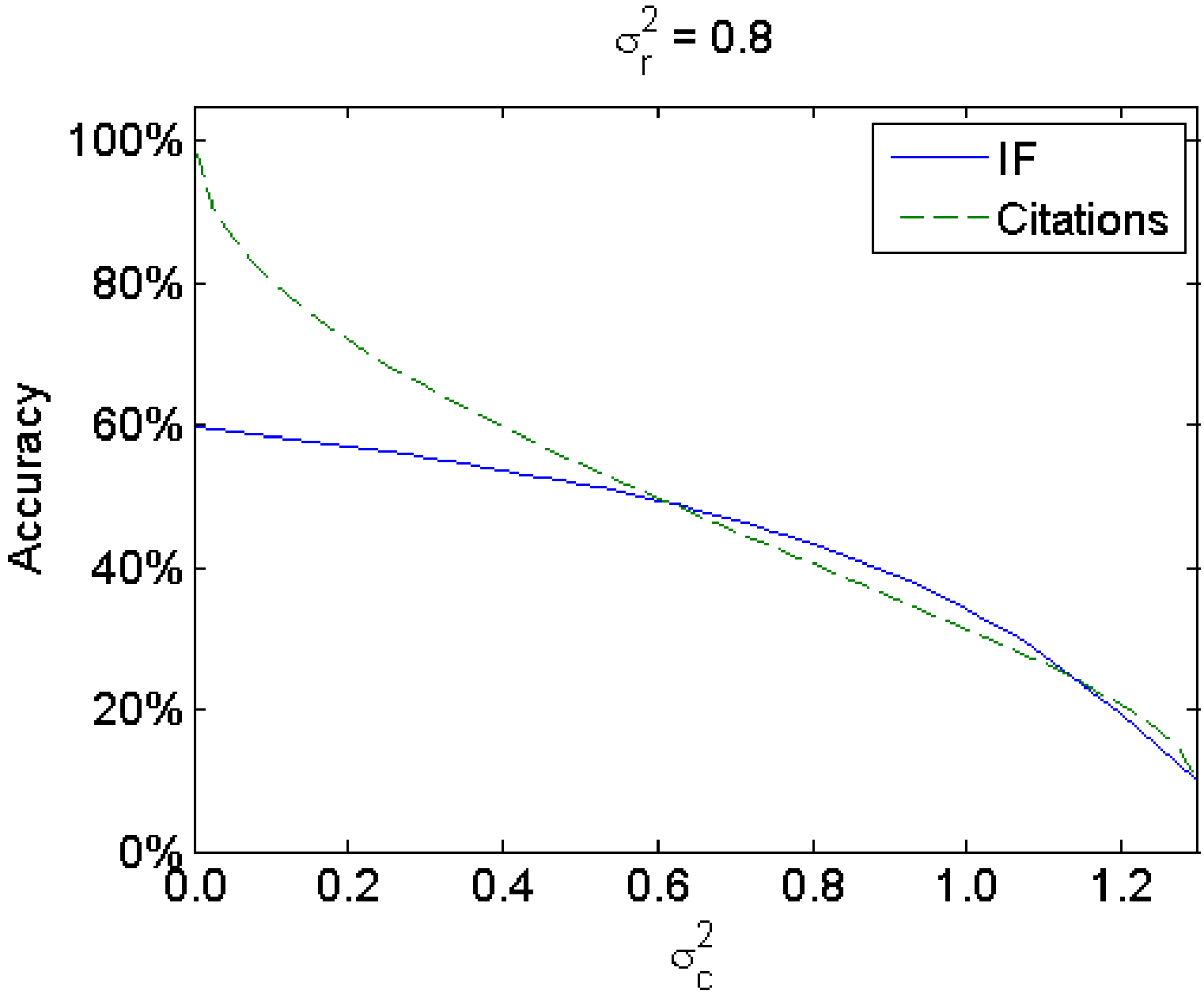
CWTS
Meaningful metrics

Results



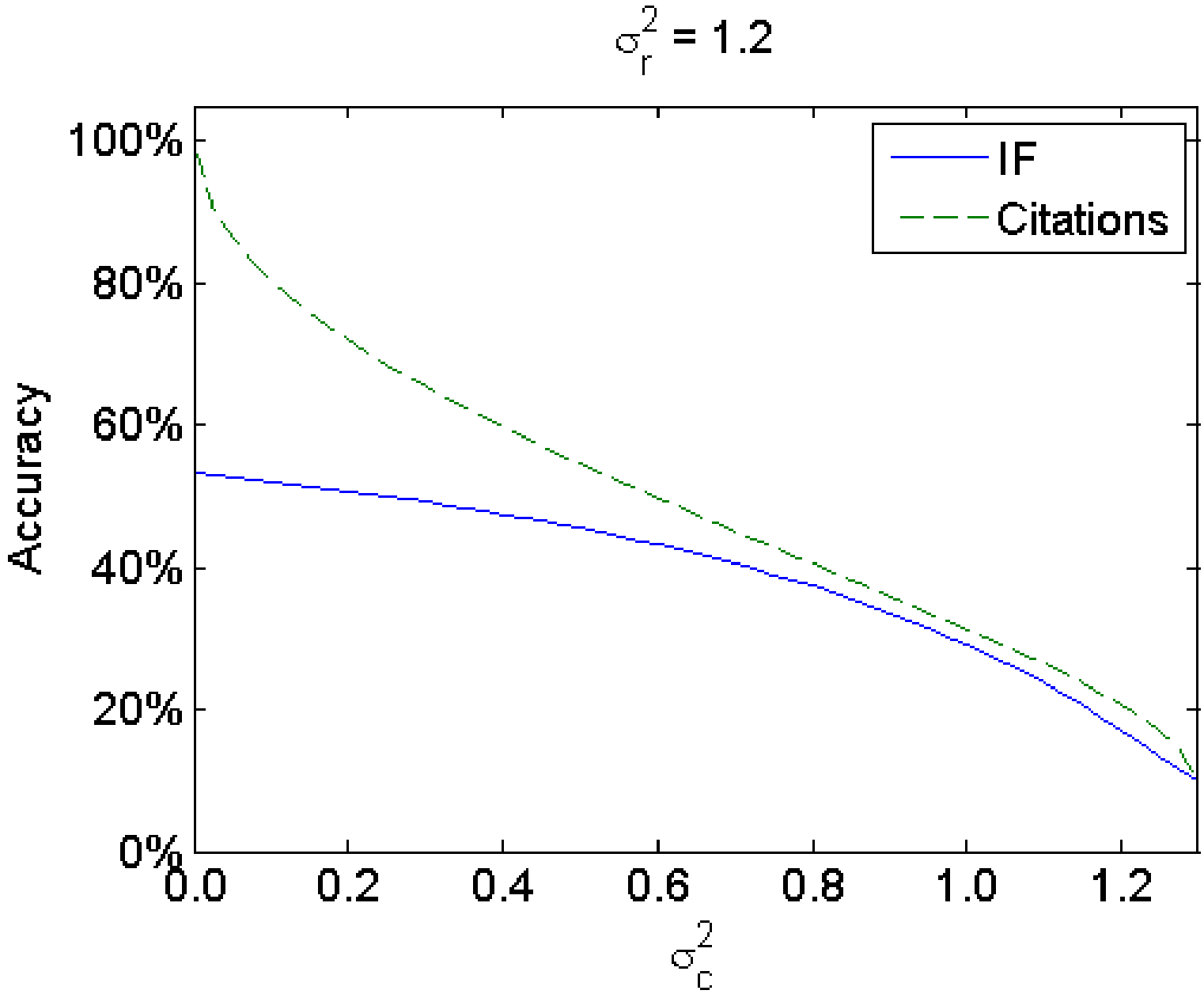
CWTS
Meaningful metrics

Results



CWTS
Meaningful metrics

Results



CWTS
Meaningful metrics

Conclusions

- Skewness alone is no argument for rejecting the IF.
- If citations are reasonably accurate while peer review is noisy, then IF may be less accurate than citations.
- If citations are noisy while peer review is reasonably accurate, then IF may be more accurate than citations.
- In practice, parameters used in this model unknown, so requires further (empirical) analysis.
- There may be good non-statistical arguments against the IF (e.g. goal displacement).

Conclusions

- Skewness alone is no argument for rejecting the IF.
- If citations are reasonably accurate while peer review is noisy, then IF may be less accurate than citations.
- If citations are noisy while peer review is reasonably accurate, then IF may be more accurate than citations.
- In practice, parameters used in this model unknown, so requires further (empirical) analysis.
- There may be good non-statistical arguments against the IF (e.g. goal displacement).

Statistics is hard.

CWTS
Meaningful metrics

Conclusions

- Skewness alone is no argument for rejecting the IF.
- If citations are reasonably accurate while peer review is noisy, then IF may be less accurate than citations.
- If citations are noisy while peer review is reasonably accurate, then IF may be more accurate than citations.
- In practice, parameters used in this model unknown, so requires further (empirical) analysis.
- There may be good non-statistical arguments against the IF (e.g. goal displacement).

Statistics is hard. Make assumptions and arguments explicit.

CWTS
Meaningful metrics

Thank you!

Questions?

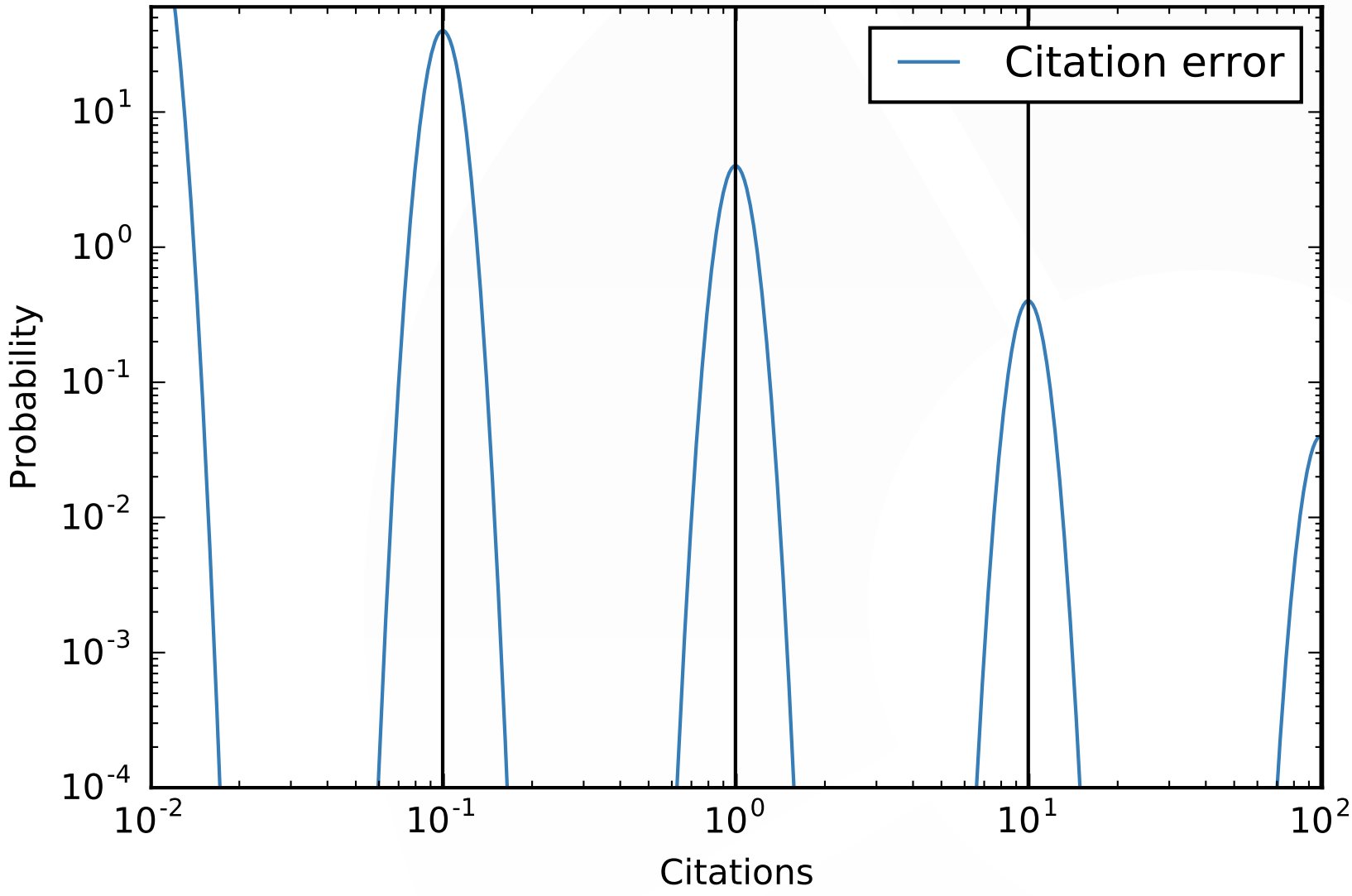
arχiv:1703.02334

✉ v.a.traag@cwts.leidenuniv.nl

@vtraag

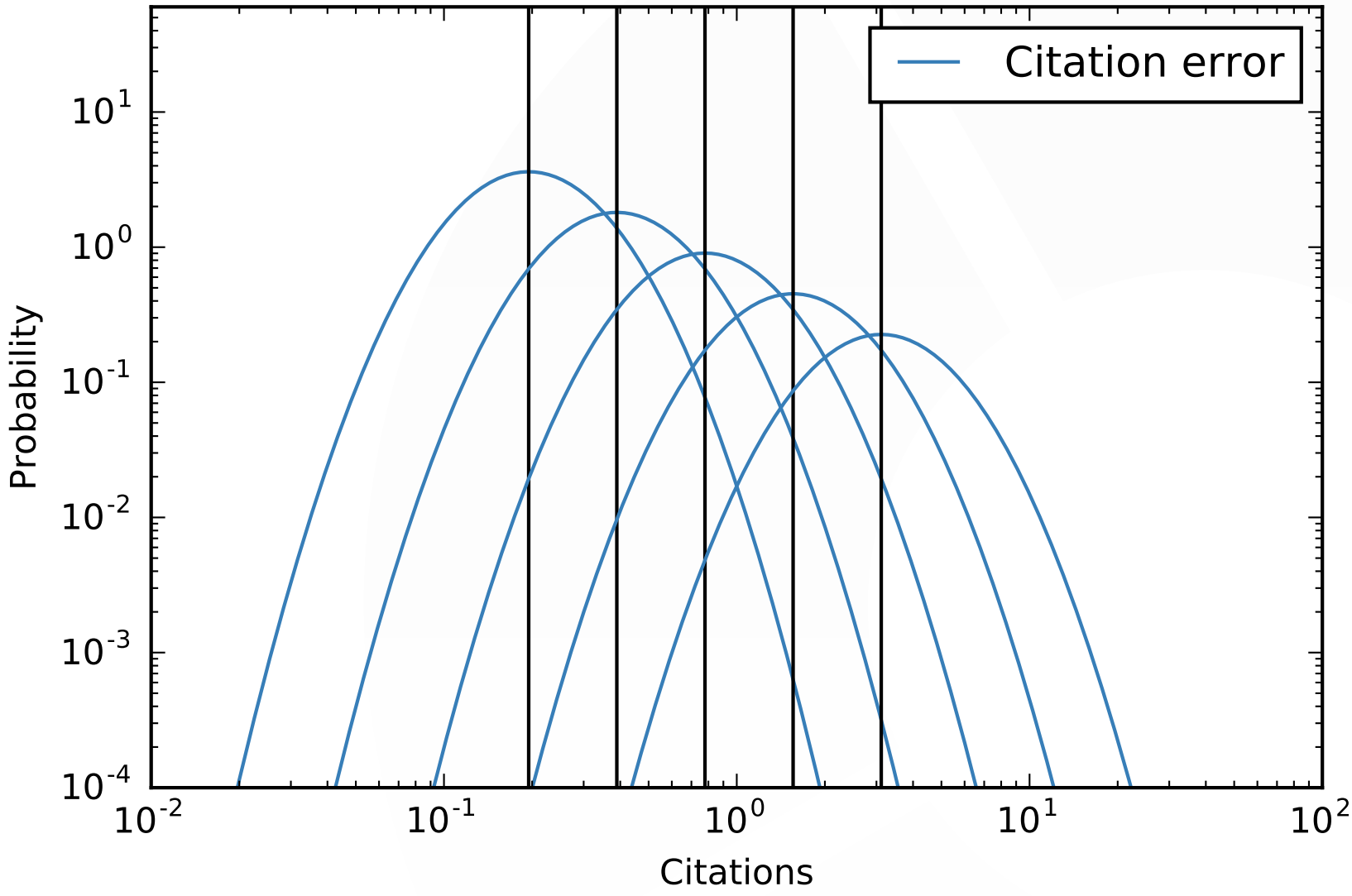
CWTS
Meaningful metrics

Scenario 1 (Citations accurate) illustration



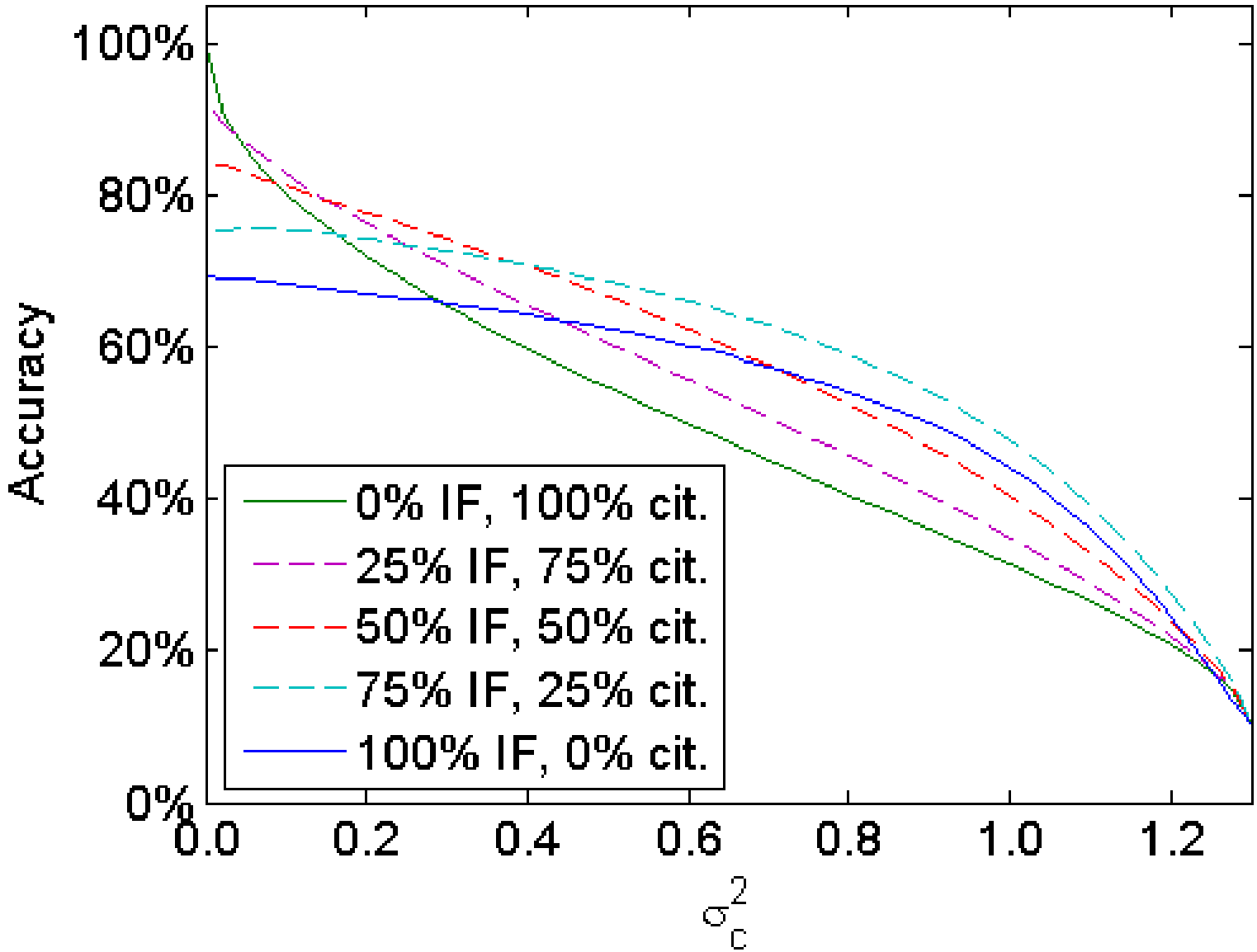
CWTS
Meaningful metrics

Scenario 2 (Citations noisy) illustration



CWTS
Meaningful metrics

Results hybrid case



CWTS
Meaningful metrics