The Economics of Citation

Jeong-Yoo Kim (Kyung Hee)
Insik Min (Kyung Hee)
Christian Zimmermann (UConn)

UConn-Wesleyan 2007 Workshop

The Questions

What strategies do authors use when choosing whom to cite?

The Questions

What strategies do authors use when choosing whom to cite?

- Become more convincing by citing respected authors

The Questions

What strategies do authors use when choosing whom to cite?

- Become more convincing by citing respected authors
- Cater to editors and potential referees

The Questions

What strategies do authors use when choosing whom to cite?

- Become more convincing by citing respected authors
- Cater to editors and potential referees
- Look more competent that cited authors

The Effects

The Effects

Correlation effect

The Effects

Correlation effect

Reputation effect

The Effects

Correlation effect
Cite authors ranked higher
Reputation effect

The Effects

Correlation effect

Cite authors ranked higher

Reputation effect

Cite authors ranked lower

The Environment

An author makes claim ω_1

The Environment

An author makes claim ω_1, which is true with probability μ_1

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability).

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2, if it raises the posterior probability that ω_1 is true

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2, if it raises the posterior probability that ω_1 is true:
$\alpha_T \equiv P(\omega_1 = T|\omega_2 = T)$, $\alpha_F \equiv P(\omega_1 = F|\omega_2 = F)$

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2, if it raises the posterior probability that ω_1 is true:
$\alpha_T \equiv P(\omega_1 = T | \omega_2 = T)$, $\alpha_F \equiv P(\omega_1 = F | \omega_2 = F)$, $\alpha_T, \alpha_F > \frac{1}{2}$

The Environment

An author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2, if it raises the posterior probability that ω_1 is true: $\alpha_T \equiv P(\omega_1 = T|\omega_2 = T)$, $\alpha_F \equiv P(\omega_1 = F|\omega_2 = F)$, $\alpha_T, \alpha_F > \frac{1}{2}$, common knowledge.

The Environment

A risk-neutral author makes claim ω_1, which is true with probability μ_1 (ability). Author cites ω_2, true with probability μ_2, if it raises the posterior probability that ω_1 is true: $\alpha_T \equiv P(\omega_1 = T | \omega_2 = T)$, $\alpha_F \equiv P(\omega_1 = F | \omega_2 = F)$, $\alpha_T, \alpha_F > \frac{1}{2}$, common knowledge.

Complete Information

μ_1 and μ_2 are common knowledge.

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T \mu_2 + (1 - \alpha_F)(1 - \mu_2)$,

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T \mu_2 + (1 - \alpha_F)(1 - \mu_2)$,
$P(\omega_1 = T) = \mu_1$.

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T \mu_2 + (1 - \alpha_F)(1 - \mu_2)$,
$P(\omega_1 = T) = \mu_1$.
Cite if $\mu_1 < \bar{\mu}_1 \equiv \alpha_T \mu_2 + (1 - \alpha_F)(1 - \mu_2)$

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2)$,
$P(\omega_1 = T) = \mu_1$.
Cite if $\mu_1 < \bar{\mu}_1 \equiv \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2)$
or $\mu_2 > \bar{\mu}_2 \equiv \frac{\mu_1+\alpha_F-1}{\alpha_T+\alpha_F-1}$.

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T\mu_2 + (1-\alpha_F)(1-\mu_2)$,
$P(\omega_1 = T) = \mu_1$.
Cite if $\mu_1 < \bar{\mu}_1 \equiv \alpha_T\mu_2 + (1-\alpha_F)(1-\mu_2)$
or $\mu_2 > \bar{\mu}_2 \equiv \frac{\mu_1+\alpha_F-1}{\alpha_T+\alpha_F-1}$.
Thus, less capable author can improve belief on his claim by citing more capable author, and vice-versa

Complete Information

μ_1 and μ_2 are common knowledge. Then,
$P(\omega_1 = T|\omega_2) = \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2)$,
$P(\omega_1 = T) = \mu_1$.
Cite if $\mu_1 < \bar{\mu}_1 \equiv \alpha_T \mu_2 + (1-\alpha_F)(1-\mu_2)$
or $\mu_2 > \bar{\mu}_2 \equiv \frac{\mu_1+\alpha_F-1}{\alpha_T+\alpha_F-1}$.
Thus, less capable author can improve belief on his claim by citing more capable author, and vice-versa:
Correlation effect of citation.

Incomplete Information

μ_1 known to proportion λ.

Incomplete Information

μ_1 known to proportion λ. $1 - \lambda$ know $G(\mu_1)$.

Incomplete Information

μ_1 known to proportion λ. $1 - \lambda$ know $G(\mu_1)$. μ_2 common knowledge.

Incomplete Information

μ_1 known to proportion λ. $1 - \lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1.

Incomplete Information

μ_1 known to proportion λ. $1-\lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Incomplete Information

μ_1 known to proportion λ. $1-\lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Weak Perfect Bayesian Equilibrium:

Incomplete Information

μ_1 known to proportion λ. $1 - \lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Weak Perfect Bayesian Equilibrium:

(i) Author 1 cites ω_2 if $\mu_1 \leq \tilde{\mu}_1(\lambda)$

Incomplete Information

μ_1 known to proportion λ. $1-\lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Weak Perfect Bayesian Equilibrium:

(i) Author 1 cites ω_2 if $\mu_1 \leq \tilde{\mu}_1(\lambda)$

(ii) $\tilde{\mu}_1(\lambda) < \bar{\mu}_1$

Incomplete Information

μ_1 known to proportion λ. $1-\lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Weak Perfect Bayesian Equilibrium:

(i) Author 1 cites ω_2 if $\mu_1 \leq \tilde{\mu}_1(\lambda)$

(ii) $\tilde{\mu}_1(\lambda) < \bar{\mu}_1$

(iii) $\tilde{\mu}_1(\lambda)$ is increasing in λ

Incomplete Information

μ_1 known to proportion λ. $1-\lambda$ know $G(\mu_1)$. μ_2 common knowledge. Citation decision reveals something about μ_1. Let $\tilde{\mu}_1$ be indifferent between citing or not.

Weak Perfect Bayesian Equilibrium:

(i) Author 1 cites ω_2 if $\mu_1 \leq \tilde{\mu}_1(\lambda)$

(ii) $\tilde{\mu}_1(\lambda) < \bar{\mu}_1$

(iii) $\tilde{\mu}_1(\lambda)$ is increasing in λ :

Reputation effect of citation.

Data: RePEc

- RePEc Author Service: 12,205 authors

Data: RePEc

- RePEc Author Service: 12,205 authors
- RePEc listing: 450,000 works

Data: RePEc

- RePEc Author Service: 12,205 authors
- RePEc listing: 450,000 works
- Citation analysis: Impact factors

Data: RePEc

- RePEc Author Service: 12,205 authors
- RePEc listing: 450,000 works
- Citation analysis: Impact factors
- RePEc author rankings

Data used

- Authors with references

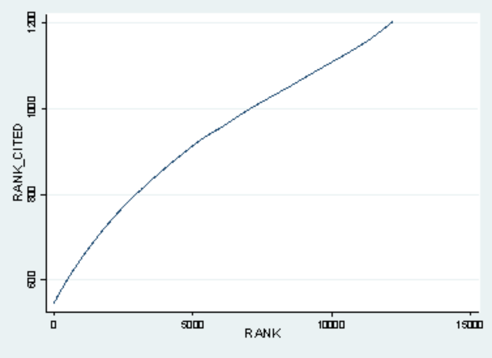
Data used

- Authors with references
- References with ranked authors

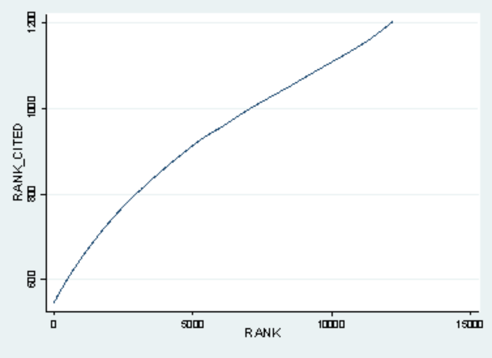
Data used

- Authors with references
- References with ranked authors
- 12,205 $\longrightarrow$ 9,127

Correlation effect

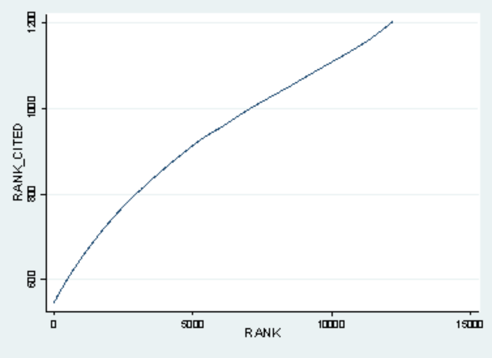


Correlation effect



$$RANK\_CITED = \beta_0 + \ \beta_1 \qquad RANK + e$$

Correlation effect



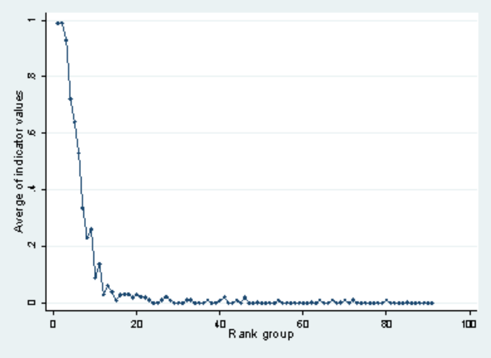
$$RANK\_CITED = \beta_0 + \underset{(0.002)}{0.05} \; RANK + e$$

Reputation effect I

$$P(RANK\_CITED > RANK)$$

Reputation effect I

$$P(RANK\_CITED > RANK)$$



Reputation effect II

$$RANK\_CITED = \beta_0 + \beta_1 RANK + \beta_2 RANK^2 + e$$

Reputation effect II

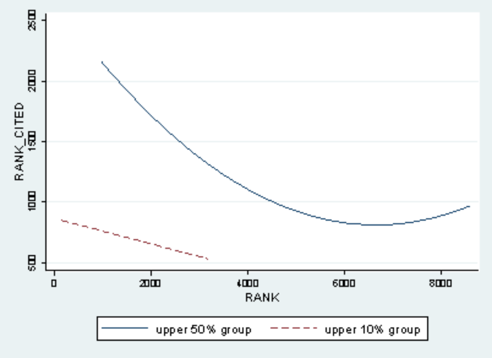
$$RANK\_CITED = \beta_0 + \beta_1 RANK + \beta_2 RANK^2 + e$$

Sort by $RANK\_NW$, top 10%, 45%–55%

Reputation effect II

$$RANK\_CITED = \beta_0 + \beta_1 RANK + \beta_2 RANK^2 + e$$

Sort by $RANK\_NW$, top 10%, 45%–55%



Regional differences

- North-Americans are higher ranked

Regional differences

- North-Americans are higher ranked
- $AVE\_CITING = \beta_0 + \beta_1 RANK\_NW + \beta_2 EUROPE + \beta_3 OTHERS + e$

Regional differences

- North-Americans are higher ranked
- $AVE\_CITING = \beta_0 + \beta_1 RANK\_NW + \beta_2 EUROPE + \beta_3 OTHERS + e$
- Tobit results:

Regional differences

- North-Americans are higher ranked
- $AVE\_CITING =$
  $\beta_0 + \beta_1 RANK\_NW + \beta_2 EUROPE + \beta_3 OTHERS + e$
- Tobit results:

Var.	Coeff	St. Err.	T
RANK_NW	-0.00066	0.00001	-60.76
EUROPE	-0.438	0.0648	-6.76
OTHERS	-0.452	0.105	-4.29

Conclusions

1. Correlation effect: cite better authors

Conclusions

1. Correlation effect: cite better authors
2. Reputation effect: cite more selectively when uncertainty about own competence

Conclusions

1. Correlation effect: cite better authors
2. Reputation effect: cite more selectively when uncertainty about own competence
3. North-American bias