

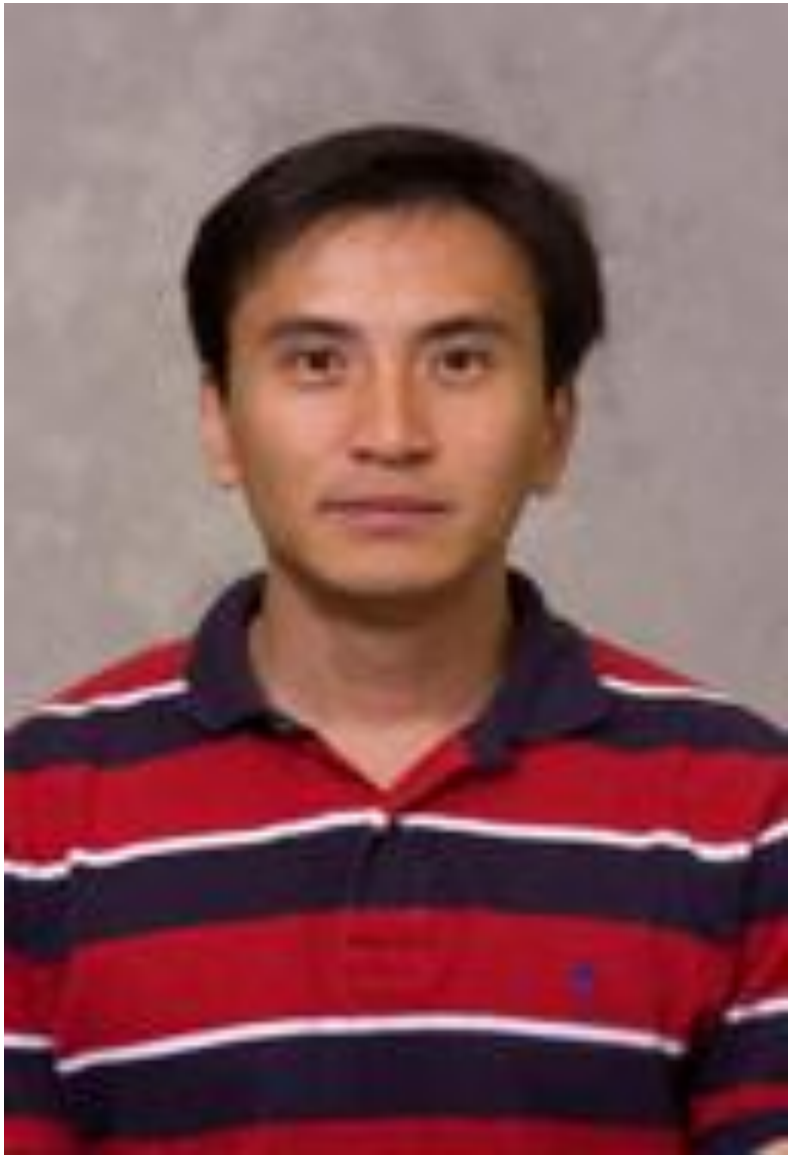
# Learning from Peers on Social Media Platforms: A Dynamic Structural Analysis

Yingda Lu, Param Vir Singh, Baohong Sun
Tepper School of Business
Carnegie Mellon University

MARCH 4, 2011

*Presented at Krannert School of Business, Purdue University*

# Thank You!!



# Web 2.0 and Knowledge Sharing

- An increasing focus on social software applications and services to
  - Break the silos
  - Encourage knowledge sharing across department and locations
- 90% are building the culture
  - IBM, CISCO, Infosys, Dell, Sun, Oracle
  - Knowledge sharing, idea generation
  - Customer service
    - Customer help each other
    - Significant savings of service costs
- Many discussions about how Software/Media revolutionizing the working world of the future

# Discussion Forum example

May 15, 2010 10:03 AM

## Photoshop CS5 Trial Runtime Error

DannyQ
7 posts since
May 15, 2010

I installed the Photoshop CS5 trial without any problems, however when i attempt to load the application a box appears stating:

*Runtime Error. This application has requested the Runtime to teminate it it an unusual way.*

I have not been able to load photoshop a single time.

I'm running Windows Vista 64bit, i'm not sure if this makes a difference or not.

Please help me to trouble shoot this problem! I have searched the net for answers without any luck.
Thanks in Advance to anyone whom can help me.

---

1. May 15, 2010 10:30 AM In response to: DannyQ

**Re: Photoshop CS5 Trial Runtime Error**

COMMUNITY PROFESSIONAL
Mylenium
10,430 posts since
Oct 26, 2006

Please provide the exact specs of your system.

Mylenium

Report Abuse

---

4. May 15, 2010 6:37 PM In response to: DannyQ

**Re: Photoshop CS5 Trial Runtime Error**

dec9
4,057 posts since
Aug 20, 2007

You can open and look at the Vista Event Viewer. There will be a hit listed in there somewhere with more detail. Just have to dig for it.

You can also run sfc /scannow at a elevated command promt. This will scan for Vista system errors and fix them if detected. Have your Vista DVD ready incase it asks you to put the DVD in for repair.

Also try this first. It detects any problems with your Vista computer and fixes it. Not sure if it will work on a runtime error but worth a try:

http://fixitcenter.support.microsoft.com/Portal

Report Abuse

---

5. May 15, 2010 10:48 PM In response to: dec9

**Re: Photoshop CS5 Trial Runtime Error**

Hudechrome-sd9sPI
3,383 posts since
Aug 11, 2002

Appears to be some sort of beta and it's closed for d/l.

Report Abuse

# Industry Background

- Fortune 500 firm
- Discussion forum adopted by a firm to support peer learning among customer-facing staff.
- Anybody can register as a user
- True identities of both knowledge sharer and seekers are revealed
  - Act as quality control in the absence of feedback
- Most of the questions are technical questions
  - "How to insert and retrieve multilingual data using ORACLE NCLOB?"
- Multiple answers are sequenced according to time stamps
- Top management use the forum to identify experts. Active contributors have higher probabilities of receiving bonus or promotion
- We focus on most basic features

# Learning Mechanism Enabled by the Platform

- Seeking knowledge
  - Ask question
  - Everybody can answer
  - Knowledge seeker gains knowledge
  - Whole community gains knowledge
- Sharing knowledge
  - Answer a question
  - Knowledge seeker gains knowledge
  - Whole community gains knowledge

Platform enables very different learning mechanism

1. Learning from peers: learning cannot be achieved without contribution from peers
2. Externality of learning: any contribution improves knowledge level of the whole community

Naturally, these properties imply decision process that is

1. Inter-dependent
2. Forward-looking
3. Future reward reciprocated by peers

"Reciprocal Altruism": an act of helping others at a cost is beneficial if there is a chance of being in a reverse situation where the person whom I helped before may perform an altruistic act towards me.

# Early Adopters Observe Two Challenges

## First Challenge: Slow adoption & Flattening of Adoption Rate Quickly after Take-off

**Figure 2A. The Total Number of Answers and Questions[a]**

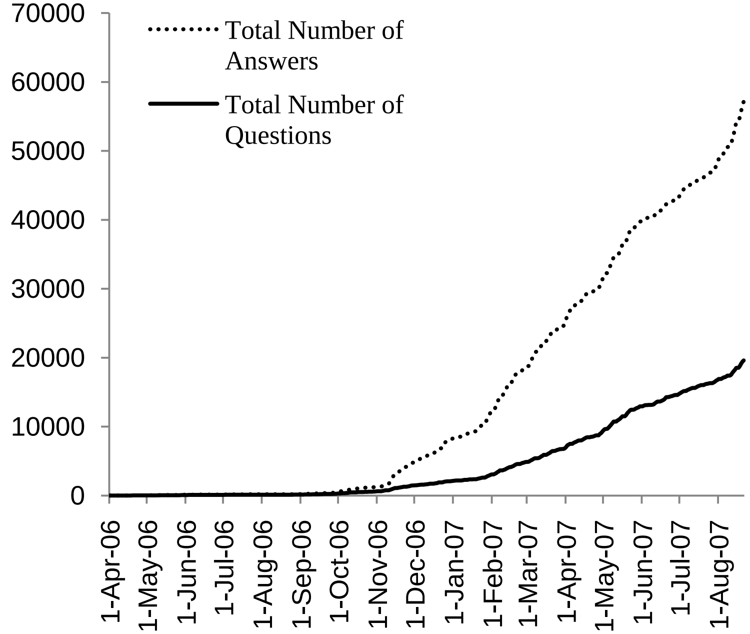
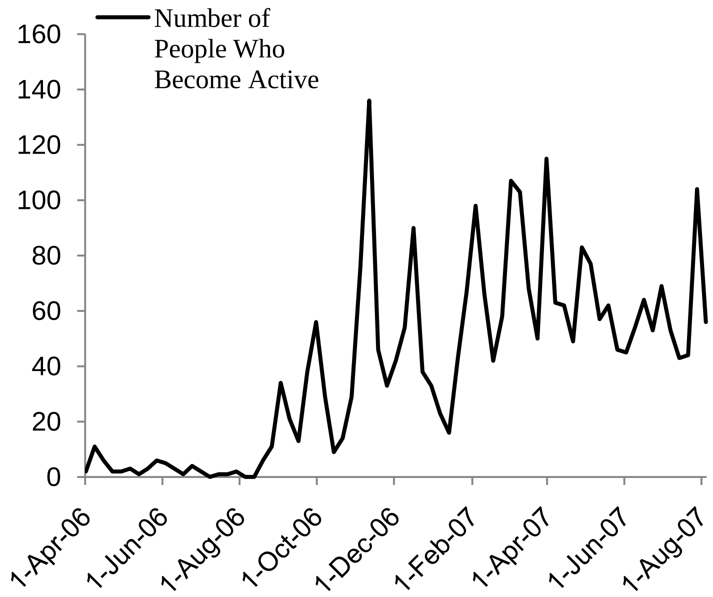
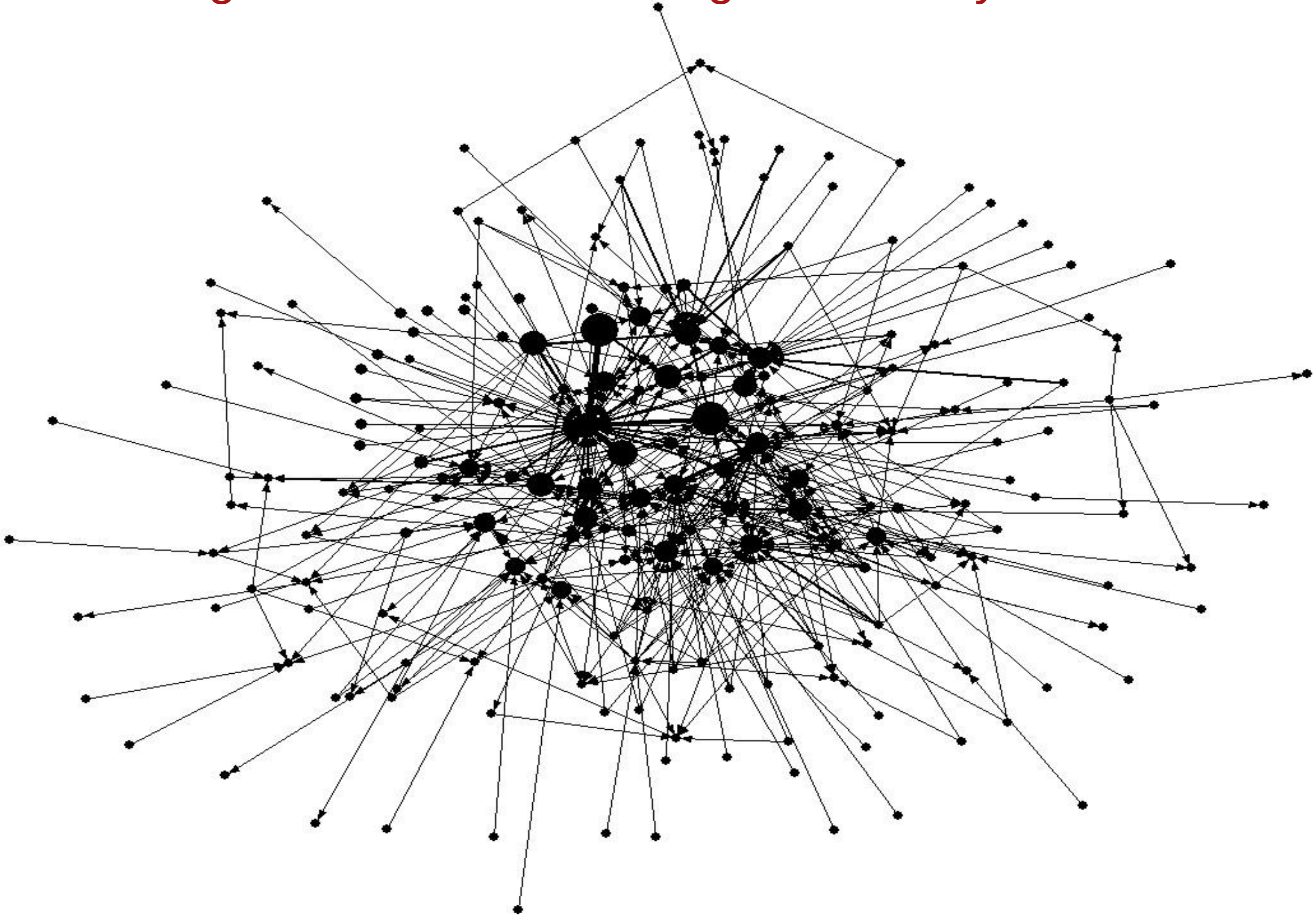



**Figure 2B. The Number of People Who Participate in the Forum for the First Time in Each Period[b]**



# Early Adopters Observe Two Challenges

Second Challenge: Most activities are generated by a few users



**Core-Periphery Network Structure**

# Research Questions

Carnegie Mellon Tepper SCHOOL OF BUSINESS

- What drives users' decisions of knowledge seeking and sharing on the social platform?
- How does an individual incorporate others' decisions of asking and answering questions into her own decisions process?
- How user incentives lead to the emergence of core/periphery social network structures among participants of web 2.0 communities?
- Is current design of the system aligned with user decision dynamics? If not, is there a way to alleviate this misalignment?

# Related Literature

- **Customer service**
  - Jones and Sasser 1995, Reinchheld 1996, Reinartz and Kumar 2000, Kamakura et al 2002, Payne and Frow 2005, Moe and Fader 2001, Anderson 2002
- **Consumer behavior in online communities**
  - Katona and Sarvary 2008, Mayzlin and Yoganarasimhan 2008, Narayan and Yang 2007, Godes and Mayzlin 2004, Chevalier and Mayzlin 2006, 2009, Weiss et a  2008
- **Network formation & Network decision making**
  - Narayan and Yang 2007, Stephen and Toubia 2009a, 2009b, Lu et al 2010, Singh et al 2010, Singh 2011,  Aral et al 2008

  **Most assume exogenous network and adopt reduced form approach**
- **Learning from peers**
  - Knowledge gained through interaction with peers improve productivity (i.e. Argote 1999, Levitt and March 1988, Benkard 2000, Argot and Epple 1990, Hatch and Mowery 1998, Bho et al 2007)

  **Do not directly observe learning**
- **Structural learning models**
  - Learning through consumption experience, signals contained in price and advertising

    (i.e. Erdem and Keane 1996, Erdem, Keane and Sun 2008)
  - Atomistic view of individual behavior
- **Dynamic structural game models**
  - Bajari et al 2007, 2008

# Data Description

## Data Description

| | .Net Framework |
|---|---|
| Number of question per day | 652 |
| Number of answer per day | 1676 |
| Number of participants | 329 |
| Percentage of employees asking questions | 66.2% |
| Average number of questions asked per employee | 1.9818 |
| Frequency of asking questions | 13.04 |
| Percentage of employees answering questions | 55.93% |
| Average number of questions answered per employee | 5.09 |
| Frequency of answering questions | 33.52 |

# A Dynamic Model

- Decision variables of user *i* at time *t*,
  - Knowledge seeking

$$a_{it} = \begin{cases} 2\,, & \text{if individual } i \text{ asks a difficult question at time } t, \\ 1\,, & \text{if individual } i \text{ asks an easy question at time } t, \\ 0\,, & \text{otherwise.} \end{cases}$$

  - Knowledge sharing

$$s_{ijt} = \begin{cases} 1 & \text{if } i \text{ answers a question from } j \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

  - We take into account question types [easy vs. difficult].
  - We allow user decisions to be dyadic.

# User Utility Function

- Per period utility

$$U_{it}(\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{X_i}, a_{it}, \boldsymbol{s_{it}}, \boldsymbol{\varepsilon_{it}}) = \alpha_1 K_{it} + \alpha_2 R_{it} - C(a_{it}, \boldsymbol{s_{it}}, K_{it}, \boldsymbol{X_i}) + \varepsilon_{it}(a_{it}, \boldsymbol{s_{it}})$$

  - $K_{it}$: the knowledge level accumulated up to time t
  - $R_{it}$: social reputation score
  - $X_i$: observed user characteristics
  - *C(.):* cost to be estimated
  - : private shock, distributed as type-I extreme value

# Knowledge State

- Knowledge $K_{it}$
  - It is about
    - Specific to a profession
    - Often related to a technical solution
    - Can saturate slowly over time
  - More knowledge increases utility
    - Solve problem
    - Job performance
    - Free time

# Knowledge State

- Knowledge updating rule

$$K_{i(t+1)} = K_{it} + \sum_{j \in N, j \neq i} k_s^w s_{jit} I\left(s_{jit} \in \{a_{it}\}\right) + \sum_{j \in N, j \neq i} k_x^w s_{jit} I\left(s_{jit} \notin \{a_{it}\}\right), where\ w \in \{O, D\}$$

  – $\sum_{j \in N, j \neq i} s_{jit} I\left(s_{jit} \in \{a_{it}\}\right)$ total number of answers provided to her question
  – $\sum_{j \in N, j \neq i} s_{jit} I\left(s_{jit} \notin \{a_{it}\}\right)$ total number of answers provided to her peers' questions (excluding herself)
  – Assumptions:
    - Same quality
    - Knowledge is additive
    - Everybody read all the answers
  – Inherent dynamics:
    - Knowledge seeker needs to predict whether her question will be answered
    - One user's decision affects knowledge of all peers and hence their future decisions
    - In the long run, she can benefit from higher community knowledge

# Social Status State

- Social Status $R_{it}$
  - It is about
    - Being perceived as an active community member
  - High status increases utility
    - Social recognition
    - Perceived as valuable in internal labor market
    - Bonus and promotion
  - It depends on
    - Frequency
    - Association
    - Relative ranking

# Social Status State

- Social Status updating rule
  - Affiliation matrix A (NXN matrix)

element $A_{ij}$ represents the number of i's questions answered by j

$$A_t = A_{1,t} + rA_{2,t}$$

- $A_1$,i: diagonal element represents number of easy questions asked by i
- $A_2$,i: diagonal element represents number of difficult questions asked by i

  - Eigenvector centrality

$$x_{ti} = A_{t,1i}x_{t1} + A_{t,2i}x_{t2} + \cdots + A_{t,ni}x_{tn}$$

$$A_t{}^T x_t = x_t$$

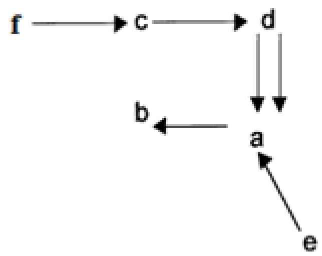
$$x_t = \alpha A_t{}^T x_t + y_t$$

$$x_t^* = \left(I - \alpha A_t^{*T}\right)^{-1}(y_t + \alpha' y_t')$$

  - Social Status score

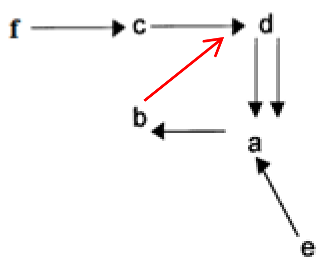
$$R_{it} = \frac{x_{ti} - 1}{\max(x_{ti}) - 1} + 1$$

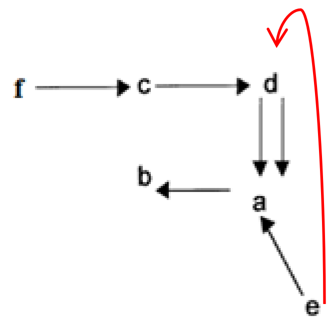
# An Example




$$A_{1t} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Affiliation matrix




Whose question should d answer: b or e?

| | **Baseline** | ***d* answers Question from *b*** | ***d* answers Question from *e*** |
|---|---|---|---|
| *a* | 2.0000 | 2.0000 | 2.0000 |
| *b* | 1.5646 | 1.5451 | 1.5525 |
| *c* | 1.2646 | 1.6537 | 1.5051 |
| *d* | 1.3439 | 1.4412 | 1.4040 |
| *e* | 1.0000 | 1.0000 | 1.0000 |
| *f* | 1.0000 | 1.0000 | 1.0000 |

Reputation score

•Social status improves more by answering question asked by a high status user than by a  low status user

# Cost Function

- Cost C(.)

$$C(a_{it}, \boldsymbol{s_{it}}) = C_a(a_{it}) + C_s(\boldsymbol{s_{it}})$$

- Cost of asking question

$$C_a(a_{it}) = \sum_{w} I(a_{it} \in w) * \left(c_{a,0}^{w} + c_{a,1}Gender + c_{a,2}POSITION\right),\ \ where\ w \in \{O, D\}$$

- Cost of answering Question

$$C_s(\boldsymbol{s_{it}}) = \sum_{j \in N} I(a_{jt} \in w) * s_{ijt} * \left(c_{s,0}^{w} + c_{s,1}Gender + c_{s,2}POSITION\right),\quad where\ w \in \{O, D\}$$

# Dynamic Problem

- User's dynamic problem:

$$\max_{(a_{i\tau},s_{i\tau})_{\tau=t}^{T}} E[\sum_{\tau=t}^{T} \gamma^{\tau-t}\, U_i(\boldsymbol{K_t}, \boldsymbol{R_t}, a_{i\tau}, s_{i\tau}, \boldsymbol{\varepsilon_{it}})|\boldsymbol{K_t}, \boldsymbol{R_t}]$$

- State variables:
  - Knowledge of self and of peers
  - Social status of self and of peers

$$V_i(\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{\sigma}; \theta) = \max_{\boldsymbol{\sigma}_i(\boldsymbol{K},\boldsymbol{R})} U_i(\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{\sigma}(\boldsymbol{K}, \boldsymbol{R})) + \gamma \int V_i(\boldsymbol{K}', \boldsymbol{R}'|\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{\sigma}_{-i}) dP(\boldsymbol{K}', \boldsymbol{R}'|\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{\sigma}_{-i}, \boldsymbol{\sigma}_{\boldsymbol{i}})$$

# Timeline of Decisions

- Everyone observes their own states as well as the states of everyone else in the community.
- Everyone receives their private shocks on the decision of asking question.
- Everyone makes predictions on their peers' decisions based on equilibrium strategy given their information of others' states in current period. Using this prediction everyone simultaneously makes decisions on whether they are going to ask question.
- Everyone observes the outcomes of asking question decisions—they know who asked questions in current period.
- Everyone receives their private shocks on the decision of answering question.
- Given information on who ask questions in current period and the predictions on others' decisions of answering questions, everyone simultaneously makes decisions on whether to provide answer for each one of the questions proposed.
- State variables: Knowledge  and Social Status are both updated.

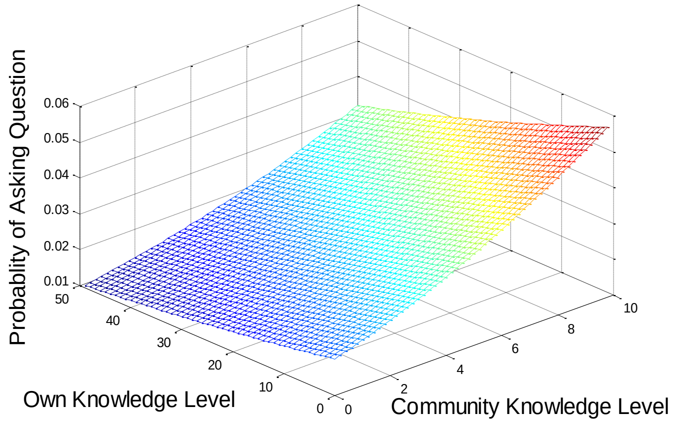
# Results for Structural Parameters

| Variable | Coefficient |
|---|---|
| **Knowledge Updating Rule** | |
| Knowledge increments from own question ($k_s^o$) | 0.5401*** |
| Knowledge increments from others' question ($k_x^o$) | 0.0036*** |
| Knowledge increments from own hard question ($k_s^D$) | 1.1703*** |
| Knowledge increments from others' hard question ($k_x^D$) | 0.0036*** |
| **Reputation Updating Rule** | |
| Reputation increment from asking hard question ($\alpha'$) | -0.0201 |
| Reputation increment from answering hard question ($r$) | 6.8783*** |
| **Utility Function Parameters** | |
| Impact from Knowledge ($\alpha_1$) | 0.2805*** |
| Impact from Social Status($\alpha_2$) | 3.6399*** |
| **Cost of asking a question** | |
| Constant for asking an easy question | 5.0030*** |
| Constant for asking a difficult question | 8.8917*** |
| Position | -0.0256* |
| Gender | -0.8052*** |
| **Cost of answering question** | |
| Constant for answering an easy question | 7.5703*** |
| Constant for answering a difficult question | 12.5224*** |
| Position | -0.1411*** |
| Gender | -0.5539*** |

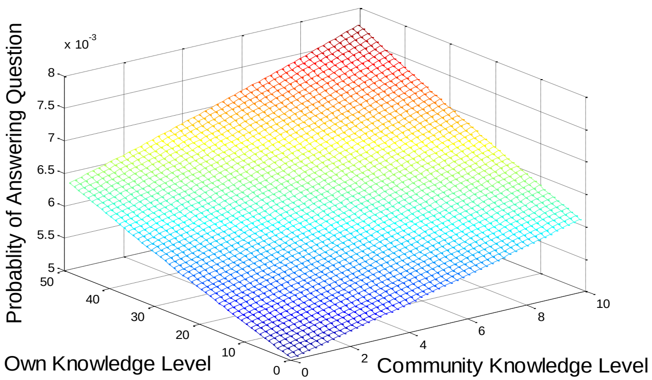
– Social Status effect is not trivial
– Cost of answering is higher than cost of asking.
– Benefit cannot compensate for the cost in the current period.

# Equilibrium Policy Functions: Reciprocal Rewards Depends on Knowledge of Peers

- The probability of asking question given her own knowledge level and average knowledge level



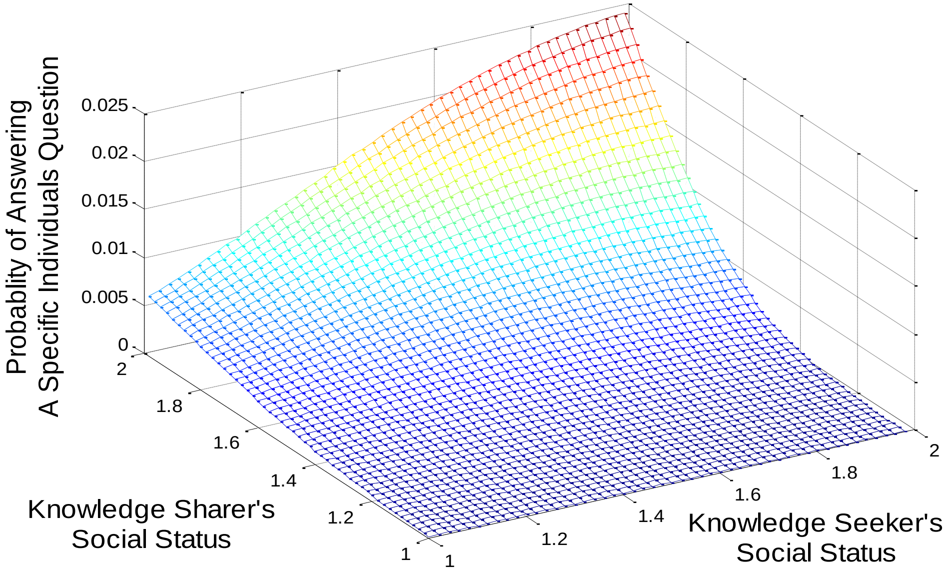

- The probability of answering question given her own knowledge level and average knowledge level




- K seeking increases with peer K
- <u>K sharing increases with peer K</u>
- <u>Decision process: dynamic and inter-dependent in anticipation of future reciprocal reward</u>

# Whose Questions to Answer?

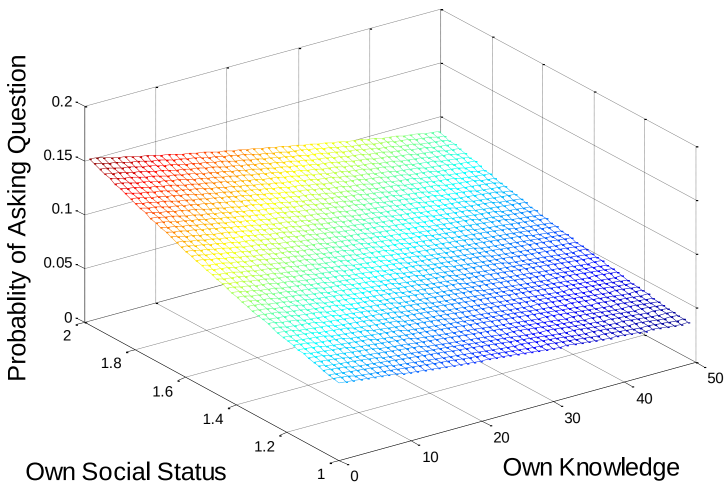
The probability of answering a question given knowledge seeker and sharer's social status



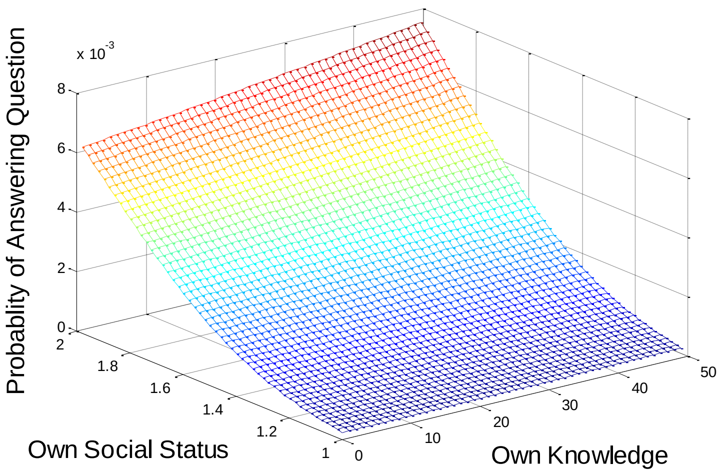

- Questions from high social status users are more likely to be answered
- High status users are more likely to share knowledge
- Implication: high status users are likely to answer each other's questions and thus forms a cohort

# Whether to Participate?

The probability of **asking a question** given knowledge and social status levels



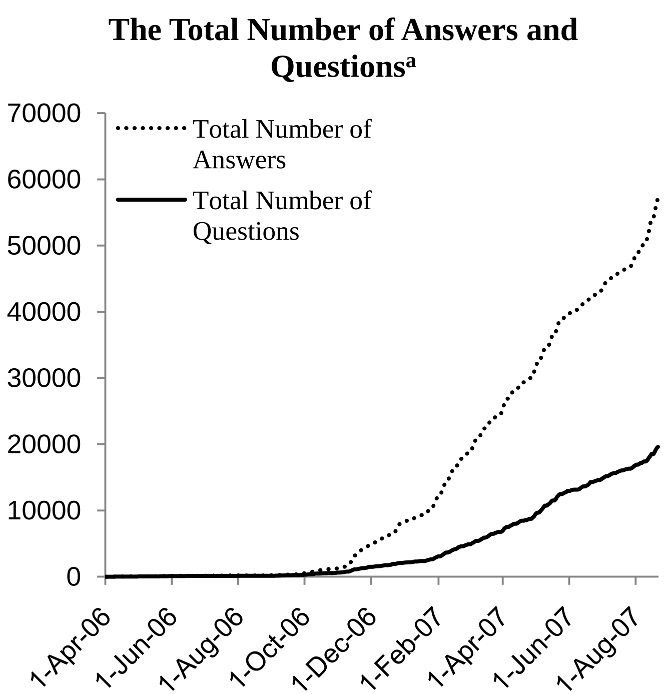

The probability of **answering a question** given knowledge and social status levels

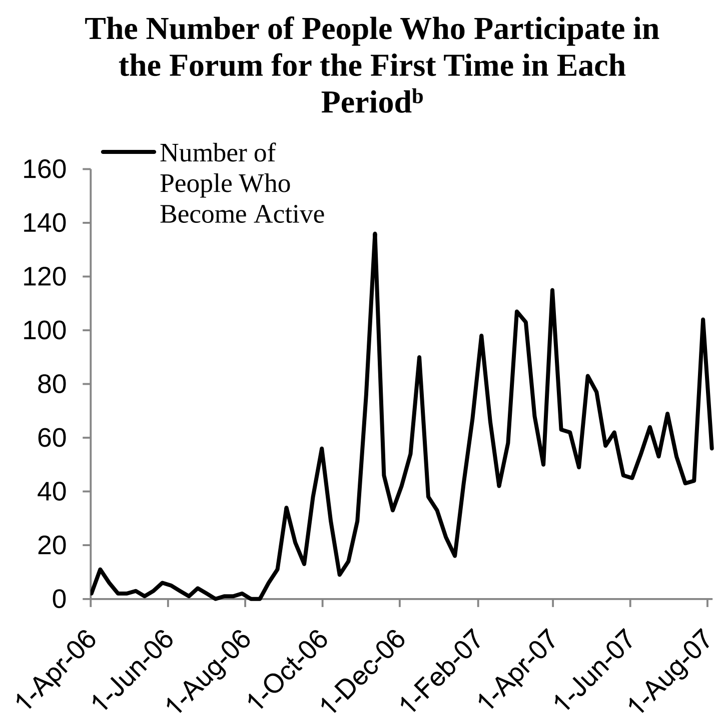



- Prob of K seeking decreases with knowledge, increases with reputation
- Prob of K sharing increases with knowledge, increases with reputation

# Explain "Free Riding" Behavior

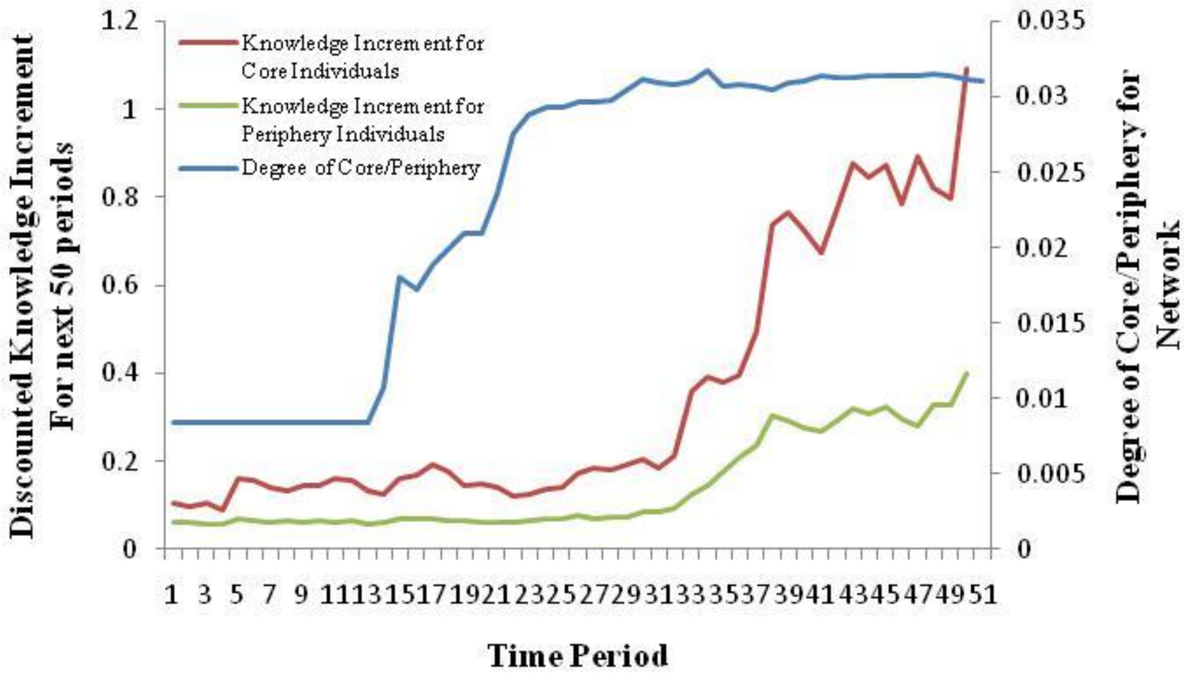
CarnegieMellon Tepper SCHOOL OF BUSINESS




- Findings:
  - Slow start: low overall knowledge level
  - Low increasing rate: formation of cohort discourage users with low status from participating
- Managerial implications:
  - Some users are "forced" not to participate
  - The adoption does not really take off

## Formation of Cohort and Speed of Knowledge Increments




Findings:

- Only centralized users benefit from the formation of cohort.

Managerial implication:

- New users do not get help
- Flow of knowledge is not from high K users to low K users
- K sharing is hurt when cohort structure is carried over to other question categories

# Reactive versus Proactive Learning

- Ripple effect
  - i ask or answer a question
  - The state variables of everybody in the whole community change
  - All peers alter their decisions about asking and answering questions
  - ……
  - The whole process continues

# Decomposition of Knowledge Increments from Answering and Asking a Question (Other state values at mean level)

Answering

| Period | % Change of Asking Questions | % Change of Answer Questions | % Knowledge Increment of User $i$ | % Knowledge Increment of Community |
|---|---|---|---|---|
| $t = 1$ | 0.0037% | 0.0313% | 0% | 0.5377% |
| $t = 2$ | 0.0035% | 0.0296% | 0.3689% | 0.0054% |
| $t = 3$ | 0.0032% | 0.0278% | 0.3584% | 0.0049% |
| $t = 4$ | 0.0031% | 0.0270% | 0.3289% | 0.0047% |
| …. | ….. | …… | …… | …... |
| Cumulative[a] | 0.0722% | 0.6025% | 7.308% | 0.6407% |

Asking

| Period | % Change of Asking Question | % Change of Answer Question | % Knowledge Increment of User $i$ | % Knowledge Increment of Community |
|---|---|---|---|---|
| $t = 1$ | 0.0019% | 0.0244% | 11.02% | 0.2579% |
| $t = 2$ | 0.0019% | 0.0235% | 0.0461% | 0.0039% |
| $t = 3$ | 0.0018% | 0.0218% | 0.0444% | 0.0037% |
| $t = 4$ | 0.0016% | 0.0210% | 0.0409% | 0.0035% |
| …. | ….. | …… | …… | ….. |
| Cumulative | 0.0351% | 0.4788% | 11.93% | 0.3347% |

Findings:

- Individual benefits more than the community
- Higher knowledge increment when asking than answering questions

Managerial implications:

- It is not about donation, it is about learning.
- Proactive learning is more effective than reactive learning

# Breaking the Cohort – A Sensitivity Study

## Alternative Design Policy

| | Probability of Asking Questions | | | Probability of Answering Questions | | | Degree of Core/ Periphery | Mean Community Knowledge | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Core[a] | Periphery | Total | Core | Periphery | Total | | Core | Periphery | Total |
| Benchmark | 0.0453 | 0.0291 | 0.0326 | 0.1955 | 0.1229 | 0.1264 | 0.0501 | 8.0472 | 3.5119 | 4.0205 |
| Anonymity | 0.0481 | 0.0556 | 0.0537 | 0.1607 | 0.1461 | 0.1518 | 0.0438 | 7.6179 | 4.4132 | 4.8429 |

Finding:

- Hiding identity of knowledge seeker can improve knowledge sharing by 36%.

Managerial implication:

- Encourage competition for reputation, but break the cohort
- Alternative design: periodically reset the record, financial incentive to encourage individuals to answer questions from users with low social status

# Summary of Major Findings

- (1) Knowledge seeking and sharing are strategic decisions driven by knowledge and network position of an individual and those of the community: users choose to seek and share knowledge for future rewards reciprocated by their peers.
- (2) Users are more likely to seek and share knowledge when their peers are more knowledgeable.
- (3) While both knowledge and reputation motivate users to share knowledge, a cohort is formed over time that has the privilege to obtain help from each other and in the meanwhile, exclude other users from participating.
- (4) "Free-riding" behavior of inactive contributors could be an equilibrium result: the earlier low community knowledge level and the later formation of cohort "force" low ranked users from participating.
- (5) Proactive learning by asking is much more effective than reactive learning by observing
- (6) Current design of the open forum is not aligned with dynamic, interrelated and inter-dependent user decision process. An alternative design that breaks the cohorts can improve the knowledge sharing by 36%.

# Conclusion

Carnegie Mellon Tepper SCHOOL OF BUSINESS

- Formally investigated knowledge seeking and sharing decisions
  - An understudied learning mechanism: learning from peers
  - Endogenize the formation of social network structure
  - Rationalize a seemingly altruism behavior: “reciprocal altruism”
  - Using observed decisions to integrate some economic, social and psychological behavior
  - Provide explanations to the two observed challenges

- Some suggestions for adaptors
  - Recognize the conspicuous nature of platform adoption
    - Knowledge Sharing Day (possibly?)
  - Build a formal reward system to recognize reputation building
    - It is a double sided sword
  - Top management should change the mind set
    - It is not a platform for donation
    - Proactive knowledge seeking behavior should be motivated
  - Be aware that silos typically formed offline also form online.
    - Modified designs that encourage competition for reputation, but break the cohort

Carnegie Mellon Tepper SCHOOL OF BUSINESS

# THANK YOU!

# Our Study …

Carnegie Mellon Tepper SCHOOL OF BUSINESS

- Explicitly model the dynamic and interdependent decision process
  - rationalize the key driving forces behind knowledge seeking and sharing decisions
  - not altruism but future reward reciprocated by peers
- Endogenize formation of a social network
  - demonstrates the formation of social network as a result of strategic interaction.
- Recognize the unique learning mechanism enabled by social media platform
  - Treats knowledge sharing as consequence of dynamic interactions of individuals.
    - learning from peers vs learning by doing
    - interdependent vs. dependent
  - Integrate social psychology, information system and marketing
- Managerially, we evaluate current design of web 2.0.

# Knowledge Updating is Different from Reputation Updating

- Reputation is directly affected by own decisions.
- Reputation ranking can go up or down
- Compete for reputation ranking.

- Thus,
  - High reputation users may not have high knowledge
  - Low reputations users may have high knowledge

# Marketing Applications of Web 2.0

Carnegie Mellon Tepper School of Business

- Customer service
  - Adobe and Oracle
  - OSI saw 22% decrease in time to solve customer support issue
- Manage business process
  - 3 months less to complete project
- Production innovation (Crowdsourcing and ideation)
  - 36% decrease in time to enact key business changes based on customer feedback
- Prediction market

- Initial values

$$(1)\ x_0 = (I - \alpha\alpha_0 A_0^T)^{-1}(y_0 + \alpha_0' y_0')$$

$$K_{i0} = k_0 + k_1 \sum_{\tau=-8}^{0} (sl_{i\tau} + sh_{i\tau})$$

# Two-Steps Estimation

- Optimality condition: equilibrium strategy is no worse than alternative perturbed strategy, while other users follow the equilibrium strategy

$$V_i(\boldsymbol{K}, \boldsymbol{R} \mid \sigma_i^*, \boldsymbol{\sigma}_{-i}^*) \geq V_i(\boldsymbol{K}, \boldsymbol{R} \mid \sigma_i', \boldsymbol{\sigma}_{-i}^*), \forall i, \boldsymbol{K}, \boldsymbol{R}, \sigma_i'$$

  - Step 1: recover policy function
  - Step 2: use the optimality constraint to construct the objective function