

Netzwerke in Online-Märkten

Stefan Wehrli Ben Jann

ETH Zürich

Tagung Soziale Netzwerke der DGS Modsim, Köln 2006



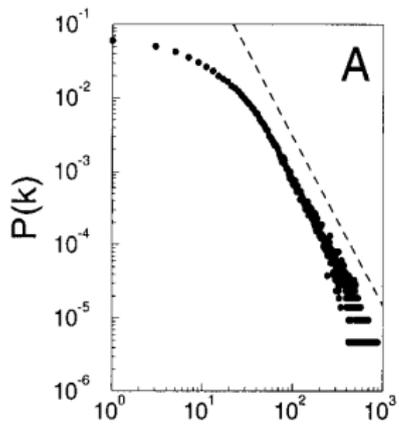
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Übersicht

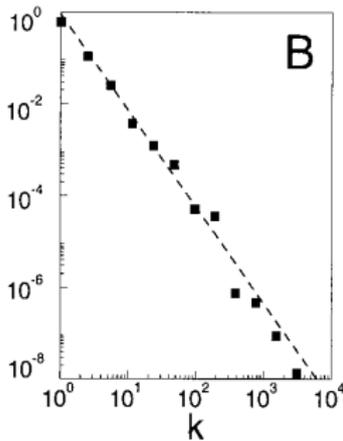
- 1 Einführung
- 2 Netzwerktopologie
- 3 Mikroprozesse
- 4 Discussion

Ausgangslage

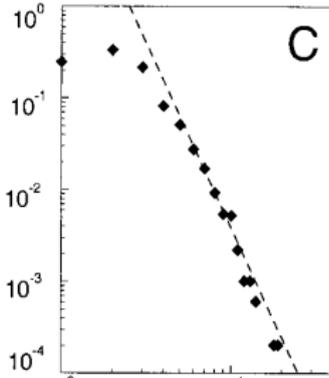
Scale-free Networks (Albert und Barabási, Science 286, 1999)



Movie Actors



WWW



Power Grid

Daten

Feedbackliste auf eBay.de

Bewertungskommentar	Von	Datum/Uhrzeit	Artikelnummer
+ Lief alles sehr zufriedenstellend.	Käufer bschattenkrieger (8)	05.10.04 11:25	8129912307
+ Schnelle Lieferung . gerne wieder	Käufer arek20 (18 ★)	05.10.04 11:00	8114687702
+ Schnell und problemlos gelaufen. Vielen Dank!	Käufer hmoppel (privat)	05.10.04 10:43	6324151694
+ Schnell und problemlos gelaufen. Vielen Dank!	Käufer hmoppel (privat)	05.10.04 10:43	6326427156
+ Hat alles gepasst, schnell und zuverlässig. Danke!	Käufer yroni_d (17 ★)	05.10.04 10:31	6321162557
+ Schnelle Lieferung und netter Kontakt. Sehr empfehlenswerter Ebay-Partner!!!	Käufer belex2000 (318 ★)	05.10.04 10:18	8135424559
+ top ebayer!! hat alles bestens geklappt. sehr netter kontakt	Käufer noja410 (206 ★)	05.10.04 09:01	6317199210
+ einfache Klasse, besser gehts nicht, schnelle korrekte Abwicklung, SUPER, danke!!	Käufer pesp (128 ★)	04.10.04 23:31	6324654423

- Prozessdaten von eBay.de
- Erhebung mit automatischem Spider-Ansatz
- Nur Auktionen, nur selektive Produktkategorien

Daten

Feedback-Profil auf eBay.de

Bewertungsprofil: player004 (63203)   

Bewertungsprofil:	63203
Positive Bewertungen:	99,7%
Mitglieder, die mich positiv bewertet haben:	63376
Mitglieder, die mich negativ bewertet haben:	185
Alle positiven Bewertungen:	85131

[Weitere Informationen](#) zur Bedeutung dieser Zahlen.

Jüngste Bewertungen:

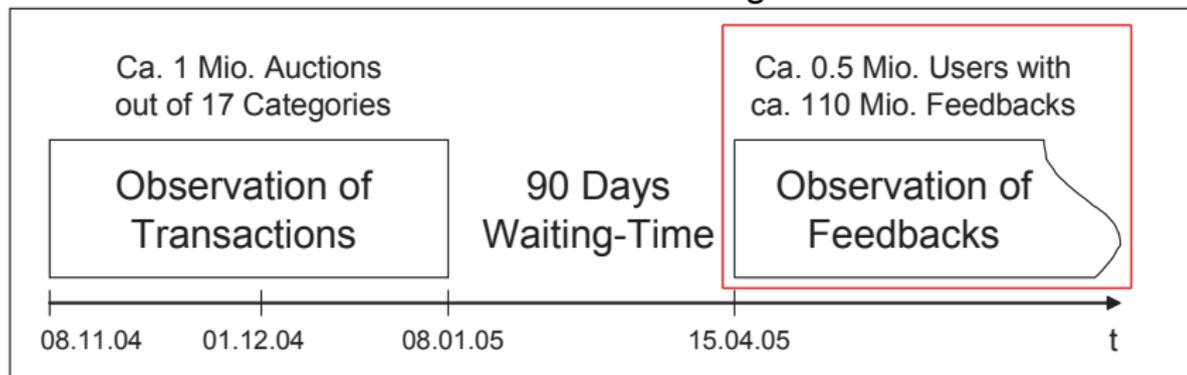
	Letzter Monat	Letzte 6 Monate	Letzte 12 Monate
 positiv	1614	10277	32063
 neutral	9	40	97
 negativ	7	27	64

Zurückgezogene Gebote (in den letzten 6 Monaten): 0

- Prozessdaten von eBay.de
- Erhebung mit automatischem Spider-Ansatz
- Nur Auktionen, nur selektive Produktkategorien

Daten

Ablauf der Erhebung



- Prozessdaten von eBay.de
- Erhebung mit automatischem Spider-Ansatz
- Nur Auktionen, nur selektive Produktkategorien

Reputationsnetzwerk

Reputation

„Reputation is a characteristic or attribute ascribed to one person (firm, industry, etc.) by another.“ (Wilson 1985:27).

Reputationsnetzwerk, Reputationssystem:

- Ein Reputationsnetzwerk ist ein bewerteter und gerichteter Graph von Zuschreibungen multilateral interagierender Akteure (Valued Digraph).
- Ein Reputationssystem ist eine institutionalisierte Aufzeichnung, Aufbereitung und Veröffentlichung von Verhaltenszuschreibungen.

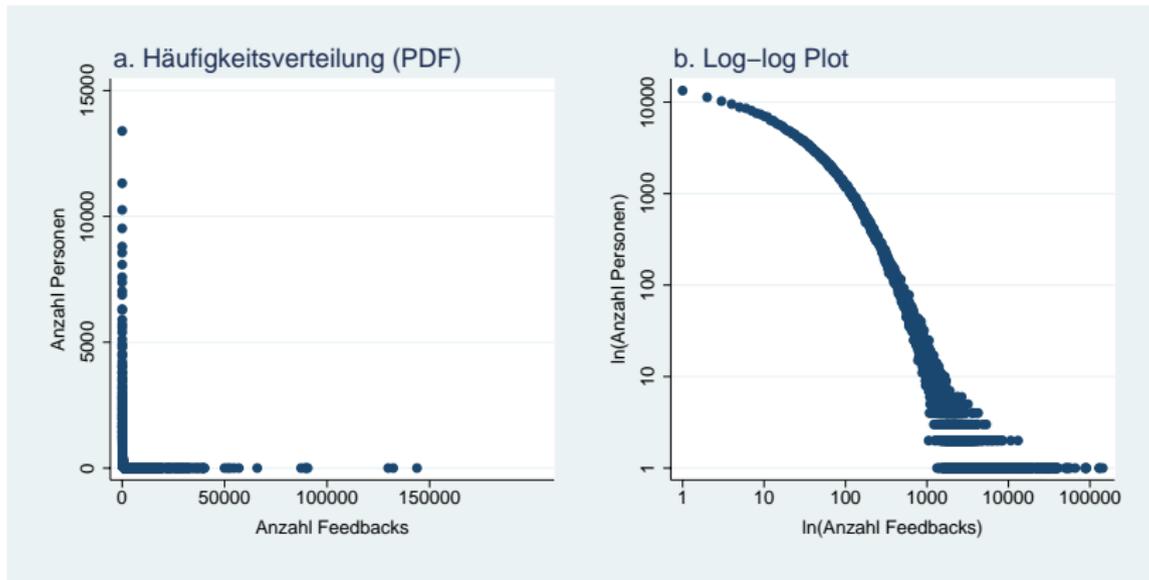
Charakteristika

Eigenschaften komplexer Netzwerke:

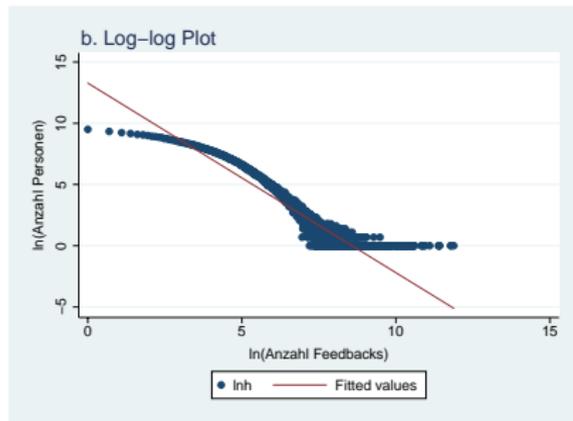
- 1 **Fat Tails: Degree-Verteilungen oft sehr rechtsschief**
- 2 Small Worlds: Kleiner Diameter und kurze durchschnittliche Pfadlängen (Watts & Strogatz 1998)
- 3 Clustering: dichte Nachbarschaften, geschlossene Triaden
- 4 Clustering der Nachbarn eines spezifischen Knotens ist indirekt proportional zu dessen Degree.

Verteilung des Reputationsnetzwerkes

PDF



Verteilung des Reputationsnetzwerkes



Power law:

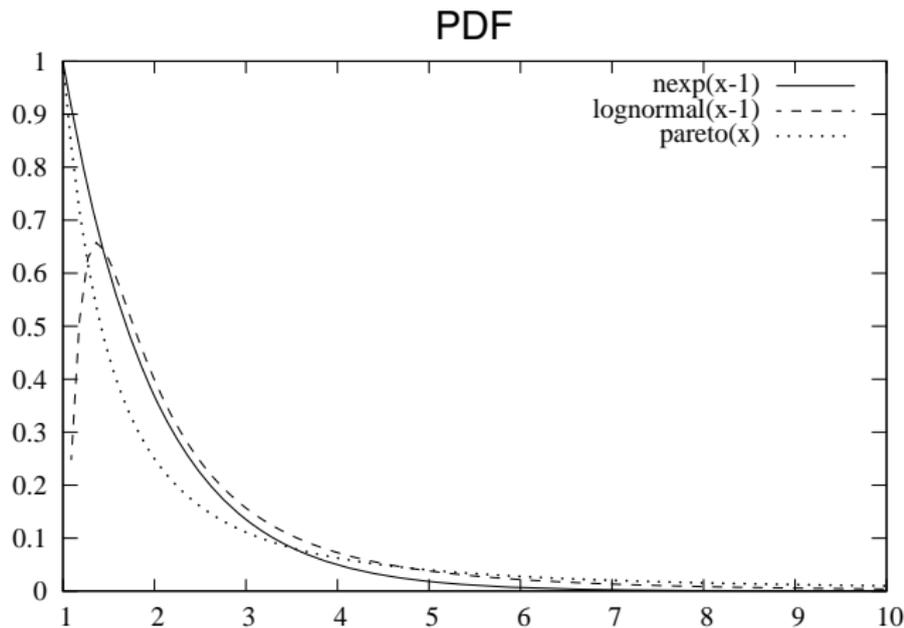
$$Pr(k) \sim k^{-\gamma}$$

OLS-Schätzung:

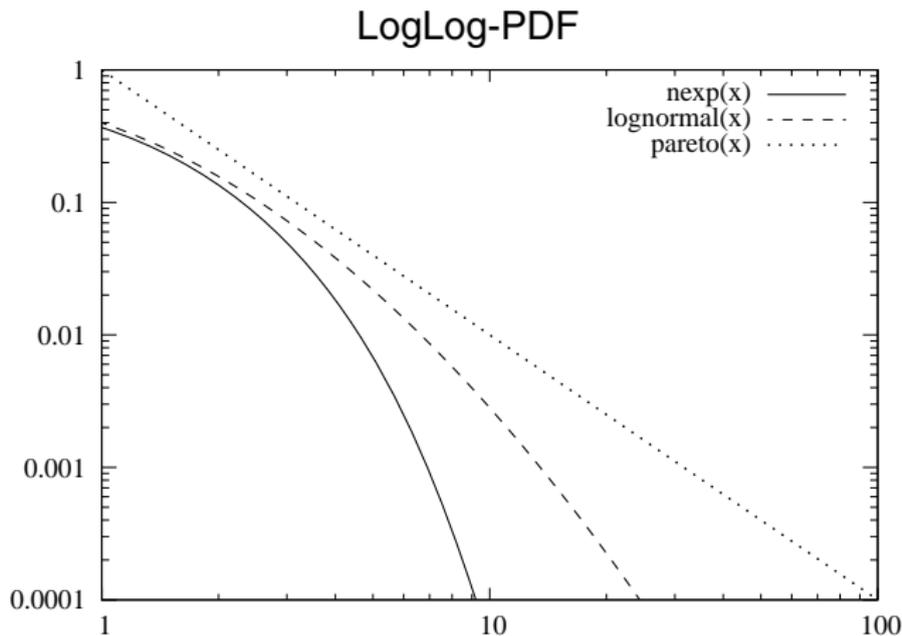
$$\log(Pr(k)) = \log(\alpha) - \gamma \cdot \log(k)$$

$$\hat{\gamma} = 1.548, R^2 = 0.841$$

Mögliche Verteilungen



Mögliche Verteilungen



Pareto Verteilung

Power law (PDF):

$$Pr[K = k] \sim k^{-\gamma}$$

Pareto (CDF und PDF):

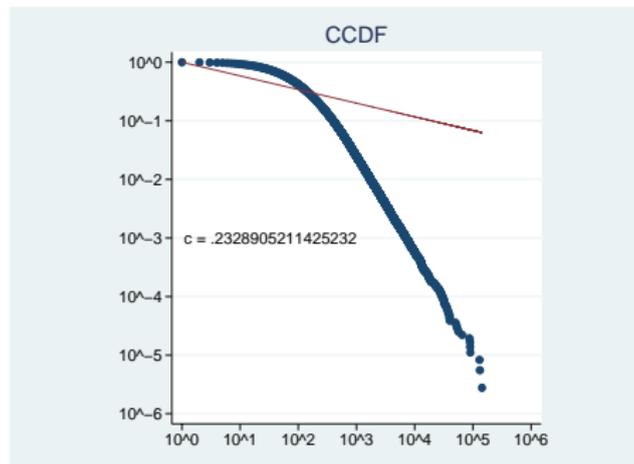
$$Pr[X \leq x] = 1 - (a/x)^c$$

$$Pr[X = x] = ca^c x^{-(c+1)}$$

$$\Rightarrow \gamma = c + 1$$

ML-Estimator:

$$\hat{c} = \frac{1}{\left(\frac{1}{n}\right) \sum_{i=1}^n \log\left(\frac{x_i}{\hat{a}}\right)}$$



Pareto Verteilung

Power law (PDF):

$$Pr[K = k] \sim k^{-\gamma}$$

Pareto (CDF und PDF):

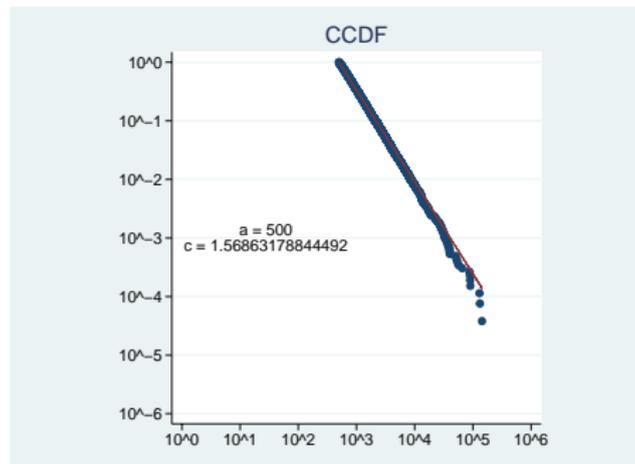
$$Pr[X \leq x] = 1 - (a/x)^c$$

$$Pr[X = x] = ca^c x^{-(c+1)}$$

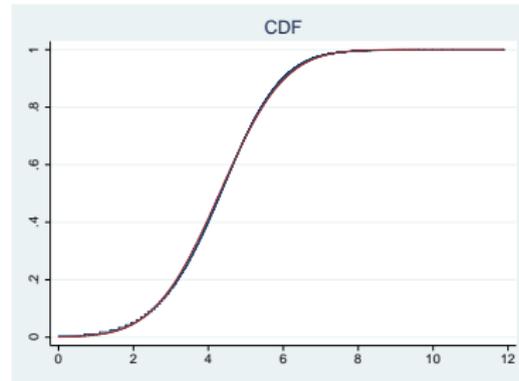
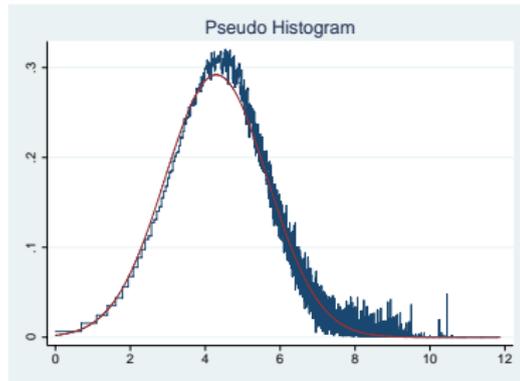
$$\Rightarrow \gamma = c + 1$$

ML-Estimator:

$$\hat{c} = \frac{1}{\left(\frac{1}{n}\right) \sum_{i=1}^n \log\left(\frac{x_i}{\hat{a}}\right)}$$

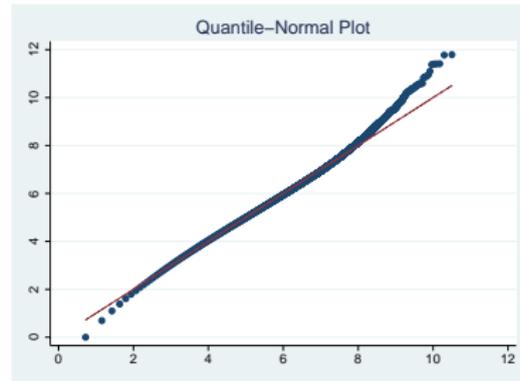
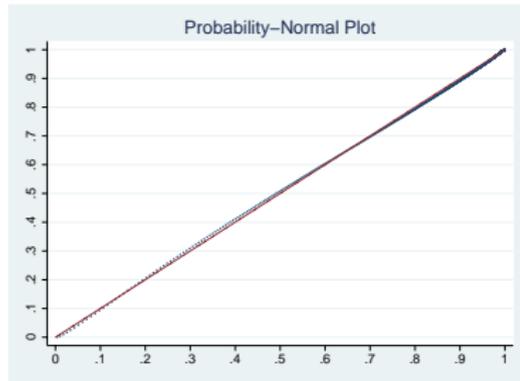


Lognormale Verteilung



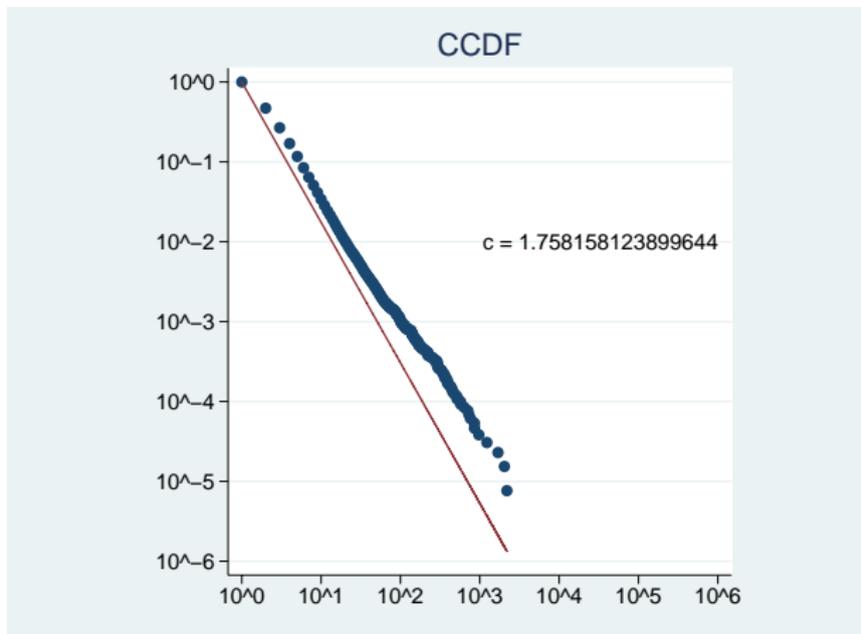
⇒ Positive Bewertungen sind annähernd lognormal verteilt.

Lognormale Verteilung



⇒ Positive Bewertungen sind annähernd lognormal verteilt.

Negative Bewertungen



Fragestellung

Was könnte der generative Prozess sein, der zu einer solchen Verteilung führt?

⇒ Ein Prozess mit proportionalem Wachstum, bzw. „Cumulative Advantage“ („Matthäus-Effekt“) mündet in einer log-normal Verteilung.

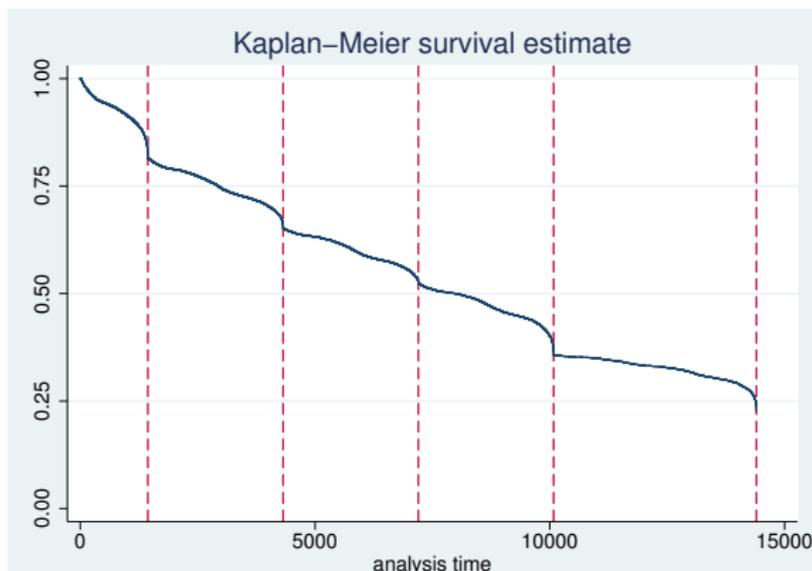
Finden wir Hinweise für Cumulative-Advantage-Mechanismen im eBay-Reputationssystem?

Zwei mögliche Effekte:

- Erhöhung Reputation durch indirekte Effekte auf die „Fitness“
- direkter Effekt der Reputation auf die weitere Reputationsbildung

Analyse 1: Bieter-Verhalten (DVD-Markt)

- N = 34'797 (10% Stichprobe)
- Anzahl zensierte Fälle: 53.5% (Auktionen ohne Gebote)
- Mediandauer bis zum ersten Gebot: 5.55 Tage (25%-Quantil: 2.03 Tage; 75%-Quantil: 9.99 Tage)



Proportional-Hazards-Modelle für das erste Gebot

	Verkäufer		Käufer	
Reputation (log. Score)	0.048***	(0.005)		
Positive Reputation (log.)			0.053***	(0.006)
Negative Reputation (log.)			-0.011	(0.007)
Verk. hat Shop	-0.081***	(0.024)	-0.077**	(0.024)
Verk. hat „Mich“-Seite	-0.216***	(0.023)	-0.210***	(0.023)
Verk. ist geprüftes Mitglied	-0.173***	(0.025)	-0.177***	(0.025)
Verk. ist Powerseller	-0.081**	(0.025)	-0.079**	(0.025)
Startpreis	-0.107***	(0.003)	-0.107***	(0.003)
Versandkosten	-0.098***	(0.006)	-0.098***	(0.006)
Beschreibungslänge (log.)	-0.014*	(0.007)	-0.014*	(0.007)
Galeriebild	0.057*	(0.024)	0.056*	(0.024)
Fettschrift	0.526***	(0.118)	0.529***	(0.118)
Fallzahl	34'797		34'797	
Anzahl „Ereignisse“	18'607		18'607	

Standardfehler in Klammern; Stratifiziert nach Unterkategorie und Wochentag;
Reputation = logarithmierte Anzahl Bewertungen

Analyse 2: Verkaufserfolg (Mobiles)

Verkaufserfolg (Logistische Regression / Mobiles)

	Neu		Gebraucht	
Positive Reputation (log.)	0.142*	(0.071)	-0.014	(0.044)
Negative Reputation (log.)	-0.353*	(0.149)	-0.118	(0.117)
Startpreis	-0.032***	(0.003)	-0.035***	(0.002)
Anbieterkonkurrenz	-0.008*	(0.003)	-0.007**	(0.003)
Auktionsdauer	0.017	(0.038)	0.051*	(0.022)
Kalenderzeit	-0.011	(0.010)	-0.002	(0.006)
Beschreibungslänge (log.)	-0.018	(0.078)	-0.039	(0.060)
Produktbild	0.134	(0.387)	0.331	(0.218)
(...)				
Anteil verkauft	95.06%		95.47%	
Fallzahl	5'338		8'738	

Standardfehler in Klammern (geklumpt); Reputation = logarithmierte Anzahl Bewertungen

Analyse 3: Verkaufspreis (Mobiles)

Effekte auf den Verkaufspreis (OLS / Mobiles)

	Neu		Gebraucht	
Positive Reputation (log.)	1.069***	(0.209)	1.047***	(0.231)
Negative Reputation (log.)	-1.995***	(0.501)	-2.738***	(0.472)
Startpreis	0.034***	(0.006)	0.127***	(0.010)
Anzahl Bieter	0.573***	(0.079)	1.495***	(0.143)
Anbieterkonkurrenz	-0.070***	(0.009)	-0.059***	(0.011)
Auktionsdauer	0.050	(0.109)	-0.105	(0.121)
Kalenderzeit	-0.312***	(0.026)	-0.260***	(0.027)
Kalenderzeit quadriert	0.023***	(0.003)	0.022***	(0.003)
Beschreibungslänge (log.)	1.001***	(0.236)	2.711***	(0.280)
Produktbild	3.736**	(1.274)	4.002*	(1.558)
(...)				
Korrigiertes R ²	0.842		0.789	
Fallzahl	5'096		8'306	

Abhängige Variable: Verkaufspreis (EUR); Standardfehler in Klammern (geklumpt); Reputation = logarithmierte Anzahl Bewertungen

Analyse 4: Vergabe von Bewertungen (DVD-Markt)

Deskriptive Übersicht zur Vergabe von Bewertungen nach Transaktionsende:

	Verkäufer		Käufer	
Anzahl Transaktionen	177'561	(100%)	177'561	(100%)
Positive Bewertung	146'693	(82.62%)	146'300	(82.39%)
Neutrale Bewertung	209	(0.12%)	702	(0.40%)
Negative Bewertung	577	(0.32%)	643	(0.36%)
Keine Bewertung	30'082	(16.94%)	29'916	(16.85%)

- In 89.6% aller Auktionen wurde mindestens eine Bewertung vergeben.
- In 53.2% hat der Käufer zuerst bewertet (Verkäufer: 36.1%; Gleichzeitig/kein Feedback: 10.7%).
- Anteil wiederholter Interaktionen \approx 5%.

Proportional-Hazards-Modelle für die Vergabe von Bewertungen

	Verkäufer		Käufer	
Partner zuerst (zeitveränd.)	2.021***	(0.118)	0.617***	(0.011)
Positive Partnerreputation	0.088***	(0.006)	0.098***	(0.004)
Negative Partnerreputation	-0.184***	(0.016)	-0.085***	(0.006)
Partner zuerst X pos. P.rep.	-0.099***	(0.010)	-0.029***	(0.005)
Partner zuerst X neg. P.rep.	0.163***	(0.023)	-0.117***	(0.010)
Eigene positive Reputation	0.050**	(0.016)	0.141***	(0.004)
Eigene negative Reputation	-0.199***	(0.029)	-0.253***	(0.010)
Wiederholte Interaktion	-0.281***	(0.077)	-0.467***	(0.033)
Rollenwechsel	0.175***	(0.051)	0.185***	(0.022)
Geprüftes Mitglied	0.210*	(0.084)	0.190***	(0.037)
Preis	-0.002	(0.001)	-0.000	(0.000)
(...)				
Fallzahl (Klumpen)	177'561	(29'816)	177'561	(99'139)
Anzahl „Ereignisse“	146'693		146'300	

Bootstrap-Standardfehler in Klammern (geklumpt); Reputation = logarithmierte Anzahl Bewertungen

Schlussfolgerung

Indirekte und direkte Cumulative-Advantage-Effekte der Reputation können empirisch gezeigt werden. Aber sind dies wirklich die zentralen Mechanismen, die hinter der Reputationsbildung stehen?

Für eine fruchtbare Analyse des Reputationssystems wird ein theoretisches Modell der Mikroprozesse benötigt, das die beobachtete Netzwerktopologie erzeugen kann und aus dem sich empirisch prüfbare Hypothesen ableiten lassen.

Interessant wäre in diesem Zusammenhang zudem, weitere Netzwerk-Eigenschaften zu bestimmen (Clustering, durchschnittliche Pfadlängen, etc.) und z.B. verschiedene eBay-Teilmärkte zu vergleichen.