# The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment

Chris Nosko Steve Tadelis

University of Chicago UC Berkeley and NBER

November 16, 2015

# Platform Markets and Quality Control

**Platform markets differ from retailers:**

- Facilitate trade between anonymous buyers and sellers
- Do not control key variables (inventory, price, transaction quality,...)
- Variance in the quality of sellers on the platform

# Platform Markets and Quality Control

**Platform markets differ from retailers:**

- Facilitate trade between anonymous buyers and sellers
- Do not control key variables (inventory, price, transaction quality,...)
- Variance in the quality of sellers on the platform

**Reputation/Feedback:**

- Lauded as facilitating trade (reveals information to participants)
  - eBay, Taobao, AirBnB, Uber (Amazon product reviews, Yelp, TripAdvisor)
- Presented as “self regulatory” mechanisms for quality control

# Platform Markets and Quality Control

**Platform markets differ from retailers:**

- Facilitate trade between anonymous buyers and sellers
- Do not control key variables (inventory, price, transaction quality,...)
- Variance in the quality of sellers on the platform

**Reputation/Feedback:**

- Lauded as facilitating trade (reveals information to participants)
  - eBay, Taobao, AirBnB, Uber (Amazon product reviews, Yelp, TripAdvisor)
- Presented as "self regulatory" mechanisms for quality control

**For reputation systems to work:**

- Reputation measures should accurately reflect quality
- Buyers should correctly perceive reputations-to-quality mapping

# Possible Concerns with Reputation/Feedback Mechanisms


http://xkcd.com/1098/

# Contributions

- **Highlight issues missing from traditional platform models:**
  - Asymmetric information (seller quality or effort)
  - Quality spillovers/externalities between sellers on platform

# Contributions

- **Highlight issues missing from traditional platform models:**
  - Asymmetric information (seller quality or effort)
  - Quality spillovers/externalities between sellers on platform
- **Highlight issues missing from traditional models of reputation:**
  - Explicit discussion of heterogeneous costs of leaving feedback
  - Often can lead to skewed or uninformative reputation systems

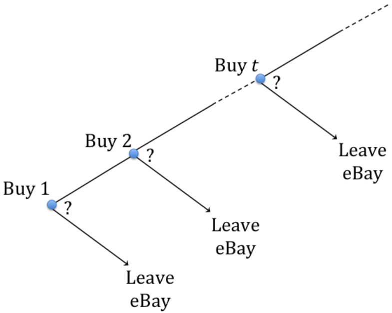
# Contributions

- **Highlight issues missing from traditional platform models:**
  - Asymmetric information (seller quality or effort)
  - Quality spillovers/externalities between sellers on platform
- **Highlight issues missing from traditional models of reputation:**
  - Explicit discussion of heterogeneous costs of leaving feedback
  - Often can lead to skewed or uninformative reputation systems
- **Argue that marketplaces need to augment feedback systems**
  - Have better incentives than individual sellers to self regulate
  - Can find information in data that indicates seller quality
  - Offer "proof of concept" not optimal solution (engineering)

# Contributions

- **Highlight issues missing from traditional platform models:**
  - Asymmetric information (seller quality or effort)
  - Quality spillovers/externalities between sellers on platform
- **Highlight issues missing from traditional models of reputation:**
  - Explicit discussion of heterogeneous costs of leaving feedback
  - Often can lead to skewed or uninformative reputation systems
- **Argue that marketplaces need to augment feedback systems**
  - Have better incentives than individual sellers to self regulate
  - Can find information in data that indicates seller quality
  - Offer "proof of concept" not optimal solution (engineering)
- **Suggest to use search to affect buyer experience and outcomes**
  - CS literature documents the impact of ranking on choice
  - Intervene in search algorithm to control for seller quality

# Conceptual Framework




- A Buyer's Dynamic Bayesian Decision Problem: buy again if,
  - Had good past experiences **relative to expectations**
- Buyers may use outcomes to update on platform, not just seller!

# What do buyers use to form expectations? Reputation!

After every eBay transaction

- Buyers **choose** to leave feedback (positive, negative, neutral, nothing)

# What do buyers use to form expectations? Reputation!

After every eBay transaction

- Buyers **choose** to leave feedback (positive, negative, neutral, nothing)

Information is aggregated and displayed to potential future buyers as:

- Percent positive: $(\frac{pos}{neg+pos})$
- Seller feedback score: $(pos - neg)$
- Seller standards: (ETRS)

## APPLE MACBOOK 13.3 HD,OSX 10.6,CORE 2 DUO,RAM 1 GB, 2.16 GHZ,120GB HD,GREAT COND

★★★★★ 7 product reviews

Item condition: **Used**
"GREAT CONDITION, TESTED AND IN GREAT WORKING CONDITION, 120 GB HDD,13.3 HD ,OSX 10.6,COMES WITH"
... Read more

Quantity: 1 5 available

Price: **US $274.99** Buy It Now Add to cart

SquareTrade 2 yr warranty **$79.99**

Best Offer: Make Offer

1 watcher Add to Watch list

BillMeLater **Spend $99+ and get 6 months to pay**
Subject to credit approval. See terms

Shipping: **FREE** Economy Shipping | See details
Item location: **Holiday, Florida, United States**
Ships to: **Worldwide**

Delivery: Estimated between **Tue. Sep. 3 and Wed. Sep. 11**

Payments: PayPal, Bill Me Later | See details

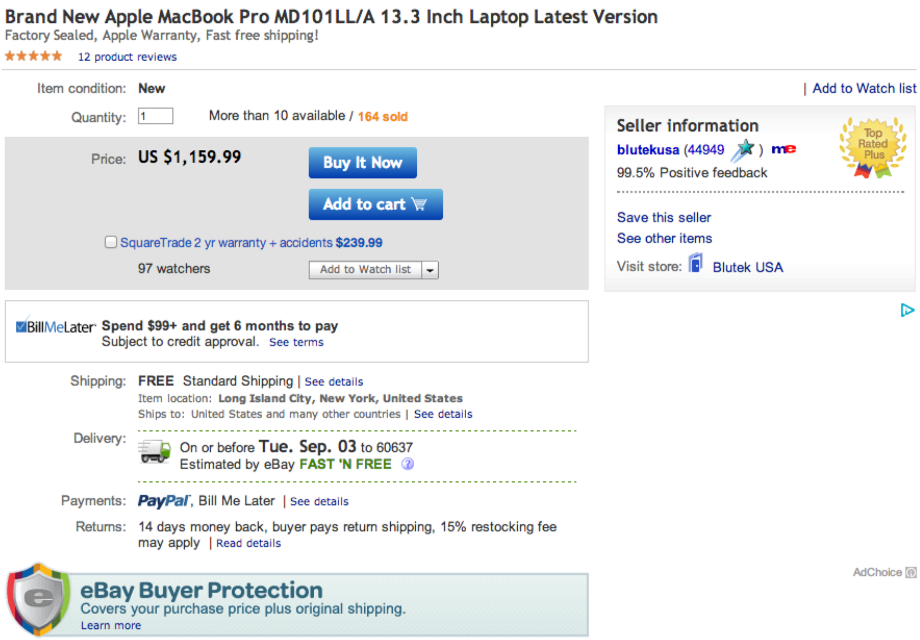
Returns: 14 days money back, buyer pays return shipping | Read details

**eBay Buyer Protection**
Covers your purchase price plus original shipping.
Learn more

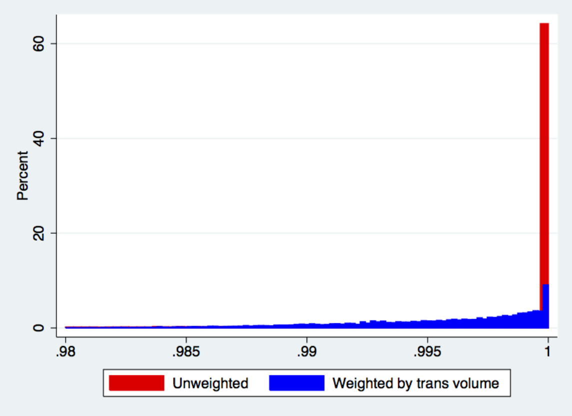
| Add to Watch list

**Seller information**
**samnas04** (317 ☆ )
96.9% Positive feedback

Save this seller
See other items


Brand New Apple MacBook Pro MD101LL/A 13.3 Inch Laptop Latest Version
Factory Sealed, Apple Warranty, Fast free shipping!
12 product reviews
Item condition: New
| Add to Watch list
Quantity: 1
More than 10 available / 164 sold
Price: US $1,159.99
Buy It Now
Add to cart
SquareTrade 2 yr warranty + accidents $239.99
97 watchers
Add to Watch list
Seller information
blutekusa (44949 )
99.5% Positive feedback
Top Rated Plus
Save this seller
See other items
Visit store: Blutek USA
BillMeLater
Spend $99+ and get 6 months to pay
Subject to credit approval. See terms
Shipping: FREE Standard Shipping | See details
Item location: Long Island City, New York, United States
Ships to: United States and many other countries | See details
Delivery: On or before Tue. Sep. 03 to 60637
Estimated by eBay FAST 'N FREE
Payments: PayPal, Bill Me Later | See details
Returns: 14 days money back, buyer pays return shipping, 15% restocking fee may apply | Read details
AdChoice
eBay Buyer Protection
Covers your purchase price plus original shipping.
Learn more

# Distribution of Reputation on ebay




- median = 100%, mean = 99.3%, $10^{th}$ percentile = 97.8%
- **Case 1:** Sellers whose reputation drops are kicked out
- **Case 2:** Feedback is heavily biased

# Is this Nirvana?

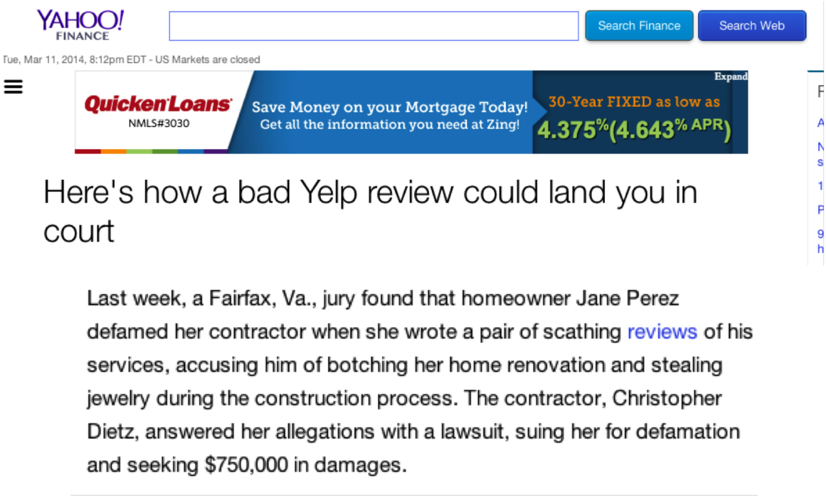
But, out of 44,604,802 transactions in October 2011:

| | No. of Trans. | % of Trans |
|---|---|---|
| Negative FB | 172,850 | 0.39% |
| email sent ex post | 3,858,757 | 8.65% |
| “Bad” email sent | 1,491,126 | 3.34% |
| Dispute filed | 466,971 | 1.05% |

# Leaving negative feedback is costly!


- The first message he saved on his voicemail: "Don't you play games with me, goddamn you. I'll follow you to your grave."
- "He knew everything about me," said Blackwelder. "My phone number, my address, my name. ... It's a little scary."

# Leaving negative feedback is costly!



YAHOO! FINANCE

Search Finance Search Web

Tue, Mar 11, 2014, 8:12pm EDT - US Markets are closed



## Here's how a bad Yelp review could land you in court

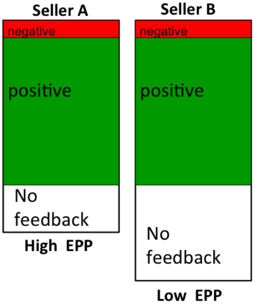
Last week, a Fairfax, Va., jury found that homeowner Jane Perez defamed her contractor when she wrote a pair of scathing reviews of his services, accusing him of botching her home renovation and stealing jewelry during the construction process. The contractor, Christopher Dietz, answered her allegations with a lawsuit, suing her for defamation and seeking $750,000 in damages.

# Feedback is Biased

- Leaving feedback is a hassle but that does not imply bias
  - Bias will happen if the cost of leaving feedback depends on the transaction quality
- **Claim:** Leaving negative feedback is “more costly” than leaving positive feedback
  - Harassing emails following negative
  - Threats of lawsuits and other harassment
  - Historical norm of reciprocity
- Implies that silence has more negative experiences than random
- **We can use this silence to help measure quality!**

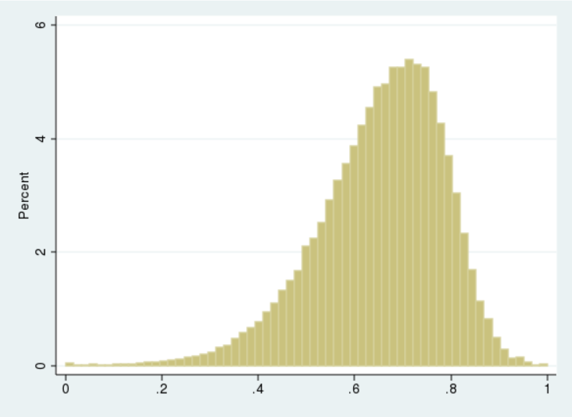
# Effective Percent Positive (EPP)

$$EPP = \frac{\#\text{ of positive feedback}}{\#\text{ of transactions}}$$




- **Seller A:** $P = 99$, $N = 1$, Silence $= 20 \rightarrow PP = 99\%$, $EPP = 82.5\%$
- **Seller B:** $P = 99$, $N = 1$, Silence $= 50 \rightarrow PP = 99\%$, $EPP = 66\%$
- Seller A is **higher quality** than seller B!

# EPP Distribution




- A lot more "spread" and information in EPP
- **But is it really a measure of seller quality?**

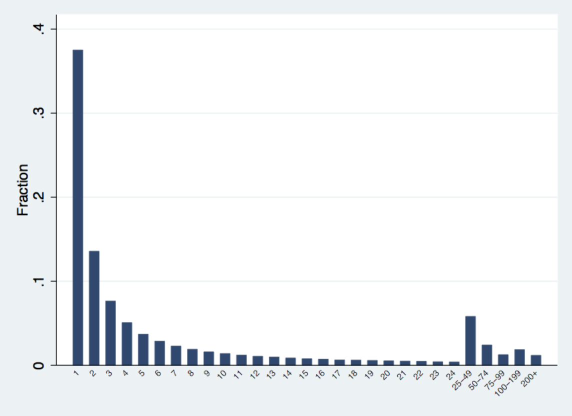
# Data

- Cohort of new users who joined the the U.S. site anytime in 2011 and purchased an item within 30 days of setting up that account. (also run the analysis on 2008, 2009, 2010)
  - 10% random sample = 935,326 buyers
  - Tracked all of their usage purchase behavior until May 31, 2014 (15,384,439 observations)
  - Data includes price, item category, title, the seller, auction or fixed price, quantity purchased, etc.

# Data

- Cohort of new users who joined the the U.S. site anytime in 2011 and purchased an item within 30 days of setting up that account. (also run the analysis on 2008, 2009, 2010)
  - 10% random sample = 935,326 buyers
  - Tracked all of their usage purchase behavior until May 31, 2014 (15,384,439 observations)
  - Data includes price, item category, title, the seller, auction or fixed price, quantity purchased, etc.
- There were a total of 1,854,813 sellers associated with all purchases
  - Seller information includes feedback score, PP, number of past transactions, etc.
  - For each transaction we look backward construct an EPP measure for that seller.
- We apply this data to our conceptual dynamic decision framework

# The Distribution of Buyer Purchases




- 38% of new buyers purchase once and leave; an additional 14% purchase twice; the mean is 16 purchases before leaving ebay.
- Large right tail: the median number of transactions is 2, the 95th percentile is 65, and the max is 19.359.

# The Scope for Externalities is real

Table: Total Transactions by Total Number of Sellers for buyers

| Total Transactions | Total Number of Sellers 00-01 | 02-05 | 06-09 | 10-19 | 20-29 | 30-49 | Total |
|---|---|---|---|---|---|---|---|
| 00-01 | 350,881 | 0 | 0 | 0 | 0 | 0 | 350,881 |
| 02-05 | 27,603 | 253,032 | 0 | 0 | 0 | 0 | 280,635 |
| 06-09 | 1,206 | 19,374 | 60,590 | 0 | 0 | 0 | 81,170 |
| 10-19 | 492 | 2,802 | 15,959 | 64,112 | 0 | 0 | 83,365 |
| 20-29 | 116 | 386 | 767 | 13,513 | 23,367 | 0 | 38,149 |
| 30-49 | 67 | 207 | 273 | 1,810 | 11,685 | 24,106 | 38,148 |
| Total | 380,365 | 275,801 | 77,589 | 79,435 | 35,052 | 24,106 | 872,348 |

- This suggests that most buyers are not "loyal" to sellers, but come to ebay to purchase from multiple sellers

# The Main Regression: "Loyalty, (Voice) and Exit"

Use a "revealed preference" approach: happy buyers are more likely to come back

$$\text{To Seller: } y_{ijt} = \alpha_0 + \alpha_1 EPP_{jt} + \beta \cdot \bar{b}_{it} + \gamma \cdot \bar{s}_{jt} + \delta \cdot \bar{d}_t + \varepsilon_{ijt}$$
$$\text{To Platform: } y_{it} = \alpha_0 + \alpha_1 EPP_{jt} + \beta \cdot \bar{b}_{it} + \gamma \cdot \bar{s}_{jt} + \delta \cdot \bar{d}_t + \varepsilon_{ijt}$$

$y_{ijt} = 1$ if buyer $i$ bought transaction $t$ from seller $j$ and returned to **seller** $j$

$y_{it} = 1$ if buyer $i$ bought transaction $t$ from seller $j$ and returns to **eBay**

$\bar{b}_{it}$ is a vector of buyer characteristics (# of transactions they completed...)

$\bar{s}_{jt}$ is a vector of seller characteristics (score, PP, ...)

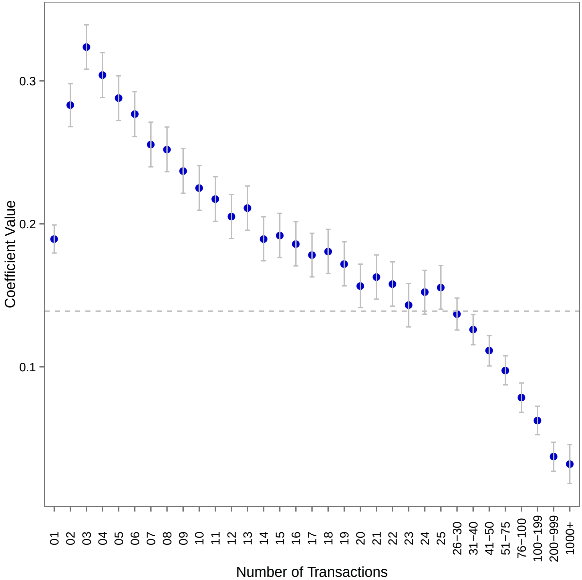
$\bar{d}_t$ is a vector of transaction characteristics (auction, price,...)

**Difference between the two is a measurement of the potential for seller externalities**

# Main Regression Results: Ever Return

| | Same Seller | eBay |
|---|---|---|
| EPP Dummy | | |
| (excluded: $0 < .517$) | | |
| $\geq .517 < .592$ | 0.00477** | 0.0192** |
| | 0.00154 | 0.000253 |
| $\geq .592 < .668$ | 0.0212*** | 0.0289** |
| | 0.00178 | 0.000285 |
| $\geq .668$ | 0.0199*** | 0.0399** |
| | 0.00221 | 0.000317 |
| Seller Feedback Score | -0.000000385*** | -1.52e-09 |
| | 2.13e-08 | 1.55e-09 |
| Percent Positive Dummy | | |
| (excluded: $0 < .994$) | | |
| $\geq .994 < 1$ | 0.0320*** | -0.00897** |
| | 0.00140 | 0.000210 |
| $= 1$ | -0.0353*** | -0.0102** |
| | 0.00162 | 0.000295 |
| Item Price | -0.000326*** | -0.000316** |
| | 0.0000151 | 0.00000381 |
| Seller Standards Dummy | | |
| (excluded: Below Standard) | | |
| Standard | -0.0908*** | -0.00840** |
| | 0.00232 | 0.000474 |
| Above Standard | -0.00534** | -0.00763** |
| | 0.00192 | 0.000412 |
| ETRS | -0.00512* | -0.0115** |
| | 0.00210 | 0.000425 |
| Constant | 0.169*** | 0.506** |
| | 0.00490 | 0.000828 |
| N | 11,879,306 | 12,820,329 |

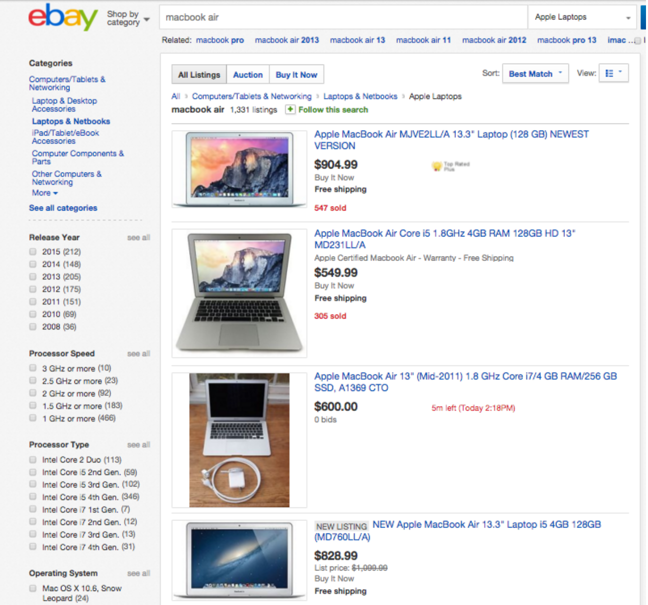
# Buyers Behave as Bayesian Learners: EPP effect over time



# Implementation: Incorporate EPP in Search

- Online marketplaces use search algorithms to direct users
  - Users put in queries for what they want to buy
  - The marketplace uses a variety of inputs to direct search (relevance, price,...)
- Incorporating seller quality can take any form between two extremes
  - **Hard hand:** very minor seller problems cause the seller to never appear (kick out)
  - **Laissez Fair:** give buyers feedback and let them decide who to buy from
- **Healthy middle ground:** sacrifice some relevance for quality

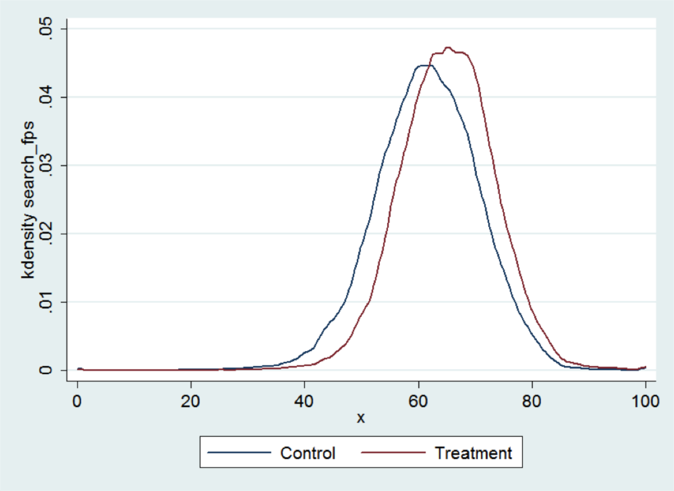
# Manipulating Search



# Large Scale Field Experiment

- We conduct an experiment to manipulate search rankings in order to:
  1. Reinforce observational data regressions
  2. Demonstrate a middle ground in platform governance
- Implementation: treatment ranking algorithm incorporates EPP
  - December 14th, 2011 though January 2, 2012
  - 10% of ebay's U.S. site traffic—about 5 million searches per day
  - selection into treatment uses GUID (cookie) → measurement error
- Collect data both during and **after** experiment to measure outcomes
  - Main analysis: Conditional on purchase, are buyers in the treatment group more likely to come back to eBay?

# Measuring Treatment Effect: Discounted Search EPP



# Intent to treat estimates

Table: Two-sample test of proportions

| Group | Obs | Mean | Std. Err. | 95% Conf. Interval | |
|---|---|---|---|---|---|
| Control | 11,486,810 | .6155062 | .0001435 | .6152249 | .6157875 |
| Treatment | 1,258,455 | .6185275 | .000433 | .6176788 | .6193762 |
| diff | | **.0030213** | .0004562 | .0021272 | .0039153 |
| diff | prop(1) - prop(0) | | | | **z = 6.6151** |

$$\frac{\Delta Pr\{\text{return}\}}{\Delta \text{DSEPP}} = \frac{(0.6185275 - 0.6155062)}{(0.6227 - 0.6157)} = 0.43$$

- Quite a bit higher than 0.14 from non-experimental OLS, but controlling for observables brings this much closer (about 0.16)
- Experimental results also support the "Bayesian Updating" framework

# Experimental Results: Effect on Treated

Table: Probability of return in 180 days

| | ols<br>b/se | ols<br>b/se | firststage<br>b/se | ivresults<br>b/se |
|---|---|---|---|---|
| EPP | | 0.261*** | | 0.246*** |
| | | 0.00174 | | 0.0985 |
| Treatment Dummy | 0.00137** | | 0.00557*** | |
| | 0.000550 | | 0.000134 | |
| Seller Feedback Score | 8.94e-09*** | 5.64e-09*** | 1.27e-08*** | 5.83e-09*** |
| | 6.07e-10 | 6.07e-10 | 1.48e-10 | 1.39e-09 |
| Percent Positive Dummy | | | | |
| excluded: $0 < .994$ | | | | |
| $\geq .994 < 1$ | 0.0145*** | -0.00760*** | 0.0847*** | -0.00631 |
| | 0.000403 | 0.000429 | 0.0000984 | 0.00835 |
| $= 1$ | 0.0203*** | -0.00740*** | 0.106*** | -0.00579 |
| | 0.000563 | 0.000592 | 0.000137 | 0.0105 |
| Item Price | -0.0000662*** | -0.0000624*** | -0.0000144*** | -0.0000626*** |
| | 0.000000943 | 0.000000941 | 0.000000230 | 0.00000170 |
| Seller Standards Dummy | | | | |
| excluded: Below Standard | | | | |
| Standard | -0.0420*** | -0.0366*** | -0.0208*** | -0.0369*** |
| | 0.00116 | 0.00116 | 0.000284 | 0.00236 |
| Above Stand | -0.0208*** | -0.0197*** | -0.00433*** | -0.0198*** |
| | 0.00106 | 0.00105 | 0.000258 | 0.00114 |
| ETRS | -0.0383*** | -0.0339*** | -0.0166*** | -0.0342*** |
| | 0.00105 | 0.00105 | 0.000256 | 0.00195 |
| Constant | 0.782*** | 0.634*** | 0.566*** | 0.643*** |
| | 0.00108 | 0.00146 | 0.000265 | 0.0558 |
| N | 5502532 | 5503316 | 5502532 | 5502532 |

# No "costs" of relevance

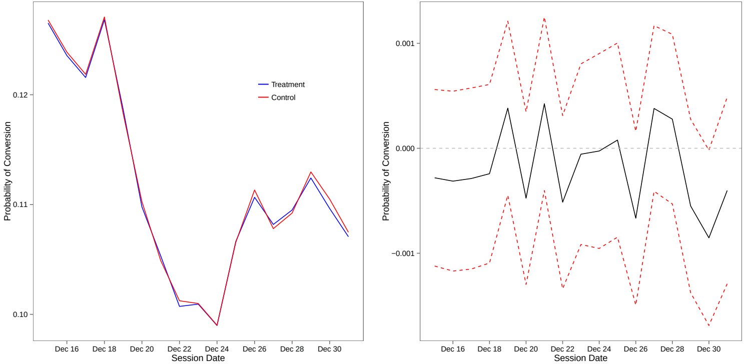



Figure: Differences in Prob. of Purchase across Groups During the Experiment

- No impact of including EPP on relevance for the treatment group
- Recall: VERY modest change in the search algorithm

# Intent to treat: Bayesian Updating

Table: Intent to treat estimates by quartile

| LHS: Prob of return | b/se |
|---|---|
| Treatment dummy | |
| Excluded: Control | |
| Treatment | 0.00249*** |
| | 0.000726 |
| Top quartile dummy | |
| Excluded: Bottom quartile | |
| Top quartile | 0.582*** |
| | 0.000326 |
| Interaction dummy | |
| Top quartile * treatment | -0.00219* |
| | 0.00104 |
| Constant | 0.294*** |
| | 0.000227 |
| N | 6,655,839 |

# Using Messages (Masterov, Mayer and Tadelis, EC15)

Should buyers ever send sellers messages *after* a transaction has completed?

# Using Messages (Masterov, Mayer and Tadelis, EC15)

Should buyers ever send sellers messages *after* a transaction has completed?

**Non Negative:**

Hallo and good morning, I am glad to tell you that the two items arrived safely today. So I am very happy because I needed it for a dinner party tomorrow evening. Thanks a lot and kind regards [name]

# Using Messages (Masterov, Mayer and Tadelis, EC15)

Should buyers ever send sellers messages *after* a transaction has completed?

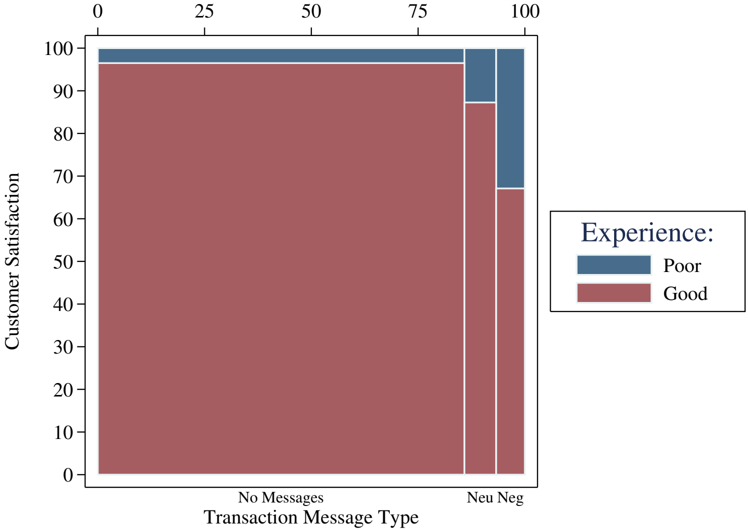
**Non Negative:**

Hallo and good morning, I am glad to tell you that the two items arrived safely today. So I am very happy because I needed it for a dinner party tomorrow evening. Thanks a lot and kind regards [name]

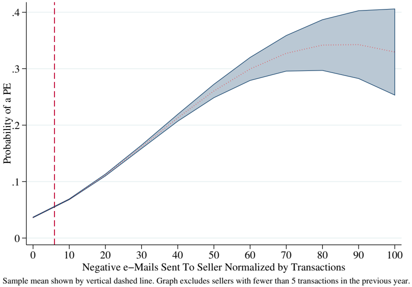
**Negative:**

I purchased two pairs of shorts from you at the same time one pair gold/red (which i have recieved) and one pair ebony/red (which i havent recieved), so i should still be recieving a refund for the ebony/red pair,as i have paid you for them and it's in my payment history, the item number is [...] and the describtion says [auction title]

# Distribution of Poor Experiences By Message Type



# Probability of Poor Experience by B2S Quality Score




- M2M message quality score can be used to flag sellers that cause poor experiences and as a result may disengage buyers.
- We show that this M2M message quality score as as much **independent** power as EPP does in predicting exit.

# Concluding Remarks

- Platform markets face challenges of asymmetric information
- **Externalities** across sellers and **bias** limit feedback effectiveness
- This discussion is missing from the academic literature
- Contributions:
  - Uncover biases and reputational externalities in a large platform market
  - Suggest a general approach of "active screening" by platforms
  - Suggest further Improvements with personalized search
  - Follow up using email messages (w/ Materov and Mayer, EC 2015)
- Growth of online marketplace will depend on how they augment biased feedback mechanisms with active screening approaches
- Implies that marketplaces have the incentives to self-regulate
- Cat and mouse game? (disequilibrium...)