

# Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem

**Kyle Soska**

Carnegie Mellon University

ECE / Cylab

[ksoska@cmu.edu](mailto:ksoska@cmu.edu)

Nicolas Christin

Carnegie Mellon University

ECE / Cylab

[nicolasc@cmu.edu](mailto:nicolasc@cmu.edu)

# Conventional Commerce



# Internet Commerce

1-24 of 509,703 results for Home & Kitchen : Home Décor : Clocks : Wall Clocks : "wall clock"

Show results for

- < Any Category
- < Home & Kitchen
- < Home Décor
- < Clocks
- Wall Clocks**

Refine by

Amazon Prime



Avg. Customer Review

- ★★★★★ & Up (4,921)
- ★★★★☆ & Up (6,334)
- ★★★☆☆ & Up (6,984)
- ★★☆☆☆ & Up (7,693)

Brand

- Howard Miller
- Infinity Instruments
- Seiko
- La Crosse Technology
- Geneva
- iCasso
- Westclox
- Advance Clock
- Decho
- Tiim
- LexMod
- nito
- Lorell
- Chaney Instruments
- Decodyne
- + See more

Price

- Under \$25 (312,090)
- \$25 to \$50 (154,383)
- \$50 to \$100 (23,065)
- \$100 to \$200 (14,692)
- \$200 & Above (6,008)

\$  to \$

Home Décor Theme

Showing results in Home & Kitchen. Show instead results in All Departments.

Related Searches: [large wall clock](#).

Home Décor Theme: [Vehicles & Transportation](#) | [Churches, Buildings & Houses](#) | [Romance & Love](#) | [Animals](#) | [Sports](#) | [See more](#)



[See Size Options](#)

Universal Indoor/Outdoor Clock, 13 1/2-Inch, Black (11381)  
by Universal

**\$16.05** ~~\$23.99~~ Prime  
Get it by **Thursday, Aug 6**

More Buying Choices  
**\$12.38** new (33 offers)  
**\$11.97** used (1 offer)

**#1 Best Seller** in Wall Clocks  
★★★★★ 418



[See Package Quantity Options](#)

ADVANCE CLOCK CO. 10" Wall Clock [Black]  
by Advance Clock

**\$9.19** Prime  
Get it by **Thursday, Aug 6**

More Buying Choices  
**\$6.29** new (28 offers)  
★★★★☆ 387



[See more choices](#)

Ikea Wall Clock, White, Clear  
by IKEA

**\$6.93** ~~\$44.99~~ Prime  
Only 12 left in stock - order soon.

More Buying Choices  
**\$1.75** new (28 offers)  
[Show only IKEA items](#)  
★★★★☆ 129



La Crosse Technology WT-8002U Digital Wall Clock  
by La Crosse Technology

**\$18.75** ~~\$49.95~~ Prime  
Get it by **Thursday, Aug 6**

More Buying Choices  
**\$18.75** new (7 offers)  
**\$14.95** used (6 offers)

★★★★☆ 716



[See Color Options](#)

12" Vintage France Paris Colourful French Country Tuscan Style Paris Wood Wall Clock  
by Decho

**\$7.80**  
More Buying Choices  
**\$7.80** new (37 offers)

FREE Shipping  
★★★★☆ 202



[See Size Options](#)

Geneva 14-Inch Plastic Decor.  
by Geneva






























**\$26.06**  
More Buying Choices  
**\$17.29** new (9 offers)  
FREE Shipping

★★★★☆ 261

# Conventional Illicit Commerce



# Illicit Internet Commerce

				
cheeba chews quad dose 0.49443757 BTC	Purple Romulan (114 g) 2.29556772 BTC	1oz. 0.77697333 BTC	AB-CHMINACA 1KG 6.71022426 BTC	DISCOUNT PRICE 7g 0.41674024 BTC
				
3.5g HIGHGRADE 0.23733003 BTC	7G OF GOLD POLM 0.29118841 BTC	3g soft squigy Afghan 0.13773618 BTC	100G OF PP100 HASH 3.25269291 BTC	CannaJuice Variety 0.51209606 BTC
				
3g Blue Dream AAA+ 0.40610983 BTC	25 Canna-Caps - 0.38848666 BTC	112 grams mazar 2.18965212 BTC	HempVap Hemp Oil vape 0.12360939 BTC	1G OF MOROCCAN 0.05512978 BTC
				
AFGHAN GOLD SEAL 0.12360939 BTC	QP 4oz / 112g / quarter 2.66643122 BTC	Norge: 5 gram Oslo 0.42380363 BTC	Afghan Seeds x 3 0.04944375 BTC	Super Strength African 0.10909411 BTC
				
3.5g OG Kush! TRUE 0.10595090 BTC	2.2g HIGHGRADE 0.12926010 BTC	3.5g BUD + MID GRADE 0.15892636 BTC	Swazi Kush 100g 1.41267879 BTC	20 Mixed Marijuana 0.17654953 BTC
				3.5g super mix
28G OF KETAMA GOLD 1.10259579 BTC	Cannatonic high-CBD Oil 0.30019424 BTC	Blue Dream Bubble Hash 0.10591559 BTC	1/4th Grand Daddy 0.31432103 BTC	3.5g super mix 0.17658484 BTC

# Anonymous Marketplaces

- Amazon.com of illegal goods
  - Drugs, CC's & Fake IDs, Weapons, etc.
  - No Child Porn
- Safety
- Convenience
- Variety
- Accountability
- Competition

# Anonymous Marketplace Technology

- Hidden Website (Tor Hidden Service, I2P)

- Customers
  - No cost of creation
  - No information needed
- Vendors
  - Vendor bonds required
  - Often invite only
  - Public feedback history



- Payments (Bitcoin)

- Marketplaces often act as escrow agent
- Escrow sometimes acts as a mixing service



- Hidden Messages(PGP)



# Market Transactions



“I’ll take the red pill”

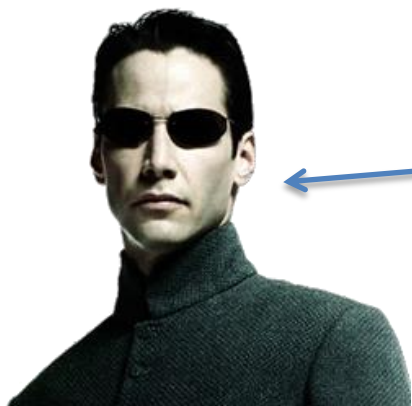




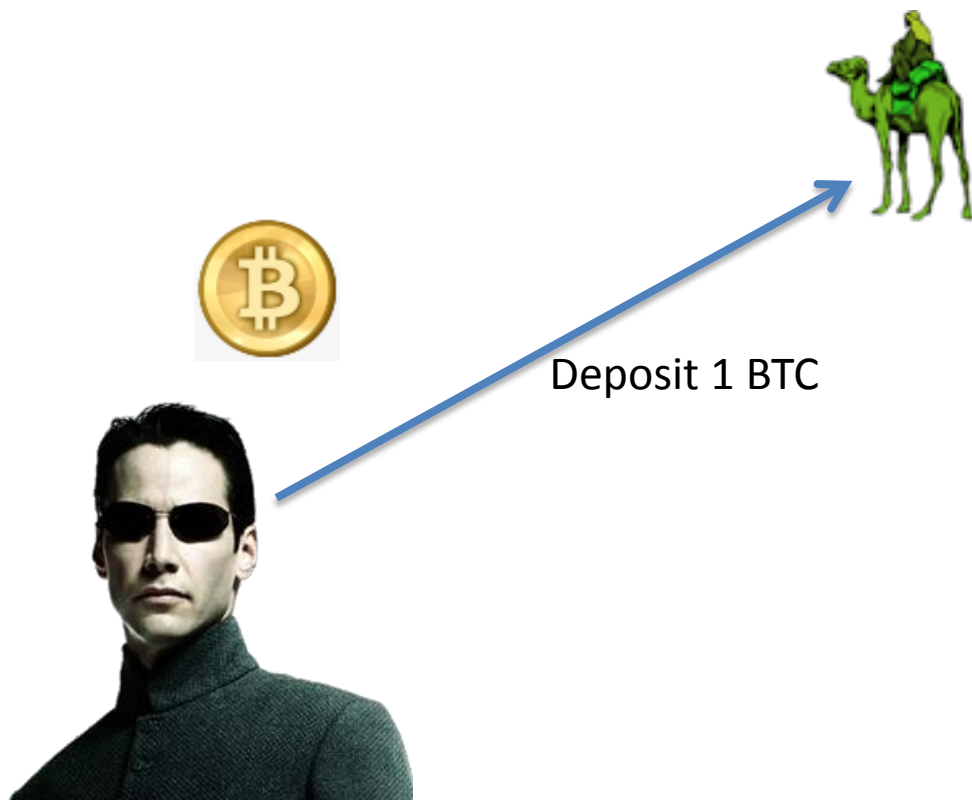
# Market Transactions



"1 BTC please"



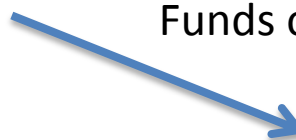
# Market Transactions



# Market Transactions



Funds ok



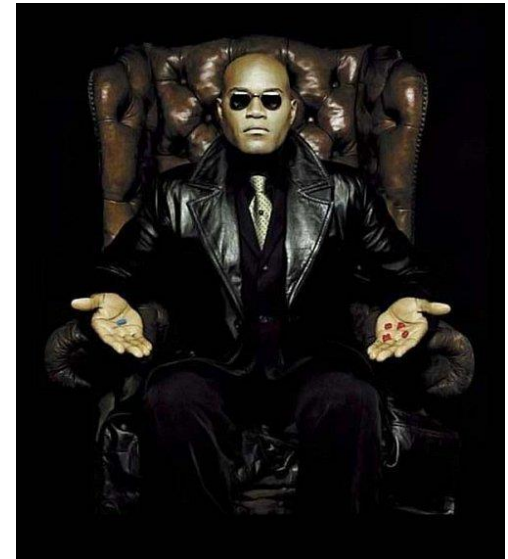
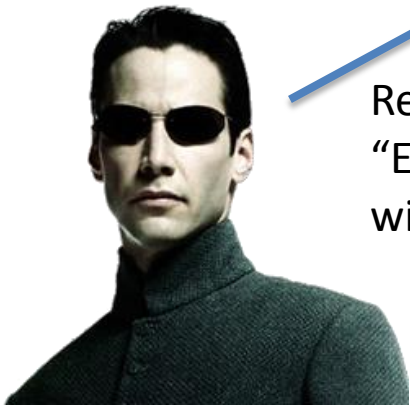
# Market Transactions



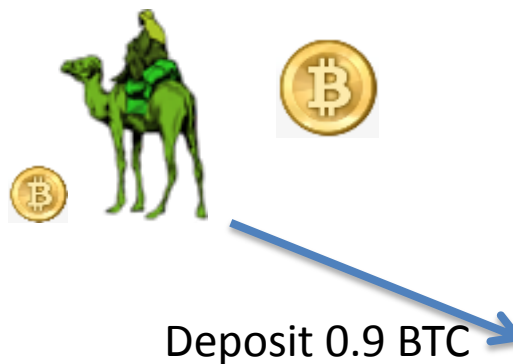
# Market Transactions



Received  
"Excellent seller, would do business  
with again. A++++"



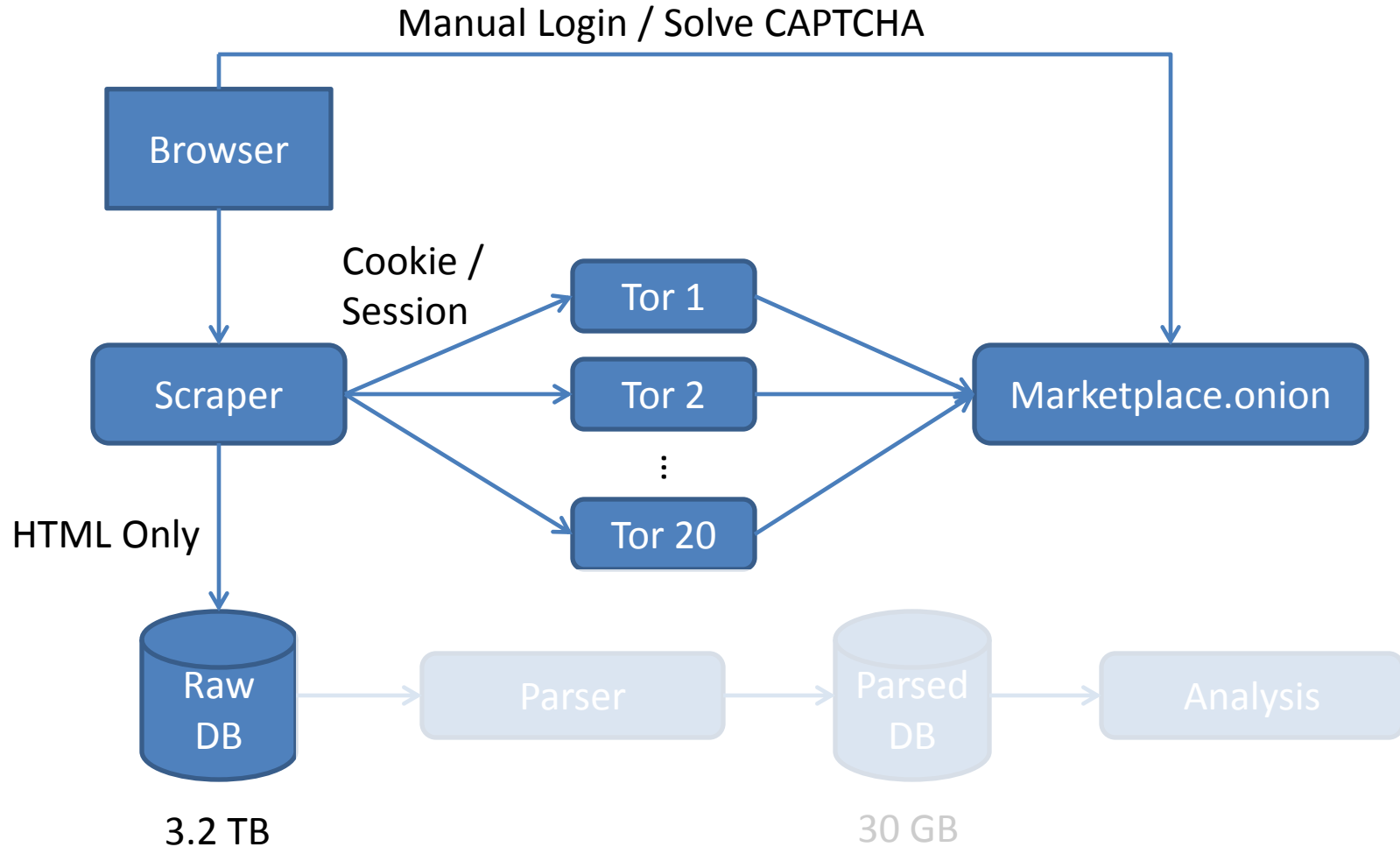
# Market Transactions



# Questions

- How much is being sold?
- What is being sold?
- How many vendors are relevant?
- What do vendors sell?

# Measurement Platform Overview





# Measurements

## ■ **Stealth**

- Indistinguishable from real user
- Random delays, scrape slowly
- Popular User Agent
- Browse website “normally”

## ■ **Complete and instantaneous**

- Dynamic marketplace, moving target
- Scrape quickly
- Site availability as low as 70%

# Measurements

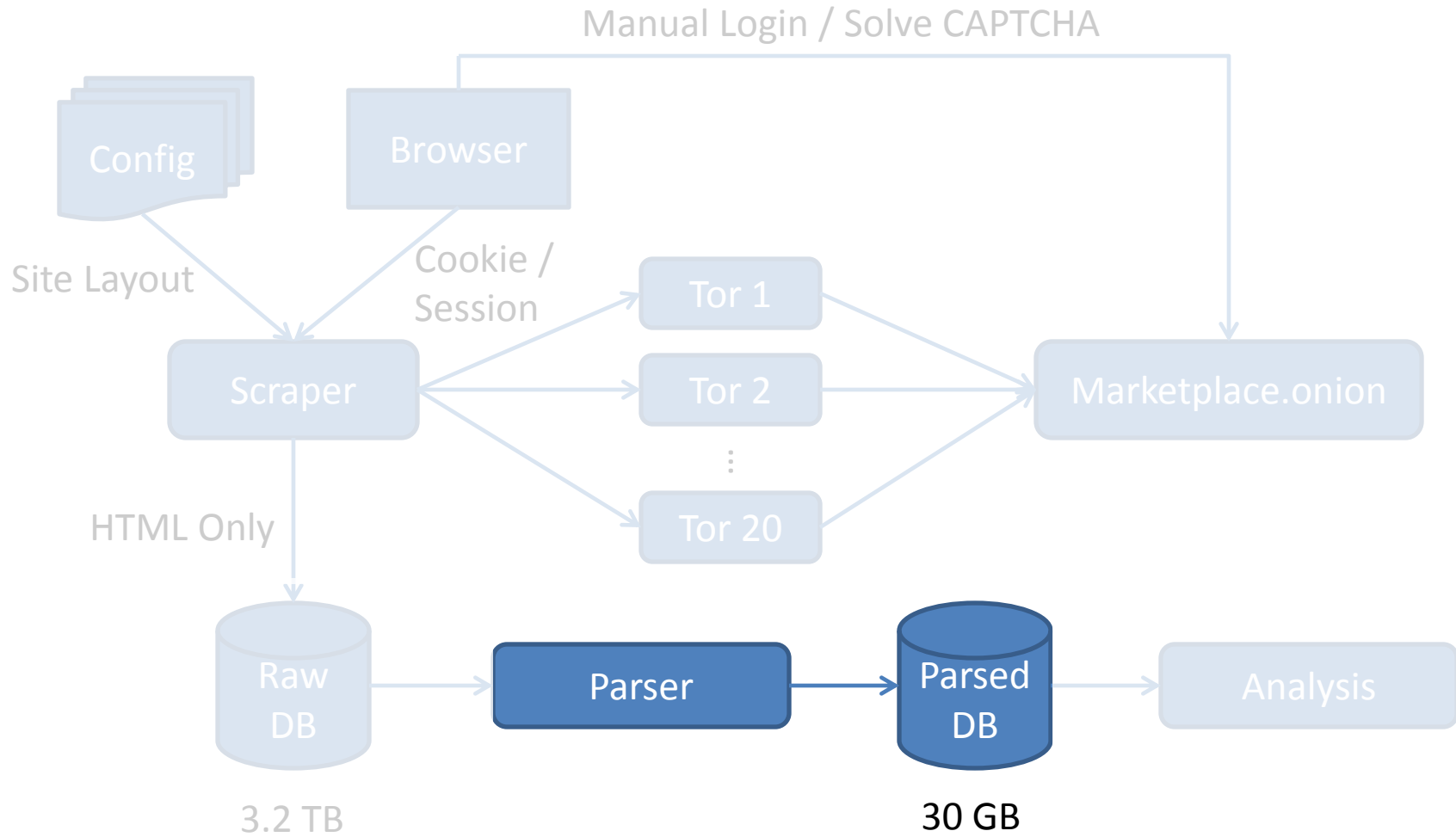
## ■ Anti-Scraping Encountered

- Rate Limits
- Cookie Timeout
- User Account Suspension

## ■ Totals

- 35 Marketplaces 1,908 scrapes total – 3.2 TB
- 27 – 331,691 pages per scrape
- 11/22/11 – present

# Parsing



# Silk Road Available Data

## Books

### Hacking for beginners

**Seller:**

**hackingtech (98)**

**Price:**

฿0.12

**Ships from:** undeclared

**Ships to:** Worldwide

**Description:**

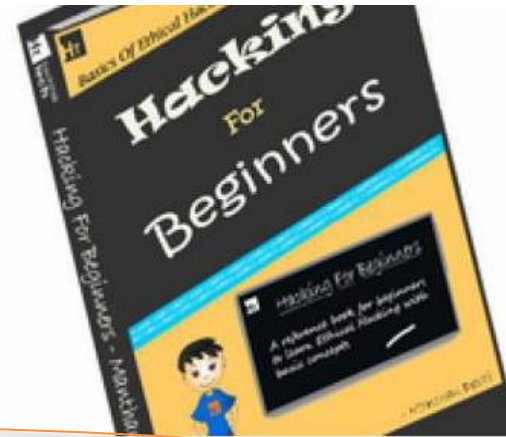
Hacking For Beginners is a reference book for beginners to learn ethical hacking for free and from basic level to clear all the fundamental concepts of ethical hacking.the book has been prepared by Hacking Tech ( www.hackingtech.co.tv ) website for the users benefit.so enjoy the book and site...

**add to cart**

**Recent feedback**

rating	feedback	freshness
5 of 5	Fast delivery	3 days
5 of 5	Thanks!	4 days
5 of 5	Leave feedback here	9 days
5 of 5	Leave feedback here	9 days
5 of 5	5 of 5	10 days

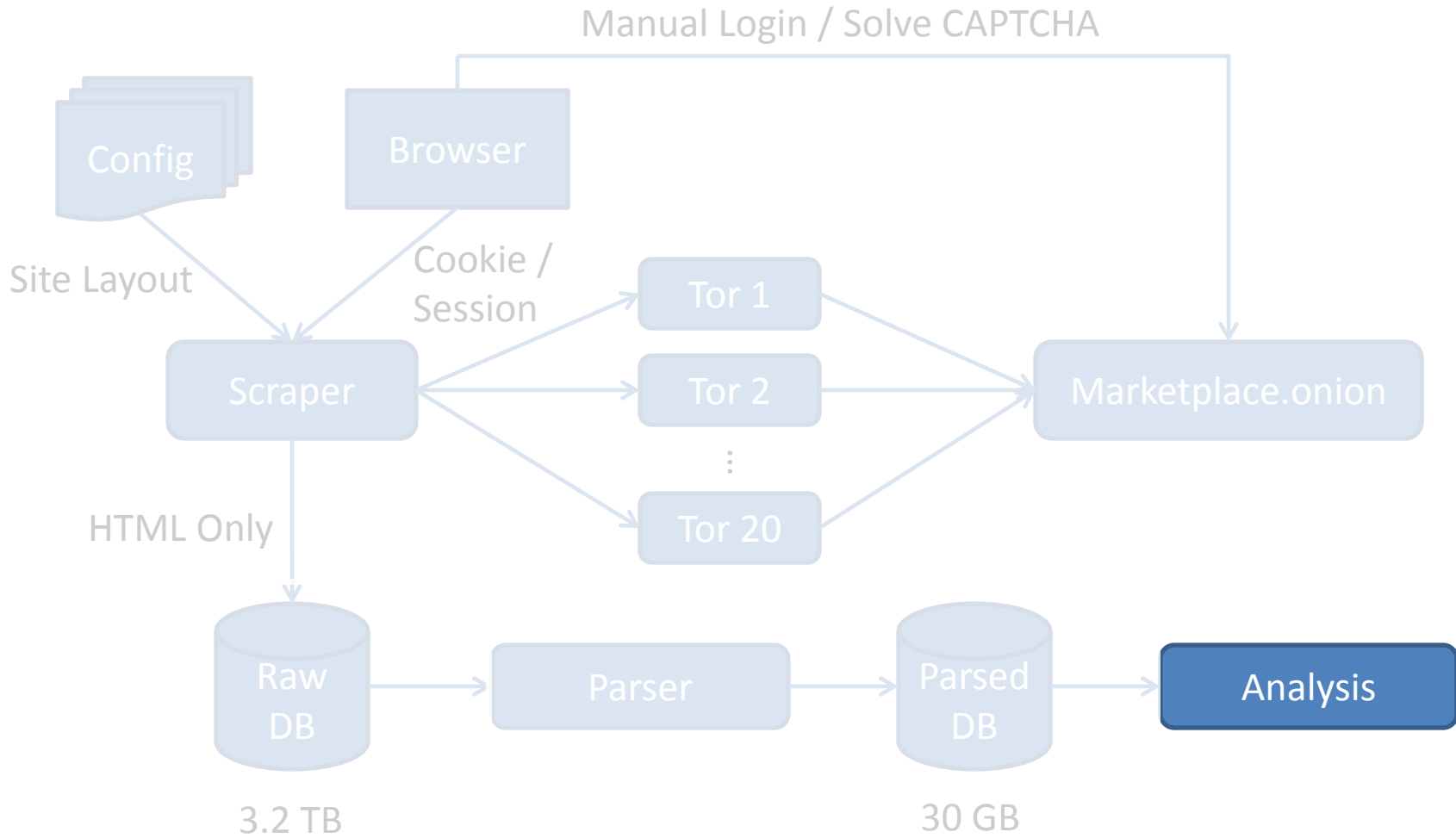
**Bookmark this item**



**Feedback is often mandatory!**

➔ Acceptable proxy for sales volume

# Analysis

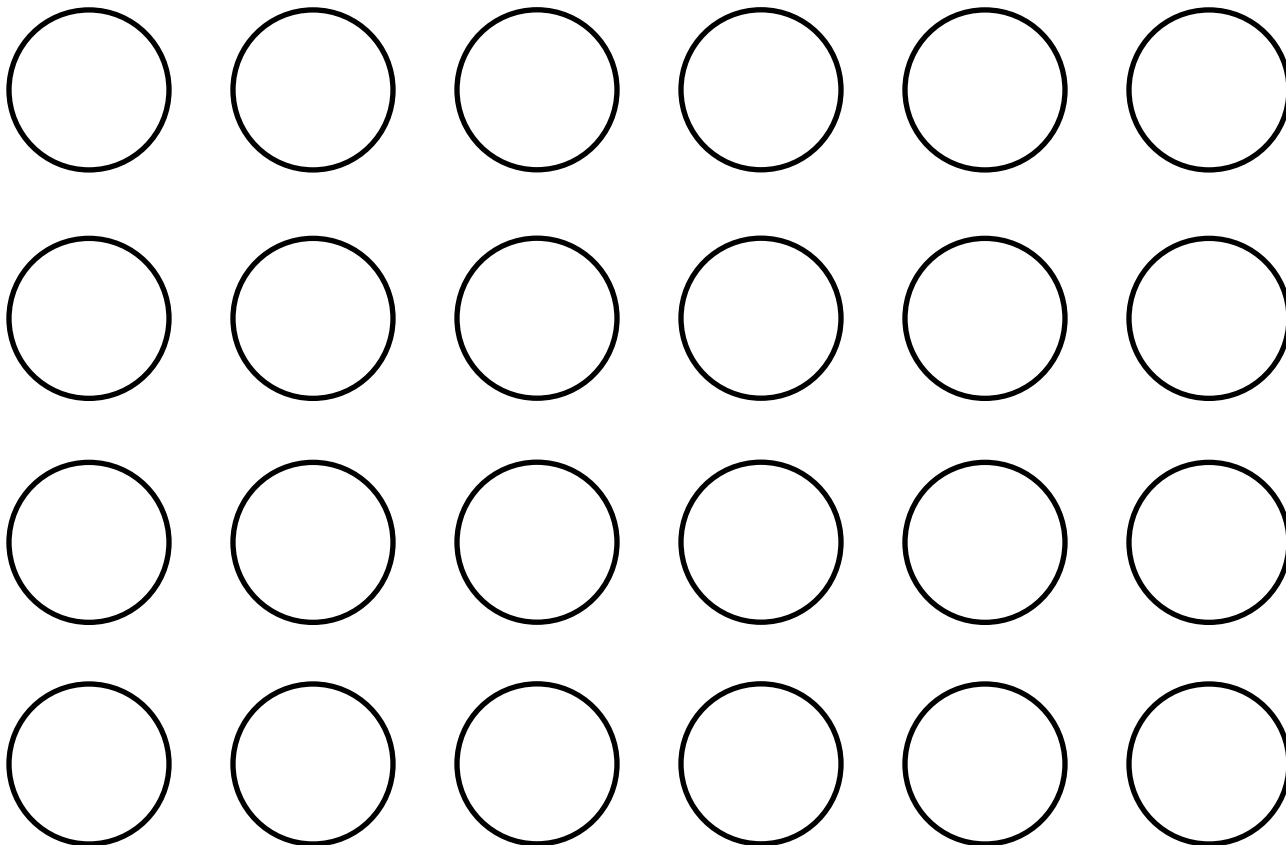


# Data Completeness

- **How complete is the data?**
  - Unreliable dynamic marketplaces that take days to scrape
  - Empirical observations - lower bound
  
- **Idea:** Estimate population via mark and recapture
  - Schnabel Estimator allows multiple recapture

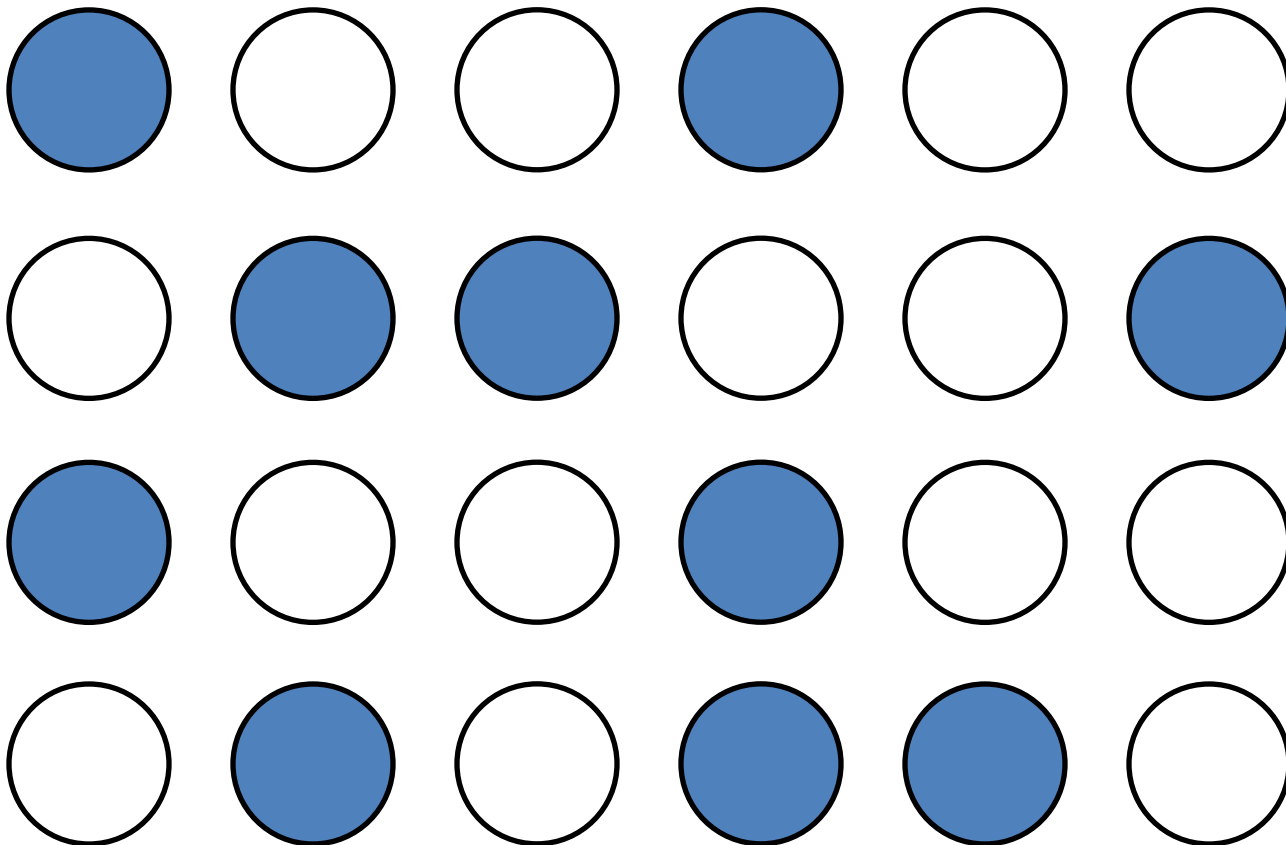
# Mark and Recapture

Population Size = 24



# Mark and Recapture

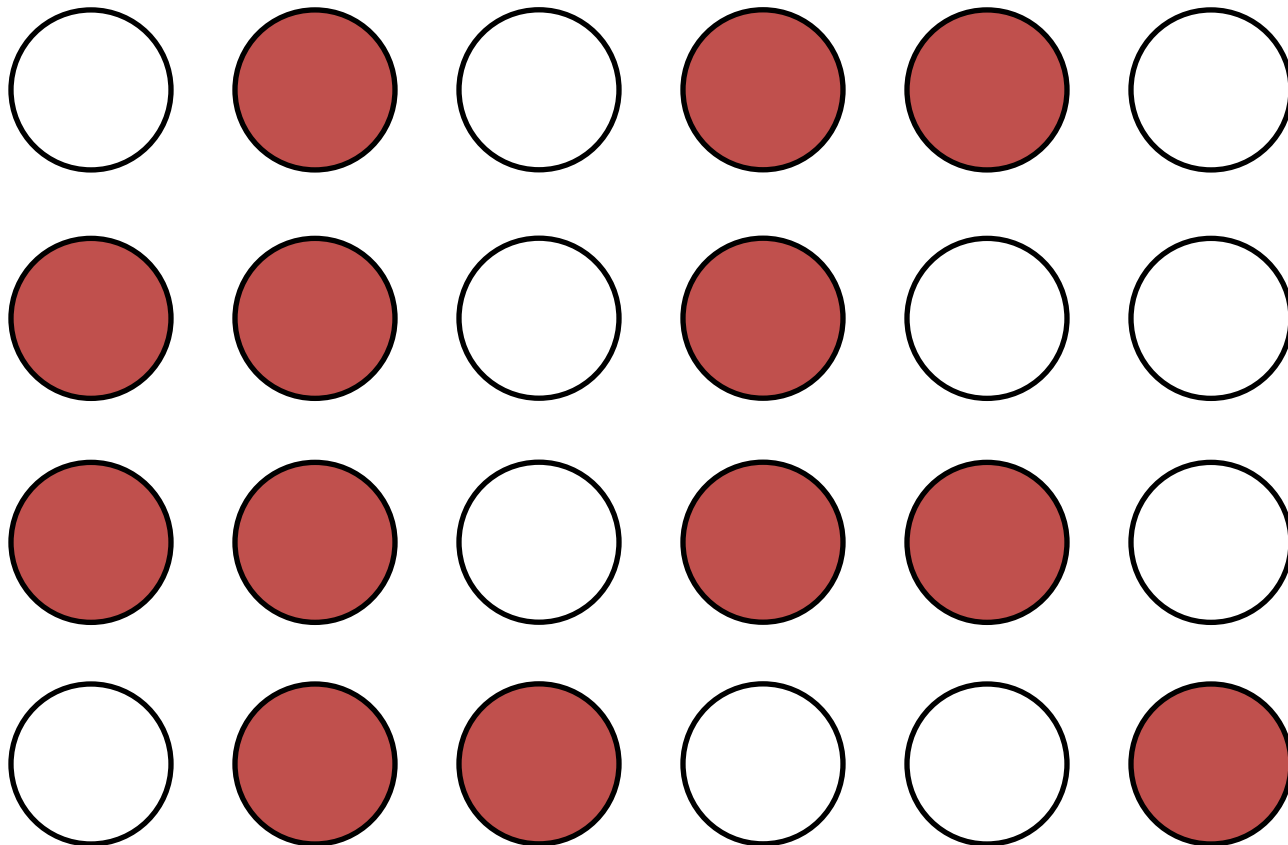
Sample Size = 10





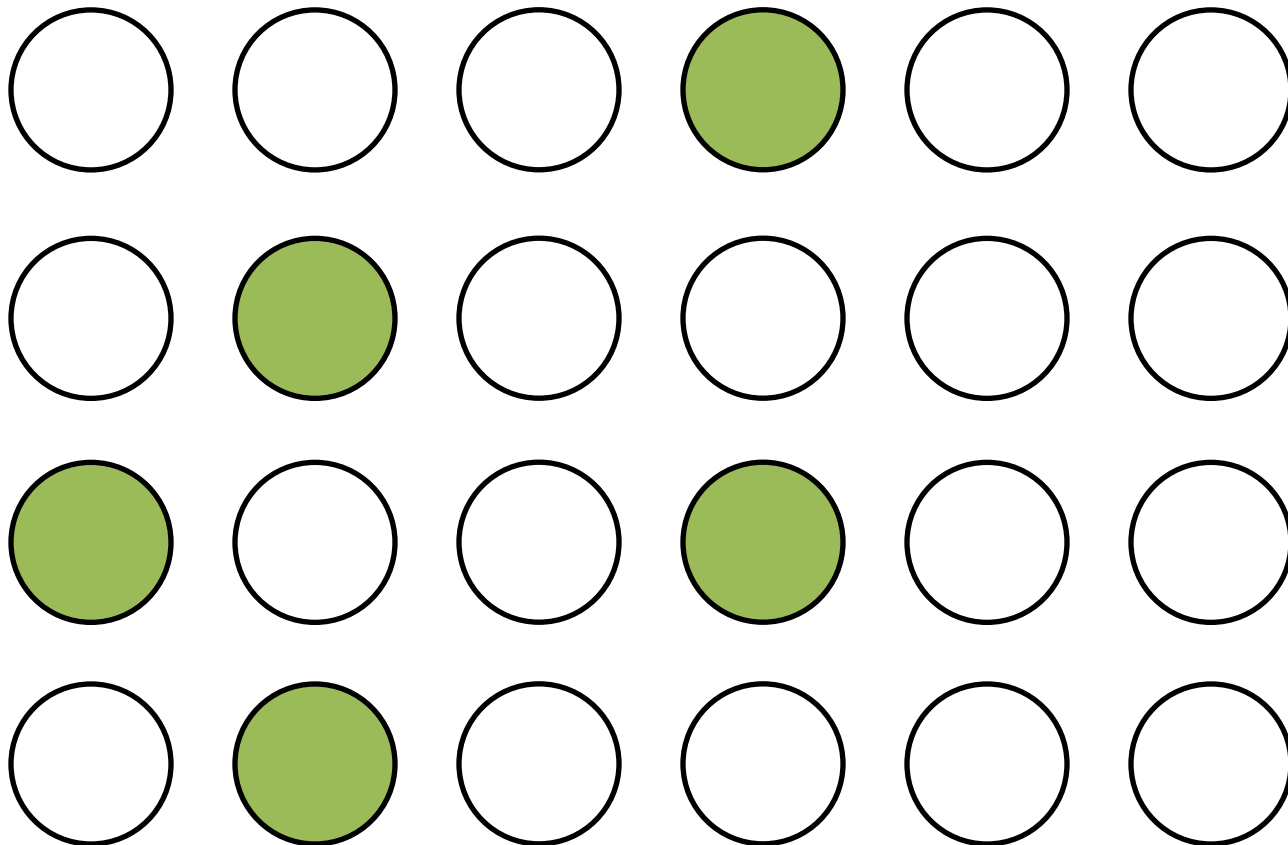
# Mark and Recapture

Sample Size = 13

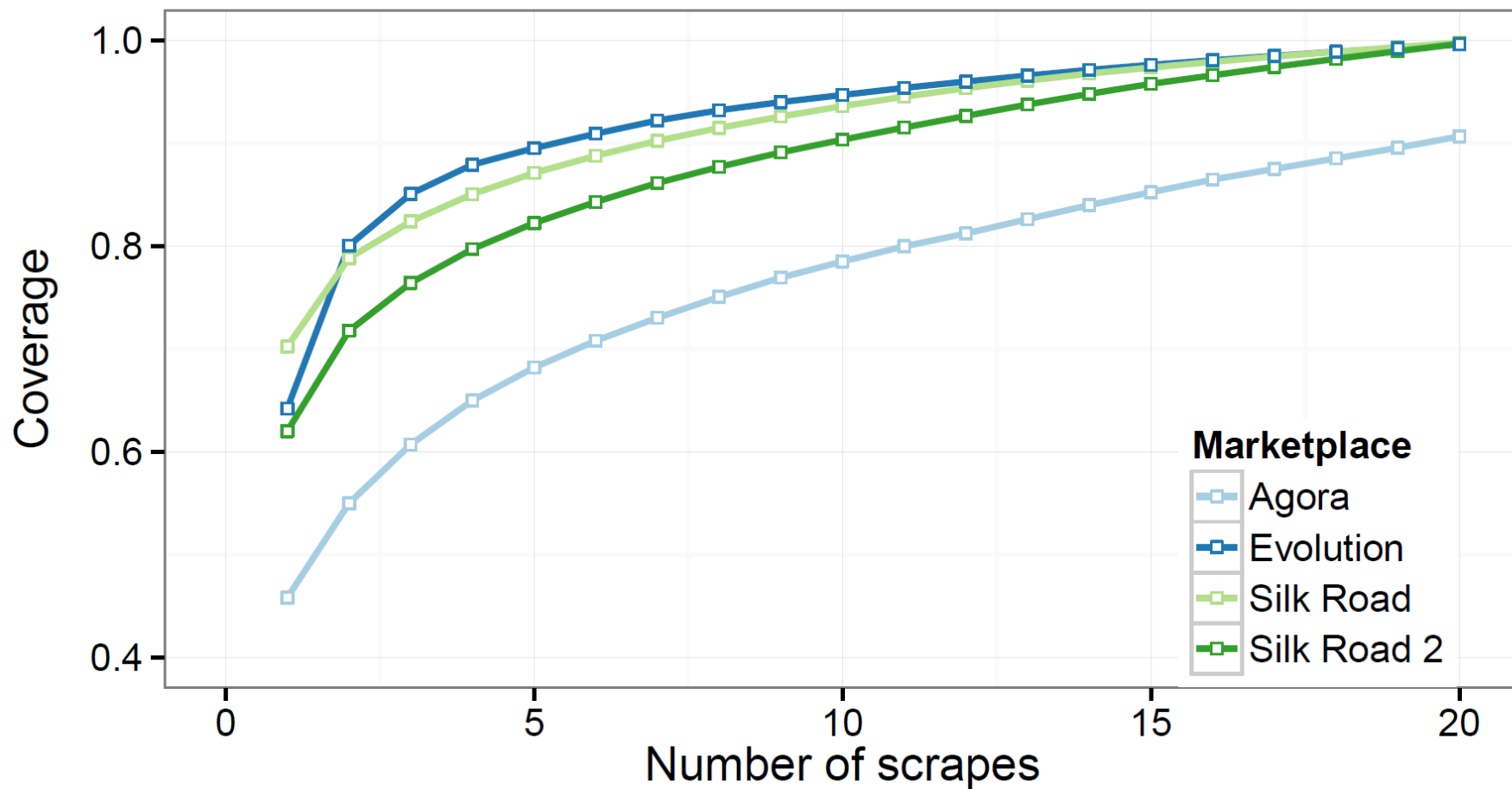


# Mark and Recapture

Overlap = 5, Population Estimate = 26



# Data Completeness



# Analysis

- **Assumption:** Each feedback corresponds to precisely one transaction
  - Anonymity requires strictly enforced feedback system to establish reputation
  - Possible on many marketplaces to purchase several quantities of item and leave 1 feedback, conservative estimate

# Alternative Transaction Proxies

## ■ Counting # Item Listings

- Very efficient and convenient
- Assumes that there exists some stable ratio between transaction volume and # listings
- Daily  $\frac{\text{volume}}{\# \text{ Listings}}$  for The Evolution Marketplace in July 2014 and September 2014 differ by factor of 4

# Uniqueness

## ■ Problem:

- 100s of observations of same feedback
- Double counting leads to over-estimations
- Feedback may be updated, deleted

## ■ Solution:

- Automatically detect updated feedbacks
  - Only keep most recent version
- Hash {timestamp, title, vendor, message, rating}

# Holding Prices

- Feedbacks are useful to vendors but are destroyed when the listing is removed
- Vendors raise listing prices prohibitively high



\$0.02 -> \$1,000.00



\$1,100.00 -> \$1,000,000.00

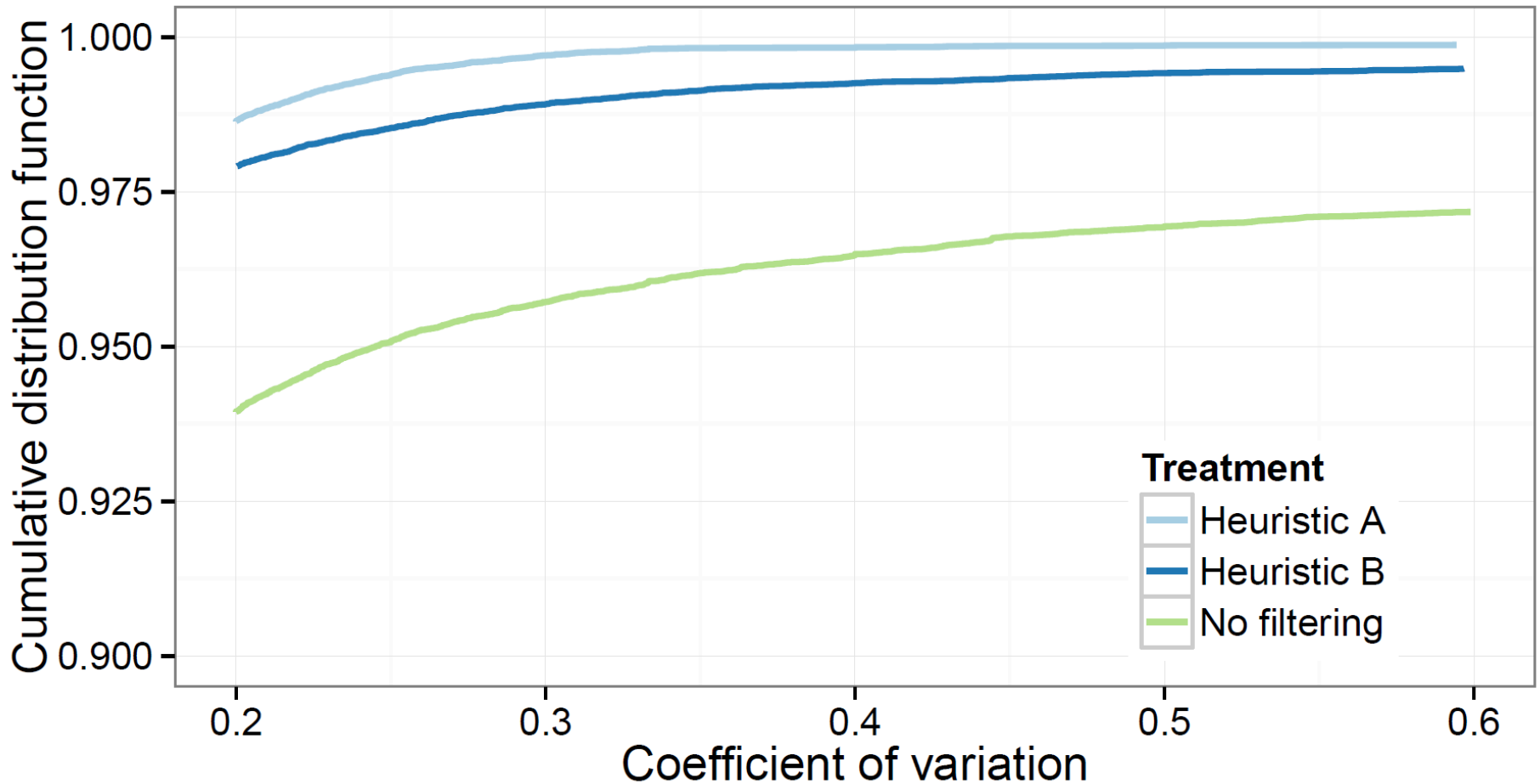
- Need to look at historical price for item

# Holding Prices

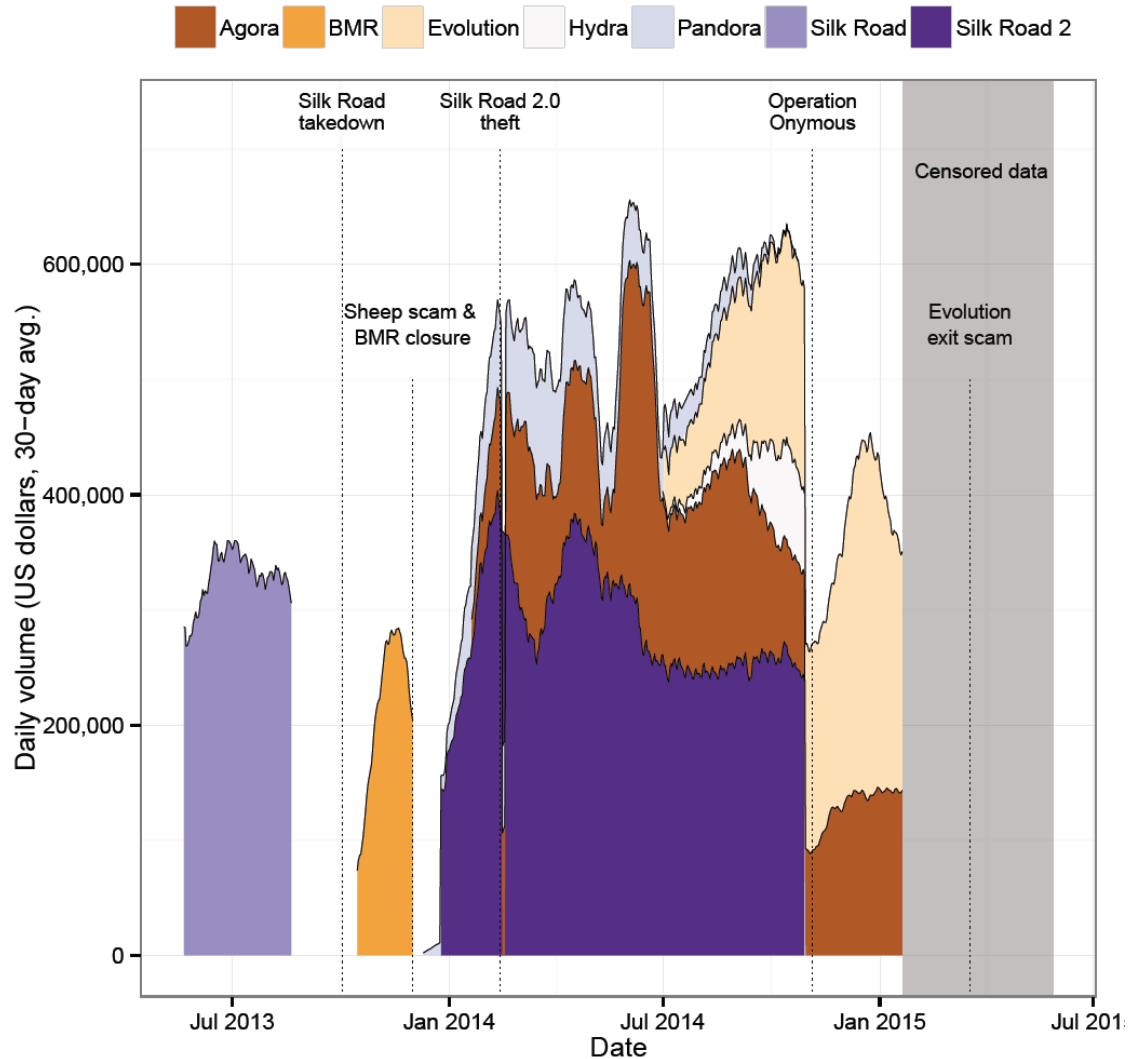
- Heuristic A:
  - Remove all free things
  - Remove all things > \$100,000
  - Calculate median of remaining prices
  - Remove everything greater than 5x median
  - Remove things less than 25% of median
- Heuristic B:
  - Remove all things > \$100,000
  - Remove upper quartile
  - Remove everything greater than 100x cheapest non-zero price
- Evaluation
  - Coefficient of Variation  $c_v = \frac{\sigma}{\mu}$



# Holding Prices CDF



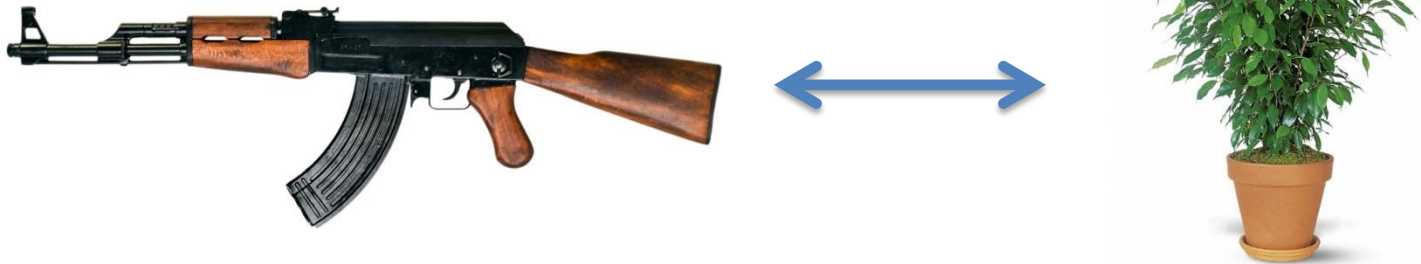
# Sales Volume



# Product Categories

## ■ What is being sold?

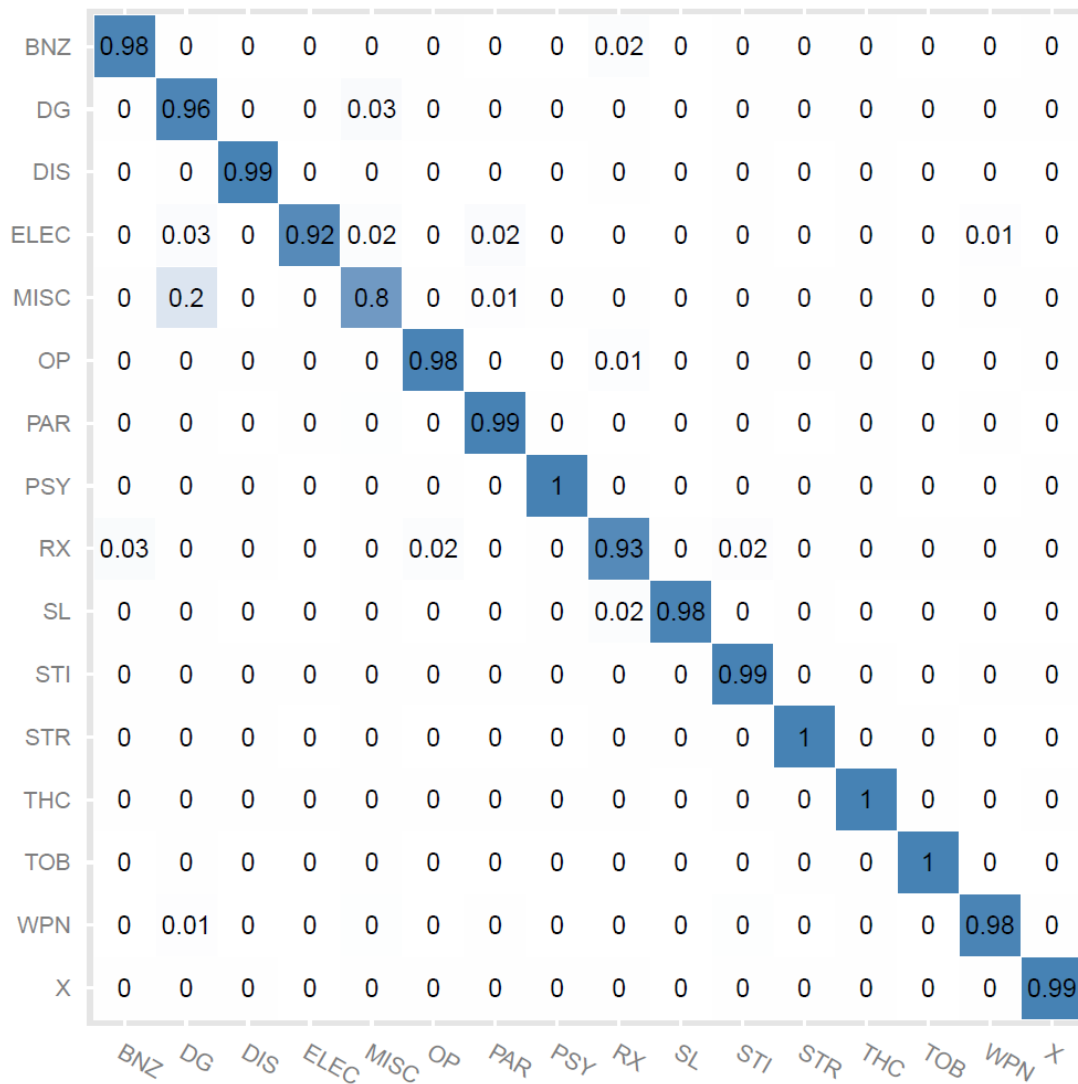
- Product labels are often unavailable or inaccurate



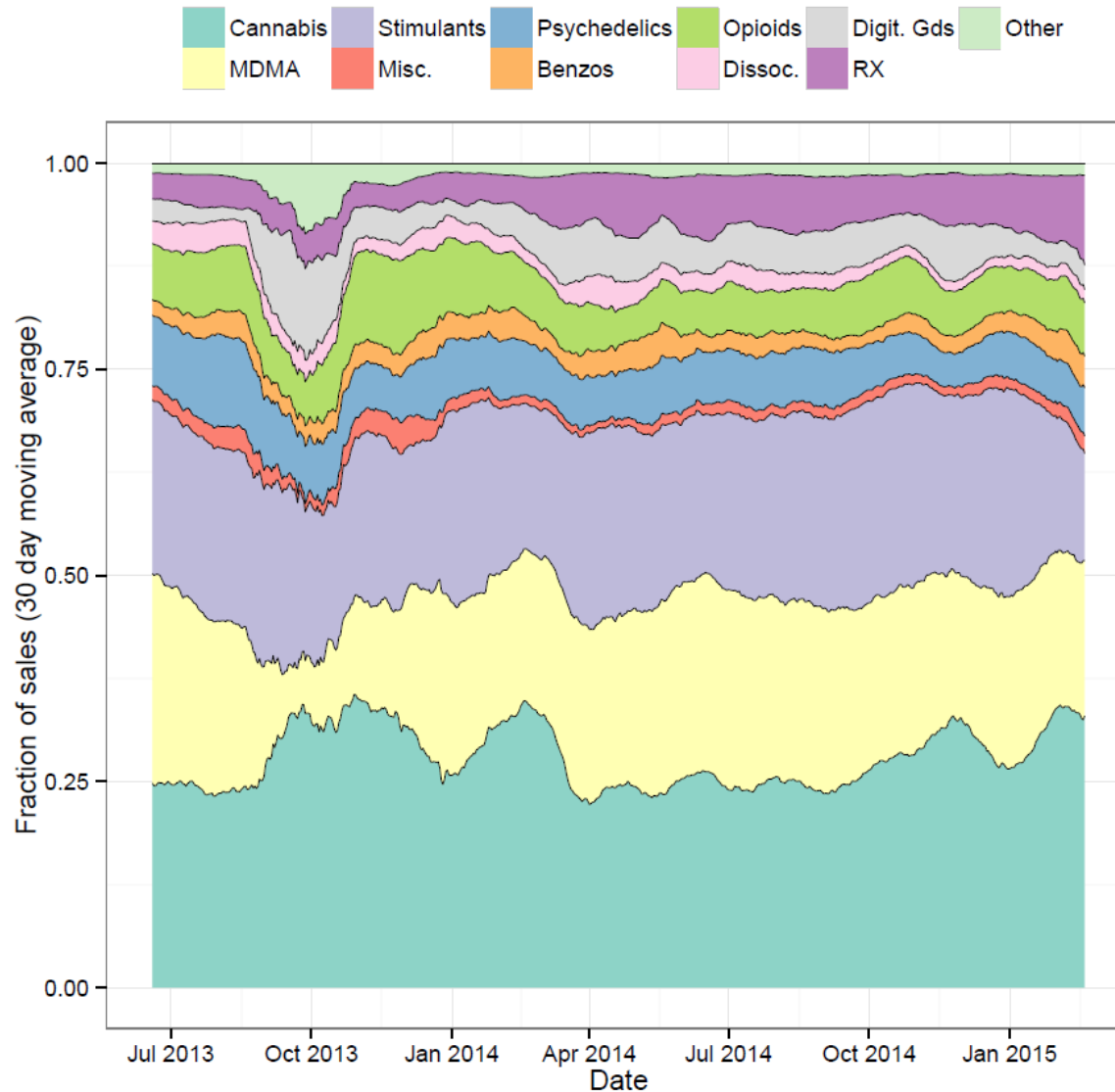
## ■ Classifier trained from Agora and The Evolution Marketplace

- Listing title and description concatenated and tfidf
- 1,941,538 unique samples, 162,198 words tokenized
- Predicts 16 class labels

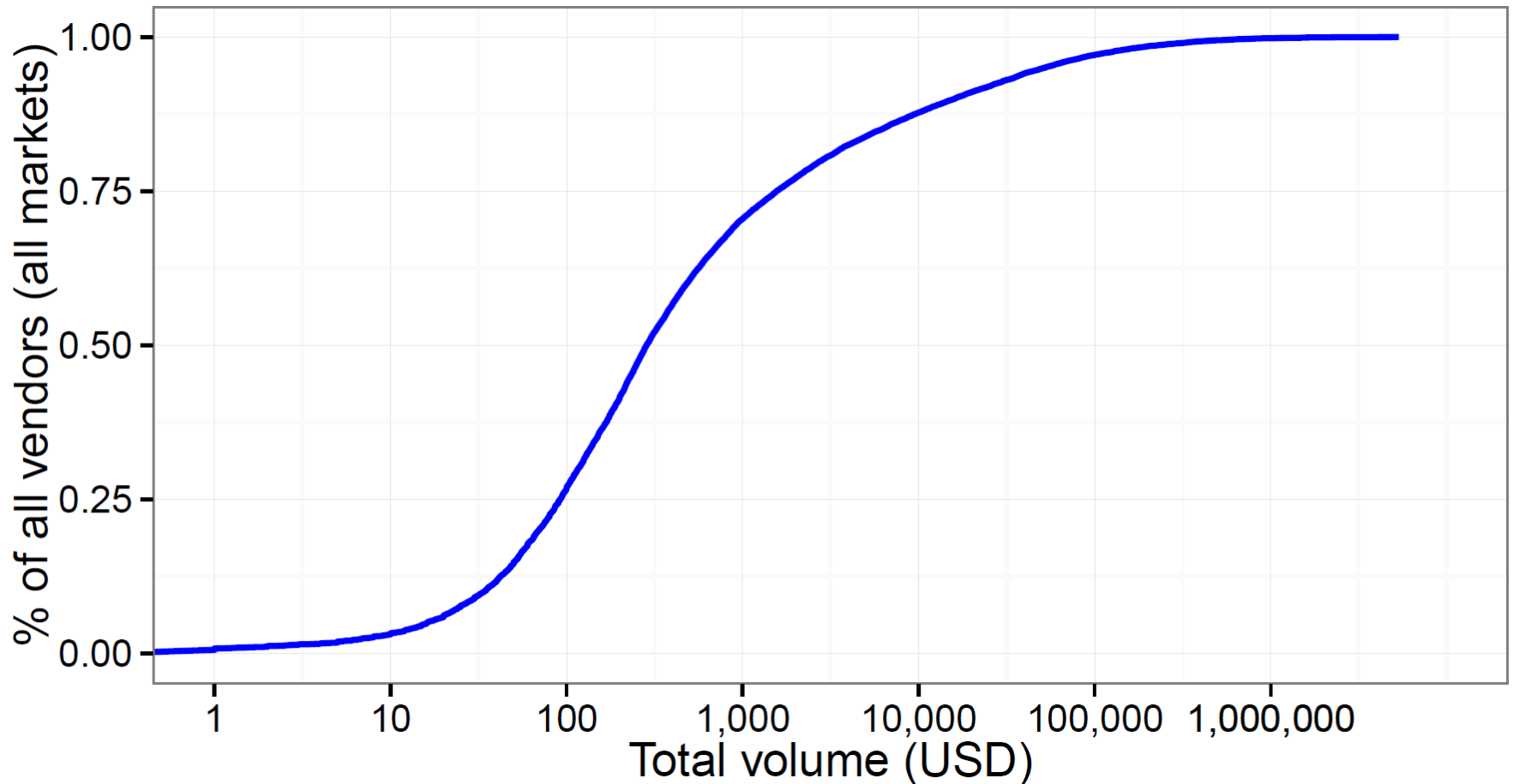
# Confusion Matrix



# Item Sales Per Category



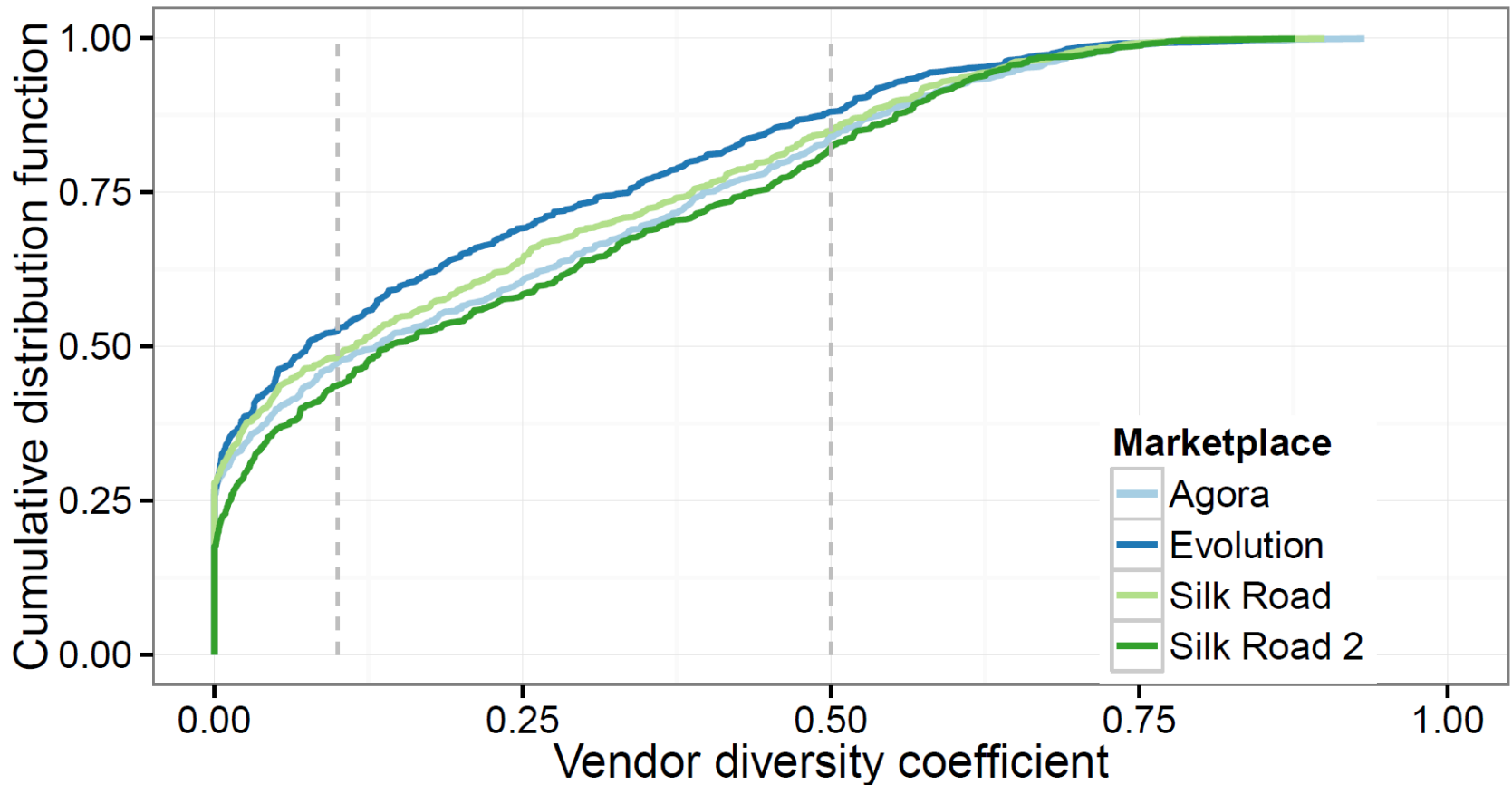
# Vendor Volumes CDF



# Vendor Diversity

- **Do vendors specialize in what they are selling?**
  - Do vendors sell what they make?
  - Does a single online presence sell goods for several diversified suppliers?
  
- **Coefficient of Diversity  $\in [0, 1]$** 
  - 0 – all sales from same category
  - 1 – equal sales from each category
  - Only vendors  $> \$10,000$  total sales considered

# Vendor Diversity CDF





# Validation

- Trial evidence GX226A, GX227C places Silk Road 1 weekly volumes at \$475,000/week in late March 2012, consistent with our estimates
- Administrator reports Silk Road 2 daily volumes of around \$250,000 in September 2014, similar to our estimated \$270,000
- Leaked Agora vendor page shows sales total on June 5, 2014 to be \$3,460, our observations yielded \$3,408

# Takeaways

- Anonymous Marketplaces are very easy to setup and use and have wide customer appeal
- Anonymous Marketplace ecosystem transacts in excess of \$500,000 / day
- Anonymous Marketplaces are primarily used (~75%) for recreational drugs
- Anonymous Marketplace ecosystem has historically recovered from takedown efforts and scams
- Anonymous Marketplaces are controlled by small set of highly influential vendors

Kyle Soska – [ksoska@cmu.edu](mailto:ksoska@cmu.edu)

# Data Completeness - Schnabel Estimator

- $F$  true feedbacks at time  $t$
- $n$  observations
- $C_i$  feedbacks in observation  $i$
- $M_i$  feedbacks in observation  $i$  previously seen
- $R_i$  total previously observed feedbacks

- Estimate  $\hat{F} = \frac{\sum_{t=1}^n C_t M_t}{\sum_{t=1}^n R_t}$

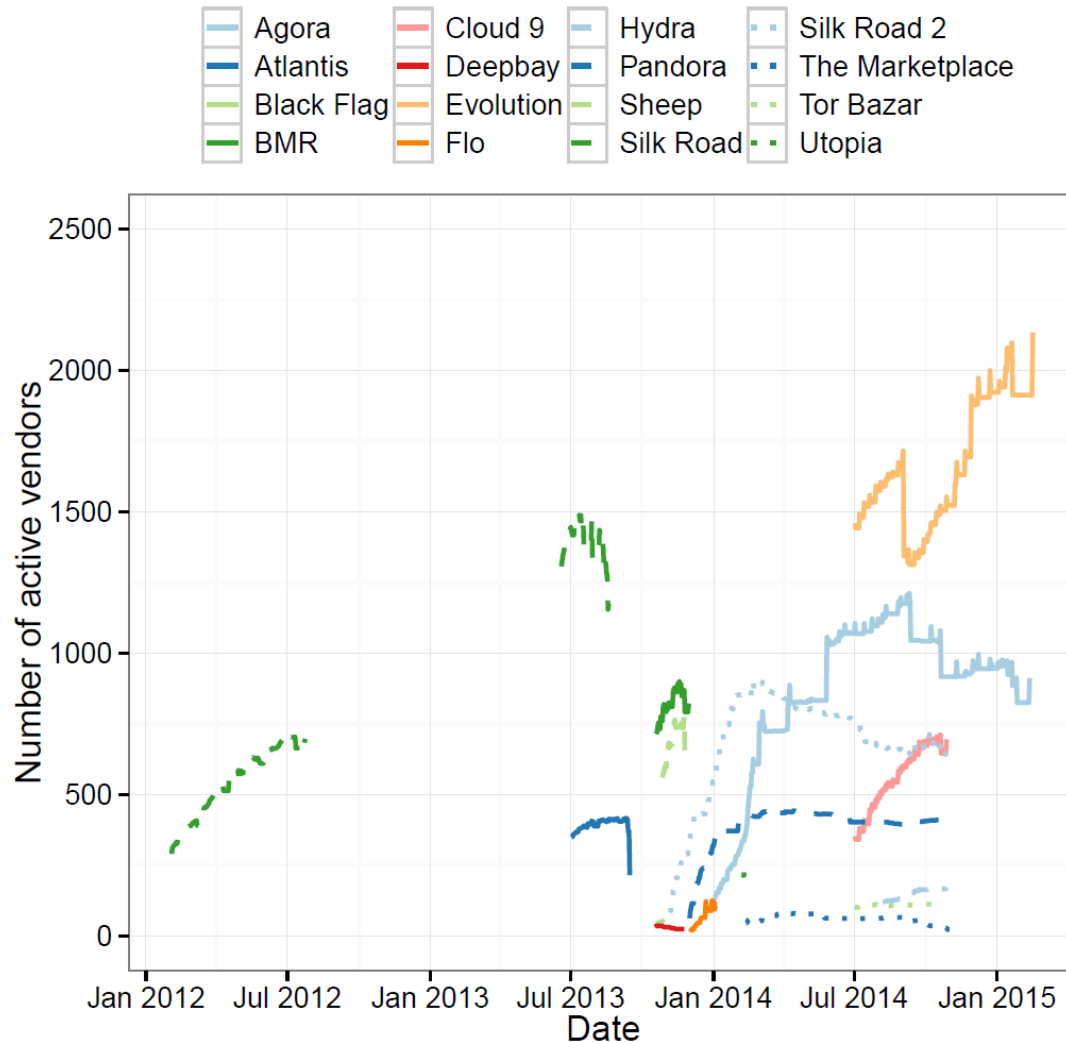
# Vendor Diversity

- $C_{S_j,i}$  = % of vendor  $j$ 's total sales that came from category  $i$

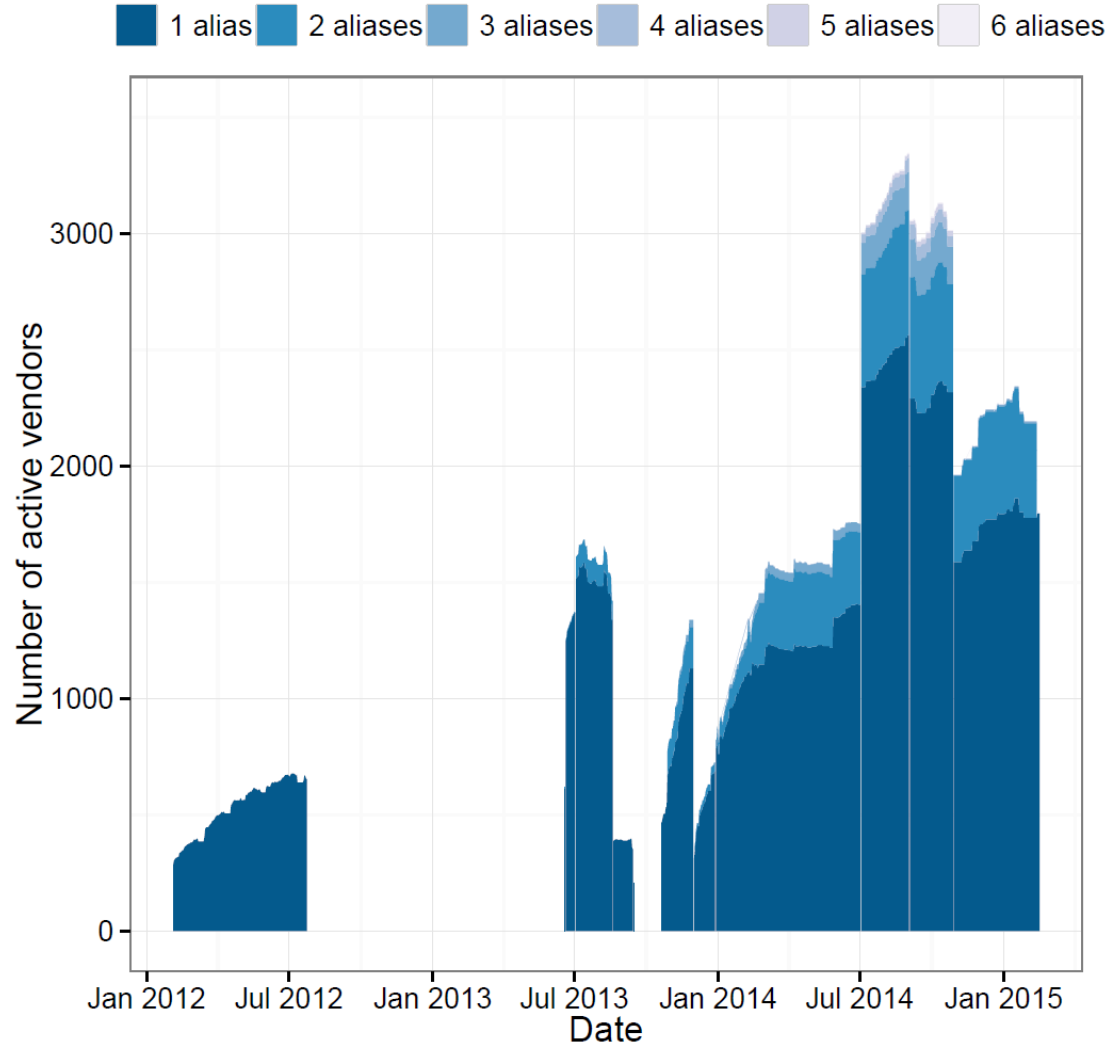
- Coefficient of Diversity =

$$c_d = \left(1 - \max(C_{S_j})\right) \frac{|C_{S_j}|}{|C_{S_j}| - 1}$$

# Active Sellers Over Time



# Aliases Per Sender



# PGP Deployment

