

# The Dark Web

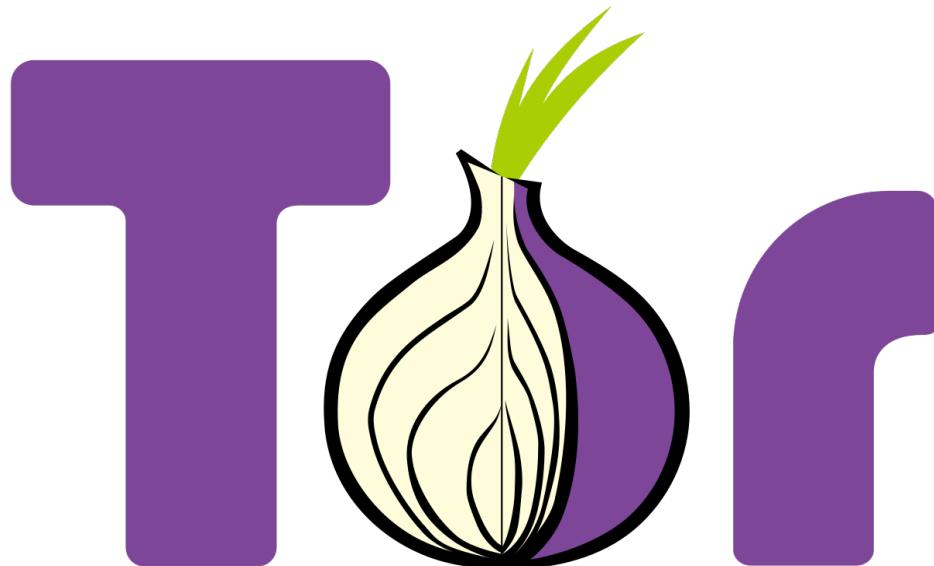
# The Dark Web

- Hollywood treats the term pretty casually...
- Generically, the part of the Web you can't access using standard browsing
- More specifically, what you can access using the Tor system



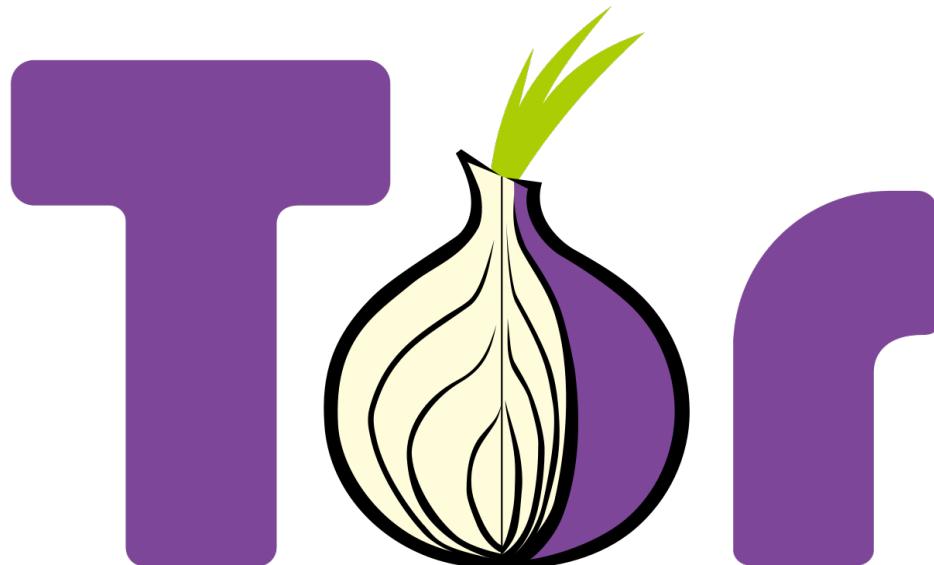
## Part I: Tor

- The Onion Router
- Used to fight traffic analysis
- (Even if you use HTTPS, your ISP or other eavesdroppers can still use traffic analysis)
- Why isn't a VPN enough?



# Part I: Tor

- **The Onion Router**
- Used to fight traffic analysis
- (Even if you use HTTPS, your ISP or other eavesdroppers can still use traffic analysis)
- Why isn't a VPN enough?
- What could you find out without the payload?





## How Tor Works: 1

Tor node  
- - -> unencrypted link  
-> encrypted link

Alice



Step 1: Alice's Tor client obtains a list of Tor nodes from a directory server.



Jane

Dave



Bob



## How Tor Works: 2

Tor node  
---> unencrypted link  
-> encrypted link

Alice



Step 2: Alice's Tor client picks a random path to destination server. **Green links** are encrypted, **red links** are in the clear.



Jane



Bob

Dave

## Ef How Tor Works: 3

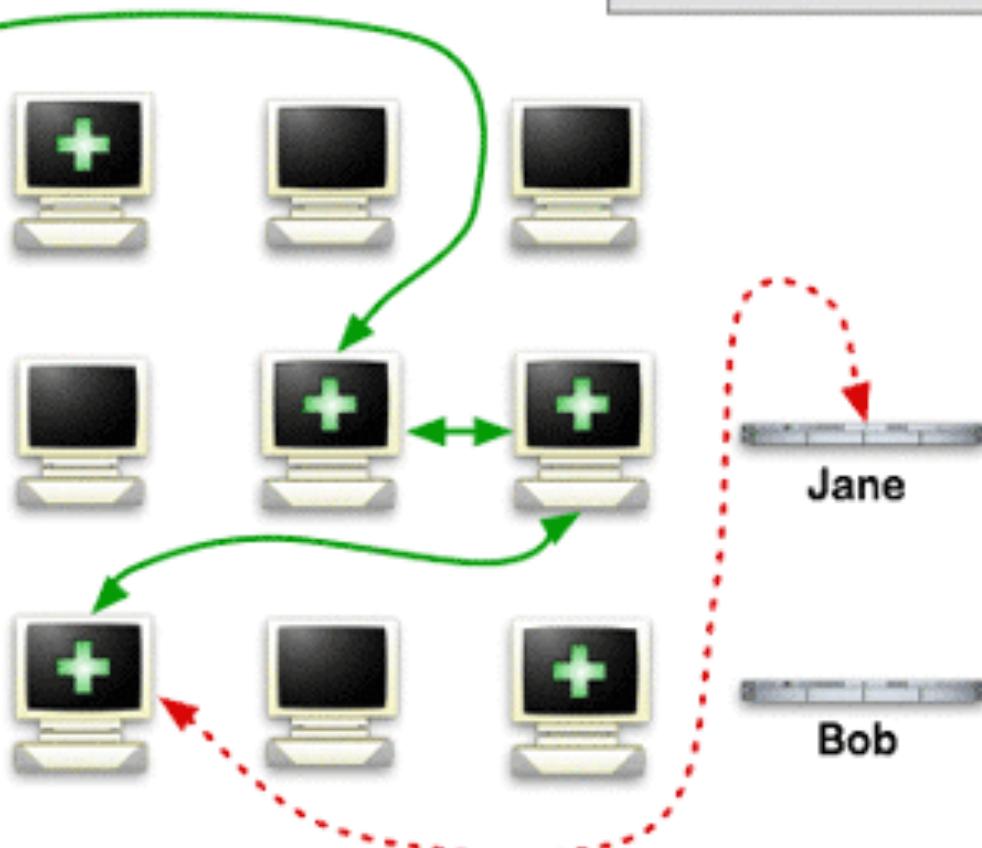
 Tor node  
---> unencrypted link  
-> encrypted link

Alice



Step 3: If at a later time, the user visits another site, Alice's tor client selects a second random path. Again, green links are encrypted, red links are in the clear.

Dave



# Why The Onion Router?

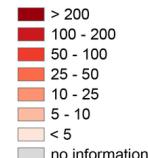
- Client encrypts outgoing data multiple times, like layers of the onion
- Each layer of encryption is peeled off by a relay node in the network.  
At each layer:
  - Decrypt the current layer
  - Forward to the next destination
- Only the first relay node knows the source IP address
- Only the last relay node knows the destination IP address
- To break anonymity, you need to surveil ALL nodes in the Tor circuit

# Tor History

- Tor is a nonprofit based in Massachusetts
- One of its major funders is the US Government
- Why???

## The anonymous Internet

Daily Tor users per 100,000 Internet users

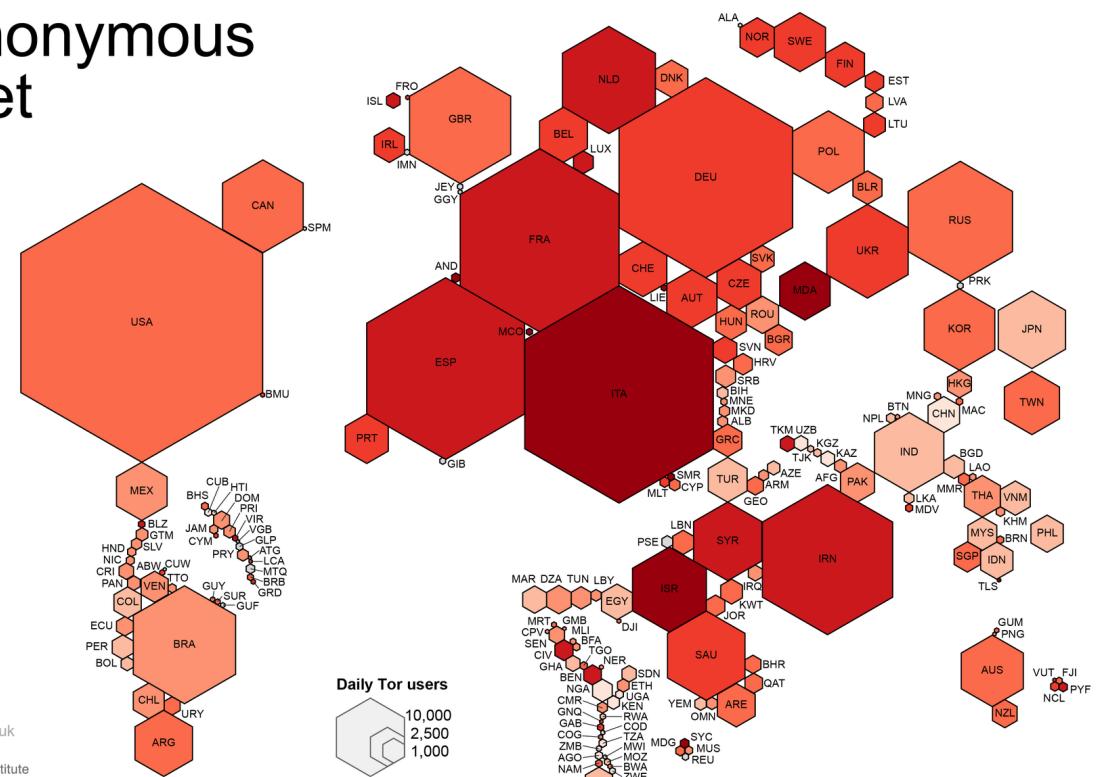


Average number of Tor users per day calculated between August 2012 and July 2013

data sources:  
Tor Metrics Portal  
metrics.torproject.org  
World Bank  
data.worldbank.org

by Mark Graham  
(@geoplace) and  
Stefano De Sabbata  
(@maps4thought)  
Internet Geographies at  
the Oxford Internet Institute  
2014 • geography.oi.ox.ac.uk

Oxford Internet Institute  
University of Oxford



## Part II: Information Extraction and Crime

- This next section covers my own research that applies Web data and machine learning to fighting online crime
- These slides aren't graphic, but they discuss a serious and distressing topic
- You will not be examined on this material: you're free to skip at any point if you

## Part II: Information Extraction and Crime

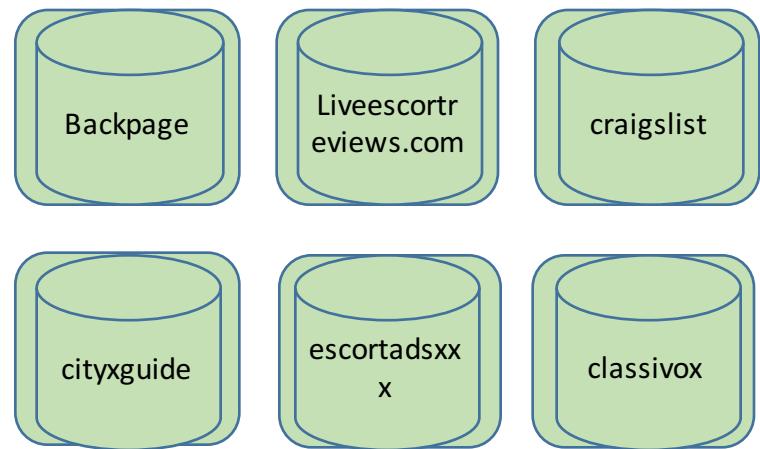
- Human trafficking is odious crime that obtains labor via physical, financial coercion
- Traffickers believed to channel many victims into sex work
- Trafficking phenomena very opaque
  - Unclear which sex workers are in fact trafficking victims
  - Unclear which sex work organizers engage in trafficking
  - Law enforcement lacks any good intelligence
  - Surprisingly old problem



Trafficking image, US Department of Justice

# Human Trafficking and Web Data

- Sex work has gone online
- >120M advertisements posted on dozens of websites
- Advertisements contain:
  - Contact info
  - Personal details
  - Location
  - Price
  - Transaction details
  - Lots of other data
- Grim, but potential analytical gold mine
- Unfortunately, few or no H/T labels



# **!!!! SWEET ASHLEY !!!! for a LIMITED TIME Only !!!!! - 27**

Posted: Thursday, September 12, 2013 12:03 AM

---

[Reply](#)

**Hello i'm Ashley and I am 100% real and would Love to make your dreams come true**

Pic's are 100% real and current, cant wait to meet you.

So call [REDACTED] for the time of your life, Outcalls only

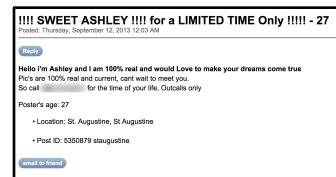
Poster's age: 27

- Location: St. Augustine, St Augustine
- Post ID: 5350879 staugustine

[email to friend](#)

# Our Goals

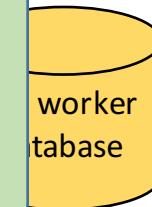
1. Transform 120M raw ads into high-quality database



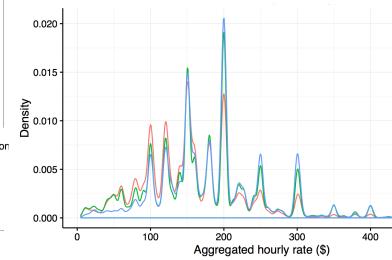
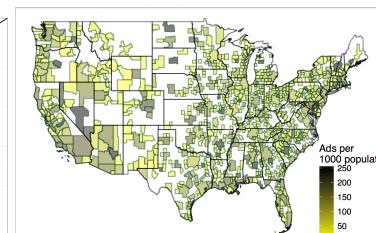
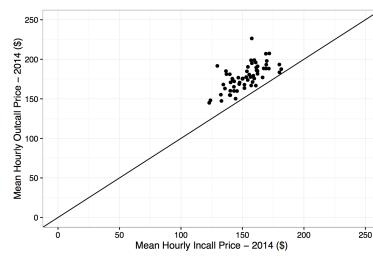
Name	Age	Location	MSA
Ashley	27	St Augustine	Jacksonville

2. Build database-backed enforcement tools to

*Novel social science and tools made possible by computational methods*



3. Long term plan: Use database for novel social science findings, further help law enforcement



# Agenda

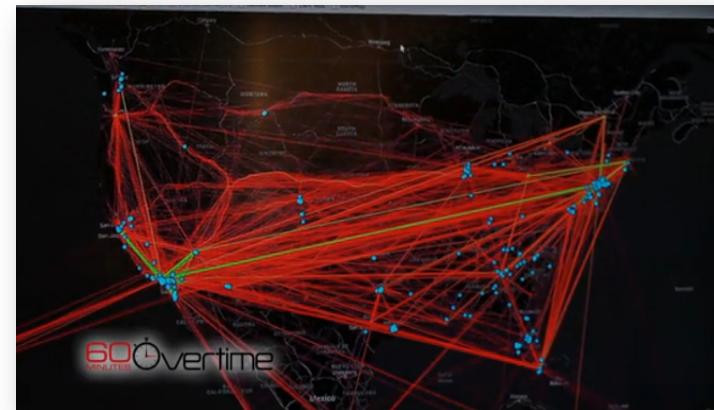
- Introduction
- Information Extraction
- Law Enforcement Tool: Evidently LE
- Analyzing Sex Market Pricing
- Conclusion

# Agenda

- Introduction
- Information Extraction
- Law Enforcement Tool: Evidently LE
- Analyzing Sex Market Pricing
- Conclusion

# Tools for Identifying Trafficking Victims with Data

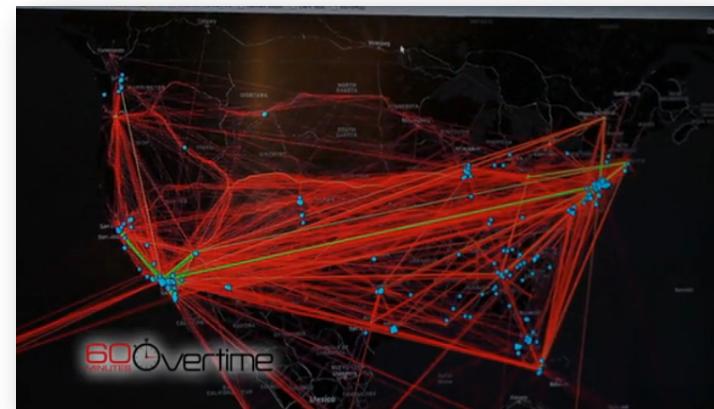
- Is sex worker:
  - Offering surprisingly low price?
  - Willing to engage in risky services?
  - Moving unusually frequently?
  - Sharing contact info with suspected h/t victim?
- Does advertisement:
  - Use numerically-nearby phone numbers?
  - Reuse ad text from a suspected h/t victim?
  - Describe someone with possible warning signs (physical abuse, drug abuse, fear of law enforcement, etc)?



60 Minutes Overtime, CBS.com

# Insight into Illicit Sex Worker Markets

- Better insight into market behavior can help better identify outliers
- What factors drive sex worker pricing decisions?
- How do opportunities in licit economy impact sex market pricing?



60 Minutes Overtime, CBS.com

# Project Background

- Project began with DARPA Memex in 2014
- Based on DeepDive project from Chris Re at Stanford
- Work done in collaboration with:
  - Christopher Re, Ce Zhang, Jaeho Shin @Stanford University
  - Feng Niu @Lattice Data, Inc.
  - Jacob Shapiro @Princeton University
  - Jeff Borowitz, Greg DeAngelo @Giant Oak, Inc.



# Agenda

- Introduction
- Information Extraction
- Law Enforcement Tool: Evidently LE
- Analyzing Sex Market Pricing
- Conclusion

## HT Advertising Table

Ad Identifier	Date Posted	(Craigslist) Region	...	Ad Text
10001	08/12/15	Wash		...
10002	08/15/15	NYC		...
10003	08/16/15	LA		...

## Aspirational Table

Ad Identifier	Price/Hour	Place(s)	Incall/Outcall	Phone	Email
10001	100	Bethesda	Incall	555...	m@...
10002	225	Nyack	Outcall	555...	c@...
10003	...	...	...	...	...

## Is the number of sex service ads increasing or decreasing?

# What large region has the most ads?

## What is the typical price for offered services?

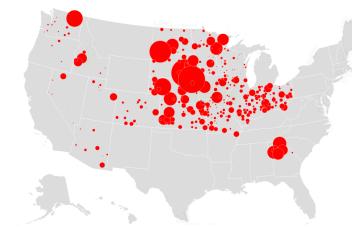
## Which providers are posting multiple times?

# Which exact locality is growing in popularity?

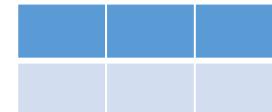
# Faceted Search

Name	Total Loss Paid	Number of Claims
kevin o'connell	1028765.95	19
roger moore	263975.02	17
mark mcmahon	58107	13
marc glanter	79337.02	6
chad krebs	217278.42	5
mark koss	205115.14	5
jeremy brantingham	98276.41	5
lisa taylor	81934.8	5
michael hupy	76577.89	5
paul aamodt	29014.79	5

# Data Analytics

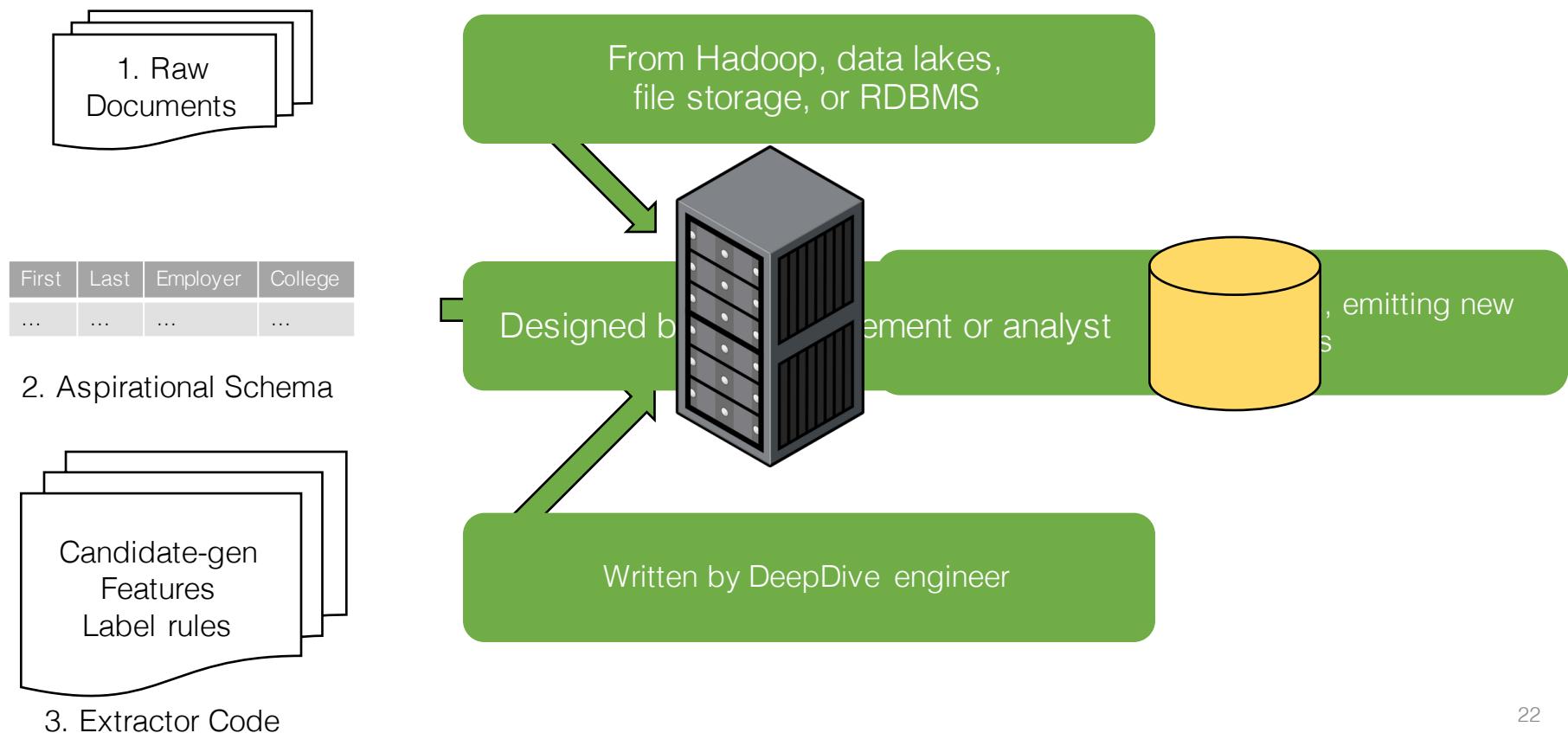


## Visualizations



## Other Tools

# Running Information Extraction



# DeepDive Internals

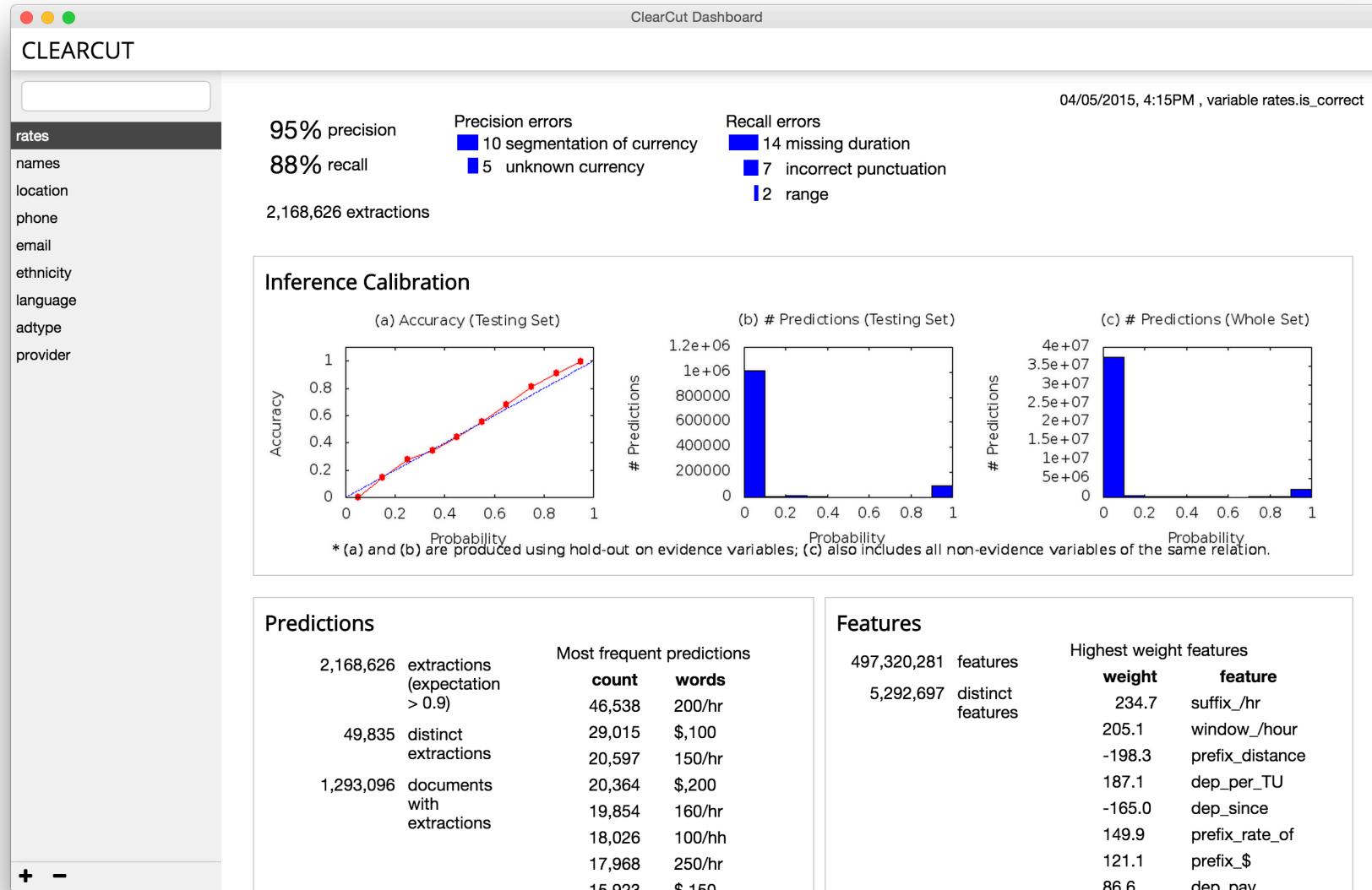
[Zhang Ph.D. Thesis, Wisconsin, 2015]

[Niu; SIGMOD 2016]

Engineer applies code change to one of:

1. Candidate generator fn
  2. Feature fns
  3. Supervision rules
1. Repeat cycle until output database is *good enough*

...  
experts, not engineers



Mindbender

Mindtagger labeling-precision-escort ▾ Showing all ▾

Tags Export

✓ ✗ ⓘ + ⚡ range Note

with expectation **0.920** appeared in:

120/HH 220/45mins 280 FH -RRB-

Show details

✓ ✗ ⓘ + ⚡ range Note

with expectation **0.992** appeared in:

30Min I suggest 150 flowers 1Hr I suggest 250 flowers • ★

Show details

✓ ✗ ⓘ + ⚡ range Note

with expectation **1.000** appeared in:

Before 4 85/hr

Show details

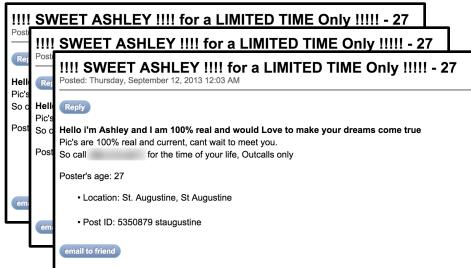
◀◀ ◀ 1 2 ▶ ▶▶ 50 /page

The screenshot shows the Mindbender application window. At the top, there's a toolbar with standard OS X window controls, followed by tabs for 'Mindtagger' and 'labeling-precision-escort', a search bar ('Showing all'), and buttons for 'Tags' and 'Export'. Below the toolbar are two rows of action buttons: a green checkmark, a red 'x', an info icon, a plus sign, and a 'range' icon, followed by a 'Note' input field. The main area contains three search results, each enclosed in a light gray box. The first result has a yellow highlight over the text '280 FH -RRB-'. The second result has a yellow highlight over the text 'I suggest 150 flowers 1Hr I suggest 250 flowers'. The third result has a yellow highlight over the text '85/hr'. Each result includes a 'Show details' button and another row of the same action buttons and note field. At the bottom of the window is a navigation bar with page control icons (double arrows, single arrows, page numbers 1, 2, etc.) and a '/page' input field.

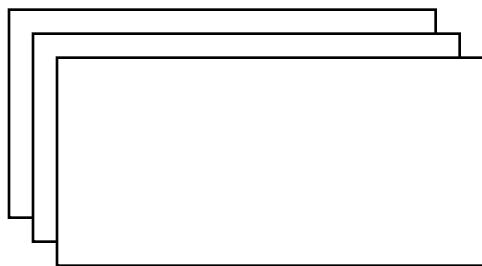
# Agenda

- Introduction
- Information Extraction
- **Law Enforcement Tool: Evidently LE**
- Analyzing Sex Market Pricing
- Conclusion

# Raw Data



Online Advertisements



Forum Posts

- 120M ads posted between 2011-2016
- Different analyses & tools used different amounts of data

- >0.5M forum discussion posts
- Corpus is crawl-limited

# Forum Posts: Evidence of Drug Abuse?

- Some reviews **strongly indicate illegal activities** that are suspected of **human trafficking**:

## Drug abuse

... Arrived at the hotel, very low budget. **Drug dealers and half naked women hanging around.** Against my better judgement after calling to get the room number I went up. She wasn't ugly and had a decent body so I went for it ...

... Danielle is **very skinny, a bit nervous**, but loosens up after multiple visits. Can be a bit out there/spacey, **almost seems drugged out**. Overall the service is good for OC, probably wont be going back ...

...Then she takes out her purse, a clear plastic one at that, **with needles clearly visible to me**. I asked if she was a diabetic to which she replied no, **took a needle out, then went into the bathroom for a couple of mintutes** ...

# Forum Posts: Forced Prostitution?

- Some reviews **strongly indicate illegal activities** that are suspected of **human trafficking**:

## Forced prostitution

...very short, she is the one in the pics but looks ugly in person, not friendly at all, looks like she's been forced to do this. she told me she has a pimp. has some wounds on her body .. got donation for full hour and had only one shot.. ...

...Tyra is young black babe with a nice body. Works for an agency that's like a factory. High volume. She's an 8 on a scale of one to ten but would give higher marks accept for this unusual scar above her tits. Looks like a knife wound...

...She's from Georgia and comes here from time to time. She didn't seem very excited like she's being forced to trick. Not really into that. Solid provider who gets the job done. Book with confidence. ...

# Forum Posts: Underage Prostitution?

- Some reviews **strongly indicate illegal activities** that are suspected of **human trafficking**:

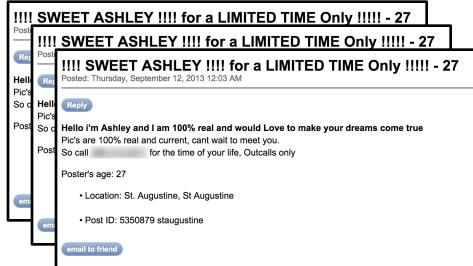
## Teenage prostitution

... Klara arrived on time and looked great. She is a dark haired little beauty who definitely looks like a teenager. (She's 19....

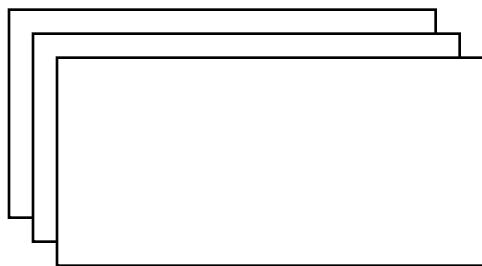
... Devlyn called me and we arranged and incall meeting. She said she would not be dressed as usual because they were in the process of moving. Devlyn greeted me at the door and I was surprised at how small she really is. She still has the face of a teenager ...

... She easily could have passed for a teenager. We sat on the couch and talked for a while, before getting started ...

# Extracted Data



Online Advertisements

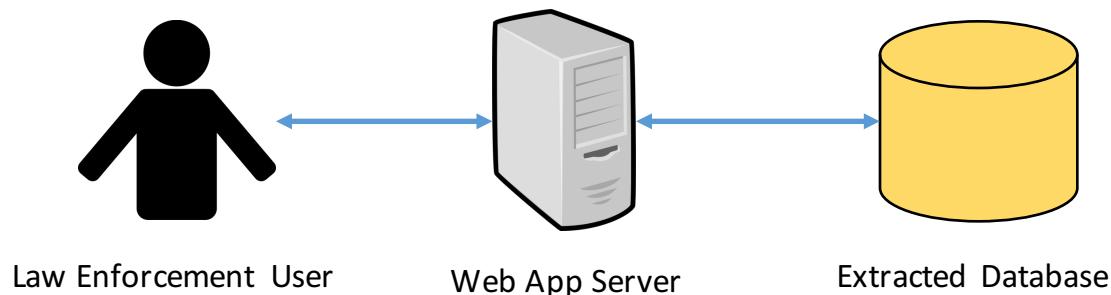


Forum Posts

- More than 2 dozen structured fields: incall/outcall, price, location, etc
- Availability of data depends on poster
- **Data shipped to law enforcement since 2014, used to make real arrests**

- Extract email, phone, text class
- Integrate using email/phone from advertisement data

# Law Enforcement Tool: Evidently LE



- Web-based tool for law enforcement agencies
- Being commercialized, piloted now
- Most intellectual content is in the data, not the user-facing tool

memex.deepdivedata.org

Evidently LE

domain\_type (28,179,090)

ads	27,951,440
reviews	227,650

flags (12,363)

New to Biz	3,755
Underage	2,907
Good English	2,071
Poor English	1,631
LE Cautious	1,521
Fake Photo	286
Drug Use	111
Physical Abuse	81

domain (28,179,090)

backpage	12,141,634
escortphonelist	3,153,196
escortsincollege	2,914,412
craigslist	2,286,441
cityxguide	1,610,211
escortedadults	1,306,929

28,179,091 documents

Evidently LE

memex.deepdivedata.org

28,179,091 documents

domain\_type (28,179,090)

ads	27,951,440
reviews	227,650

flags (12,363)

New to Biz	3,755
Underage	2,907
Good English	2,071
Poor English	1,631
LE Cautious	1,521
Fake Photo	286
Drug Use	111
Physical Abuse	81

domain (28,179,090)

backpage	12,141,634
escortphonelist	3,153,196
escortsincollege	2,914,412
craigslist	2,286,441
cityxguide	1,610,211
escortsdaddy	1,204,929

memex.deepdivedata.org

Evidently LE

domain\_type (28,179,090)

ads	27,951,440
reviews	227,650

flags (12,363)

New to Biz	3,755
Underage	2,907
Good English	2,071
Poor English	1,631
LE Cautious	1,521
Fake Photo	286
Drug Use	111
Physical Abuse	81

domain (28,179,090)

backpage	12,141,634
escortphonelist	3,153,196
escortsincollege	2,914,412
craigslist	2,286,441
cityxguide	1,610,211
escortsdaddy	1,306,929

28,179,091 documents

memex.deepdivedata.org

Evidently LE

domain\_type (28,179,090)

ads	27,951,440
reviews	227,650

flags (12,363)

New to Biz	3,755
Underage	2,907
Good English	2,071
Poor English	1,631
LE Cautious	1,521
Fake Photo	286
Drug Use	111
Physical Abuse	81

domain (28,179,090)

backpage	12,141,634
escortphonelist	3,153,196
escortsincollege	2,914,412
craigslist	2,286,441
cityxguide	1,610,211
escortedadults	1,306,929

28,179,091 documents

memex.deepdivedata.org

Evidently LE LOCATIONS: New York

1-10 of 321,393 results for [ locations:"New York" ]

1.

ads	303,987
reviews	17,406

post\_date 2014-05-29

ages 19

phones (917) 547-0723

locations New York East Side, 27th & Park Ave

917 547 0723 I'm Jessica - a college student majoring in Physical Education with a hunger for learning and experiencing new ideas. I am uninhibited, very personable, lots of fun! I'm here with Scarlett (Hot! with a sleek model's body), and some of my other room mates in our duplex loft apartment - check out some of our all-real pics on our blog. We love what we do, and we will deliver! Call and come over for the Complete Package!  
[gentlemanschoicency.com](http://gentlemanschoicency.com)

2.

ads	/ backpage
	/
	--Amazing SEXY Petite REDHEAD ♥♥Beautiful Babe ♥♥Limited Time Only
	26

post\_date 2014-05-29

ages 26

phones (305) 791-5988

locations New York Midtown West, 32 Street & 7 Ave, Madison Square Garden

service incall

1 2 3 4 5 6 7 8 9 ► ►

memex.deepdivedata.org

Evidently LE LOCATIONS: New York

1-10 of 321,393 results for [ locations:"New York" ]

1.

ads / backpage / ~~~ Jessie ~ 19 Year-Old College Student ~ Scarlett ~ All the Way to Paradise - 19 backpage:5384748

post\_date 2014-05-29

ages 19

phones (917) 547-0723

locations New York East Side, 27th & Park Ave

917 547 0723 I'm Jessica - a college student majoring in Physical Education with a hunger for learning and experiencing new ideas. I am uninhibited, very personable, lots of fun! I'm here with Scarlett (Hot! with a sleek model's body), and some of my other room mates in our duplex loft apartment - check out some of our all-real pics on our blog. We love what we do, and we will deliver! Call and come over for the Complete Package!  
[gentlemanschoicenc.com](http://gentlemanschoicenc.com)

2.

ads / backpage / --Amazing SEXY Petite REDHEAD ♥♥Beautiful Babe ♥♥Limited Time Only backpage:5385107

post\_date 2014-05-29

ages 26

phones (305) 791-5988

locations New York Midtown West, 32 Street & 7 Ave, Madison Square Garden

service incall

1 2 3 4 5 6 7 8 9 ► ►

domain\_type (321,393)

ads 303,987

reviews 17,406

flags (1,011)

New to Biz 225

Underage 214

Poor English 204

Good English 183

LE Cautious 144

Fake Photo 31

Drug Use 7

Physical Abuse 3

domain (321,393)

backpage 281,891

escortsintheus 18,749

ter-review 17,406

myproviderguide 3,254

adultsearch 93

memex.deepdivedata.org

Evidently LE LOCATIONS: New York

1-10 of 321,393 results for [ locations:"New York" ]

domain\_type (321,393)

- ads 303,987
- reviews 17,406

flags (1,011)

- New to Biz 225
- Underage 214
- Poor English 204
- Good English 183
- LE Cautious 144
- Fake Photo 31
- Drug Use 7
- Physical Abuse 3

domain (321,393)

- backpage 281,891
- escortsintheus 18,749
- ter-review 17,406
- myproviderguide 3,254
- adultsearch 93

1. ads / backpage / ~~~ Jessie ~ 19 Year-Old College Student ~~ Scarlett ~ All the Way to Paradise - 19 backpage:5384748

post\_date 2014-05-29

ages 19

phones (917) 547-0723

locations New York East Side, 27th & Park Ave

917 547 0723 I'm Jessica - a college student majoring in Physical Education with a hunger for learning and experiencing new ideas. I am uninhibited, very personable, lots of fun! I'm here with Scarlett (Hot! with a sleek model's body), and some of my other room mates in our duplex loft apartment - check out some of our all-real pics on our blog. We love what we do, and we will deliver! Call and come over for the Complete Package!  
[gentlemanschoicency.com](http://gentlemanschoicency.com)

2. ads / backpage / --Amazing SEXY Petite REDHEAD ♥♥Beautiful Babe ♥♥Limited Time Only backpage:5385107

post\_date 2014-05-29

ages 26

phones (305) 791-5988

locations New York Midtown West, 32 Street & 7 Ave, Madison Square Garden

service incall

1 2 3 4 5 6 7 8 9 ► ►

memex.deepdivedata.org

Evidently LE

domain\_type (28,179,090)

ads	27,951,440
reviews	227,650

flags (12,363)

New to Biz	3,755
Underage	2,907
Good English	2,071
Poor English	1,631
LE Cautious	1,521
Fake Photo	286
Drug Use	111
Physical Abuse	81

domain (28,179,090)

backpage	12,141,634
escortphonelist	3,153,196
escortsincollege	2,914,412
craigslist	2,286,441
cityxguide	1,610,211
escortsdaddy	1,306,929

28,179,091 documents

memex.deepdivedata.org

FLAGS: Drug Use

Evidently LE

domain\_type (111)

reviews 111

flags (116)

Drug Use 111

New to Biz 2

LE Cautious 1

Physical Abuse 1

Underage 1

domain (111)

ter-review 111

locations (107)

Los Angeles 8

New York 7

Minnesota 6

San Diego 6

Atlanta 5

Carolina 5

San Francisco 5

1-10 of 111 results for [ flags:"Drug Use" ]

1. reviews / ter-review / <http://www.theeroticreview.com/reviews/showReview.asp?Review=1081746> ter-review:214928

names TS LEXA

flags Drug Use

I wanted to see Lexa for a long time, and finally saw her ad in town.

Juicy\_Details

I set up an appointment on Monday at 9pm, drove 40 miles to see her and waited for almost two hours trying to call her with no success. I later got a text message saying that she fell asleep. I figured I would give it another shot today since she was going out of town. I called up and she said she was staying at the Sheridan in town. I showed up and after the second time I called she picked up. She acted again like she was high. She said she had to figure out how to get me in the hotel and hung up?? I waited for fifteen minutes and called her back, two calls again and she picked up. She acted as if she was completely stoned. I told her I was waiting downstairs, and she asked "who is this?". Finally she called me back and said she was coming down, after another fifteen minutes went by I got up and walked out. She called me again as I drove by the front of the hotel asking "heyyyy were are you?", and I told her I was leaving that this was the worst experience I had ever had. She laughed and asked when I was coming back, I had to look at the phone in disbelief. This was followed up by (not even joking) ten phone calls, a message and text messages stating "heyy". She is a total flake, when I saw her downstairs she looked like she just woke up, hair terrible, and I wouldn't even think about it. It was a complete disappointment.

2. reviews / ter-review / <http://www.theeroticreview.com/reviews/showReview.asp?Review=998136> ter-review:18790

username thehistorymaker

names Jennifer Kayla

ages 26 - 40

◀◀ ▶▶ 1 2 3 4 5 6 7 8 9 ▶▶

memex.deepdivedata.org

Q Evidently LE

FLAGS: Drug Use FLAGS: Physical Abuse

1-1 of 1 results for [ flags:"Drug Use" flags:"Physical Abuse" ]

1. reviews / ter-review / http://www.theeroticreview.com/reviews/showReview.asp?Review=217725 ter-review:58450

username	acosey
names	Stacy
ages	21 - 25
phones	(504) 598-4030
ethnicity	White
locations	New Orleans
service	outcall
flags	Drug Use Physical Abuse

I had seen Stacey's ad in EROS and had seen a few positive reviews so I gave the agency a call. Despite the fact it had been planned for several days, she showed up at my Canal Street hotel in Jeans and work boots and looked nothing like her pictures. It was the same girl but she was very thin, had bruises all over her like some guy had beat her up. She had a bad place on her face and neck that she said she did by accident with a curling iron. Frankly she acted like she was on drugs. She could barely hold her eyes open. She actually nodded off a time or two while she was talking to me which she did for 50 minutes telling me all her troubles. Then when we got started, she told me we were going over the hour and I would have to pay extra if I wanted to get off. What a rip off. Her hair is much shorter than in her pictures and she is not near as pretty. She also told me she was 24, not 19 like the ad says. She gave me her private number afterwards and told me to call her directly but I wont see her again and I can't recommend her. Read on for the few details!!

◀◀ 1 ▶▶

domain\_type (1)

reviews 1

flags (2)

Drug Use 1

Physical Abuse 1

domain (1)

ter-review 1

locations (1)

New Orleans 1

ethnicity (1)

White 1

phones (1)

(504) 598-4030 1

names (1)

Stacy 1

service (1)

memex.deepdivedata.org

Q Evidently LE

FLAGS: Drug Use FLAGS: Physical Abuse

1-1 of 1 results for [ flags:"Drug Use" flags:"Physical Abuse" ]

1. reviews / ter-review / http://www.theeroticreview.com/reviews/showReview.asp?Review=217725 ter-review:58450

username	acosey
names	Stacy
ages	21 - 25
phones	(504) 598-4030
ethnicity	White
locations	New Orleans
service	outcall
flags	Drug Use Physical Abuse

I had seen Stacey's ad in EROS and had seen a few positive reviews so I gave the agency a call. Despite the fact it had been planned for several days, she showed up at my Canal Street hotel in Jeans and work boots and looked nothing like her pictures. It was the same girl but she was very thin, had bruises all over her like some guy had beat her up. She had a bad place on her face and neck that she said she did by accident with a curling iron. Frankly she acted like she was on drugs. She could barely hold her eyes open. She actually nodded off a time or two while she was talking to me which she did for 50 minutes telling me all her troubles. Then when we got started, she told me we were going over the hour and I would have to pay extra if I wanted to get off. What a rip off. Her hair is much shorter than in her pictures and she is not near as pretty. She also told me she was 24, not 19 like the ad says. She gave me her private number afterwards and told me to call her directly but I wont see her again and I can't recommend her. Read on for the few details!!

◀◀ 1 ▶▶

domain\_type (1)

reviews (1)

flags (2)

Drug Use (1)

Physical Abuse (1)

domain (1)

ter-review (1)

locations (1)

New Orleans (1)

ethnicity (1)

White (1)

phones (1)

(504) 598-4030 (1)

names (1)

Stacy (1)

service (1)

username acosey

names Stacy

ages 21 - 25

phones (504) 598-4030

ethnicity White

locations New Orleans

service outcall

flags Drug Use Physical Abuse

memex.deepdivedata.org

Q Evidently LE PHONES: (504) 598-4030

1-10 of 17 results for [ phones:"(504) 598-4030" ]

1. reviews / ter-review / http://www.theeroticreview.com/reviews/showReview.asp?Review=360065 ter-review:11219

username	neworleansfreak
names	Charlotte
ages	31 - 35
phones	(504) 598-4030 (504) 944-4401
emails	abetterescort@hotmail.com
ethnicity	White
locations	New Orleans
service	incall and outcall

She's very friendly, fun to talk to if you strike up the right conversation, and enthusiastic in her performance. The service was not rushed though it may be better to get an outcall. The incall place was not hers (it was a special incall rented place) and she told me that it was the current incall renter gave a lousy place. You may consider having it at your own place, else check with her if it is in Metairie incall (I guess would have been nicer). The phone number is from the agency.

2. reviews / ter-review / http://www.theeroticreview.com/reviews/showReview.asp?Review=374491 ter-review:59797

username	pczc2003
names	Raven
ages	26 - 30

1 2

locations (17)

- New Orleans 17

ethnicity (17)

- White 15
- African American 1
- Middle Eastern 1

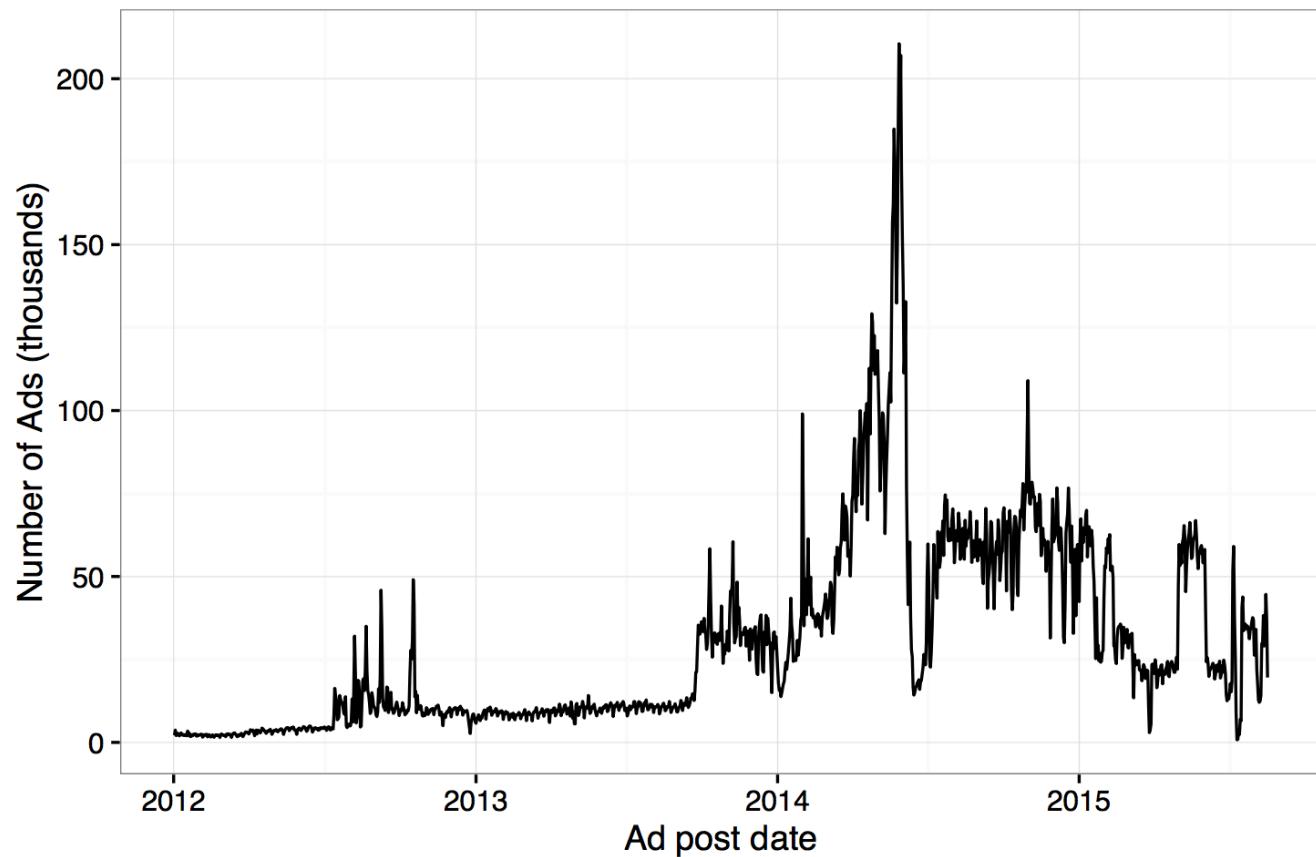
phones (19)

- (504) 598-4030 17
- (504) 944-4401 2

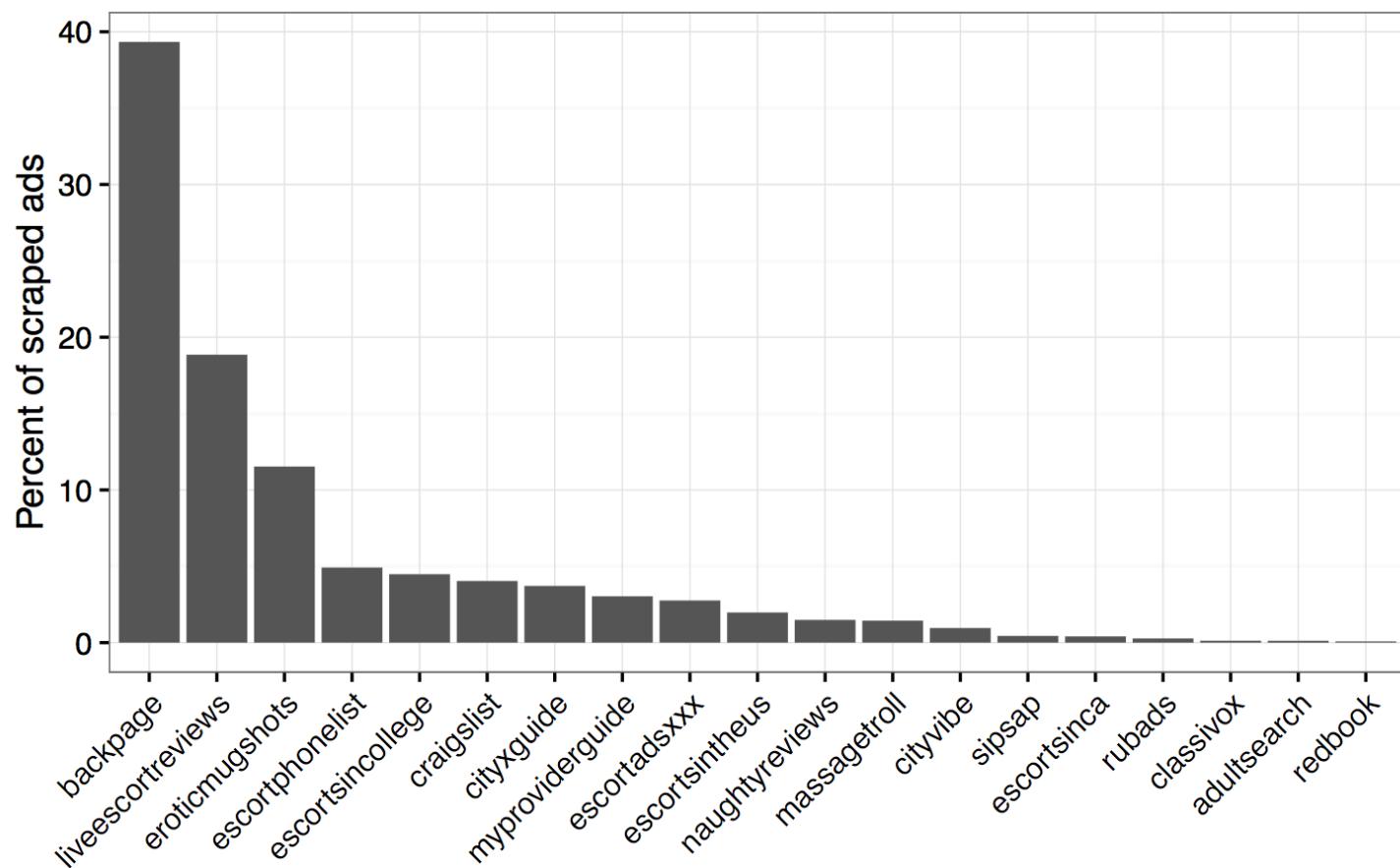
# Agenda

- Introduction
- Information Extraction
- Law Enforcement Tool: Evidently LE
- **Analyzing Sex Market Pricing**
- Conclusion

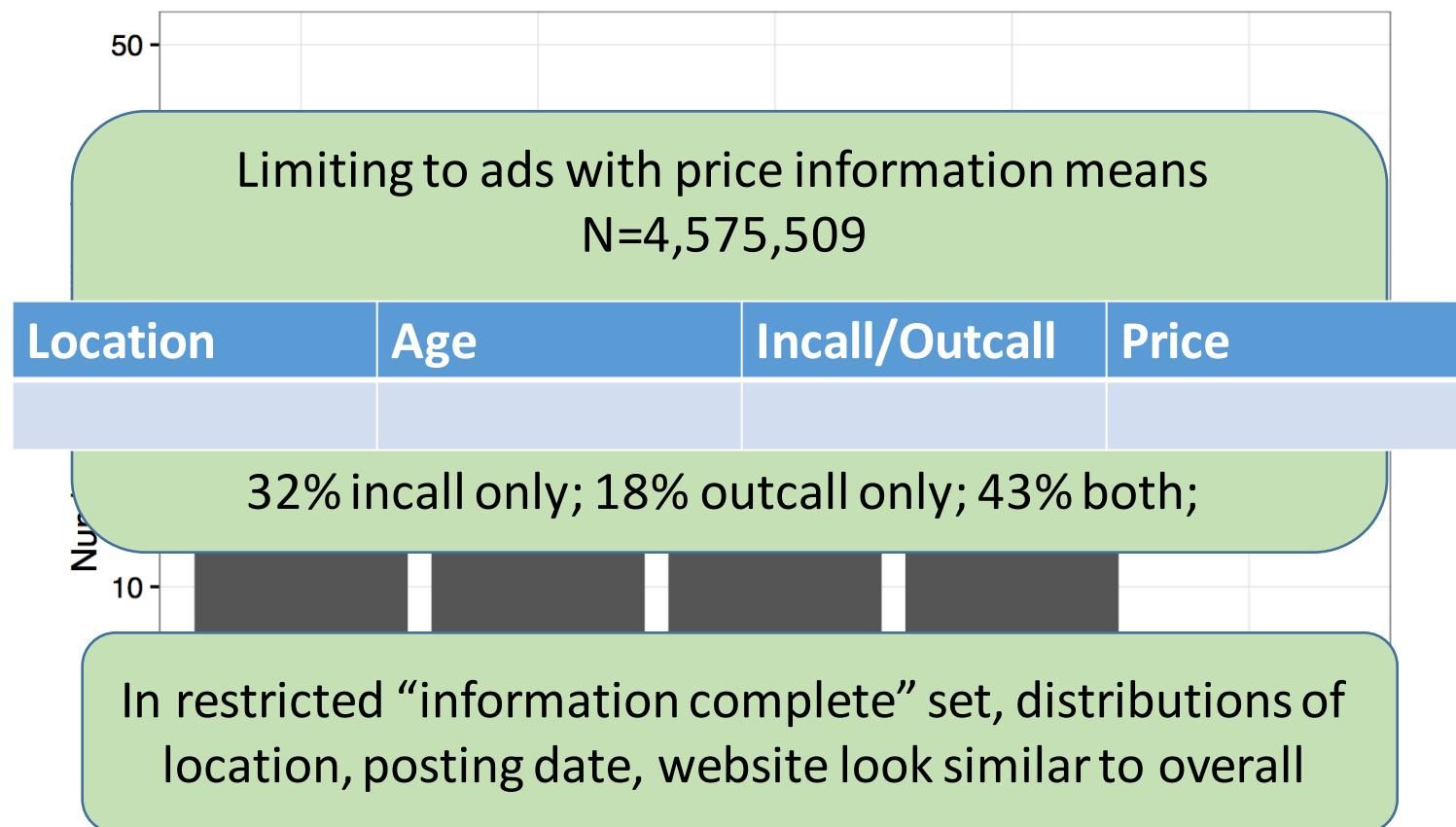
# Posting Date of Scrapped Advertisements



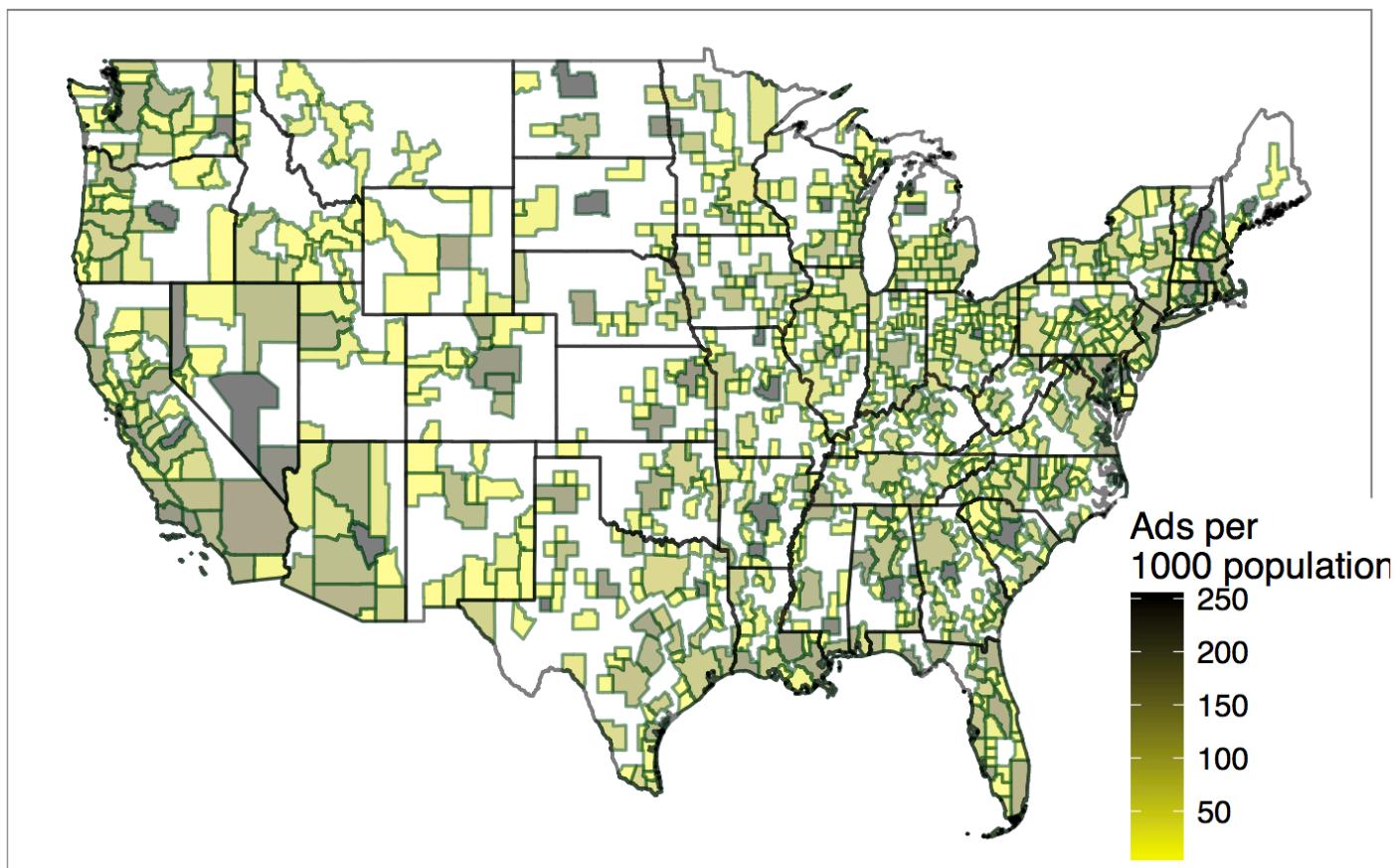
# Ad Origin Site



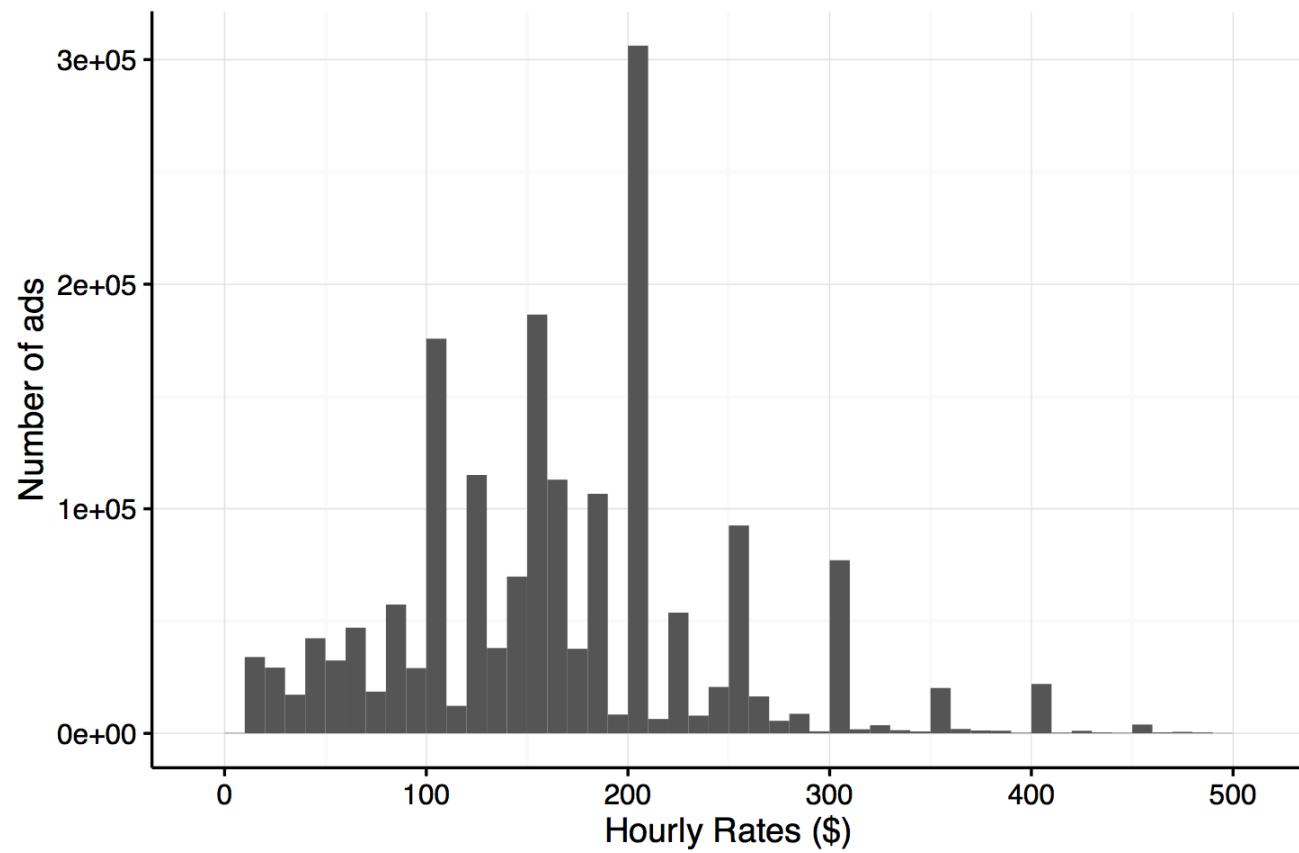
## Availability of Different Fields



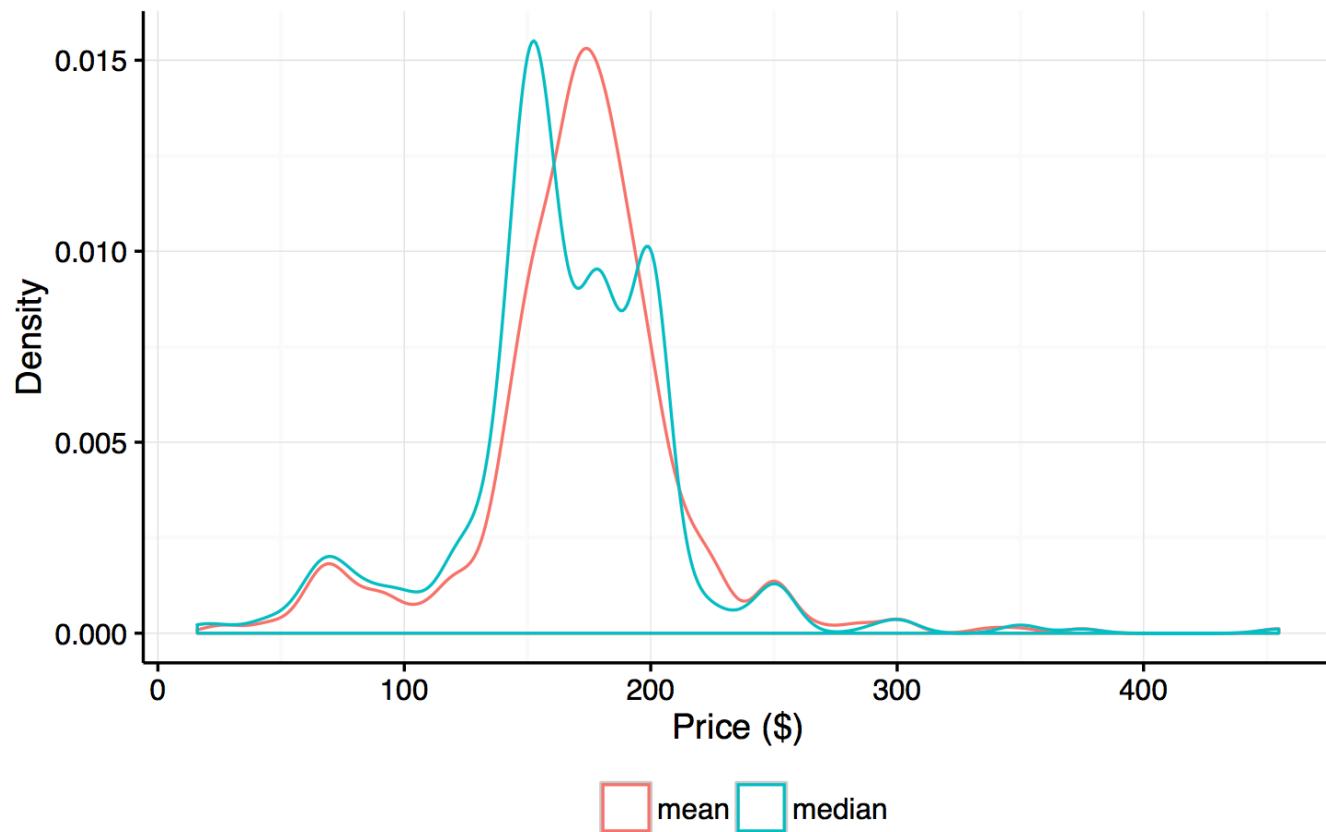
## Ads Per Capita in CBSAs



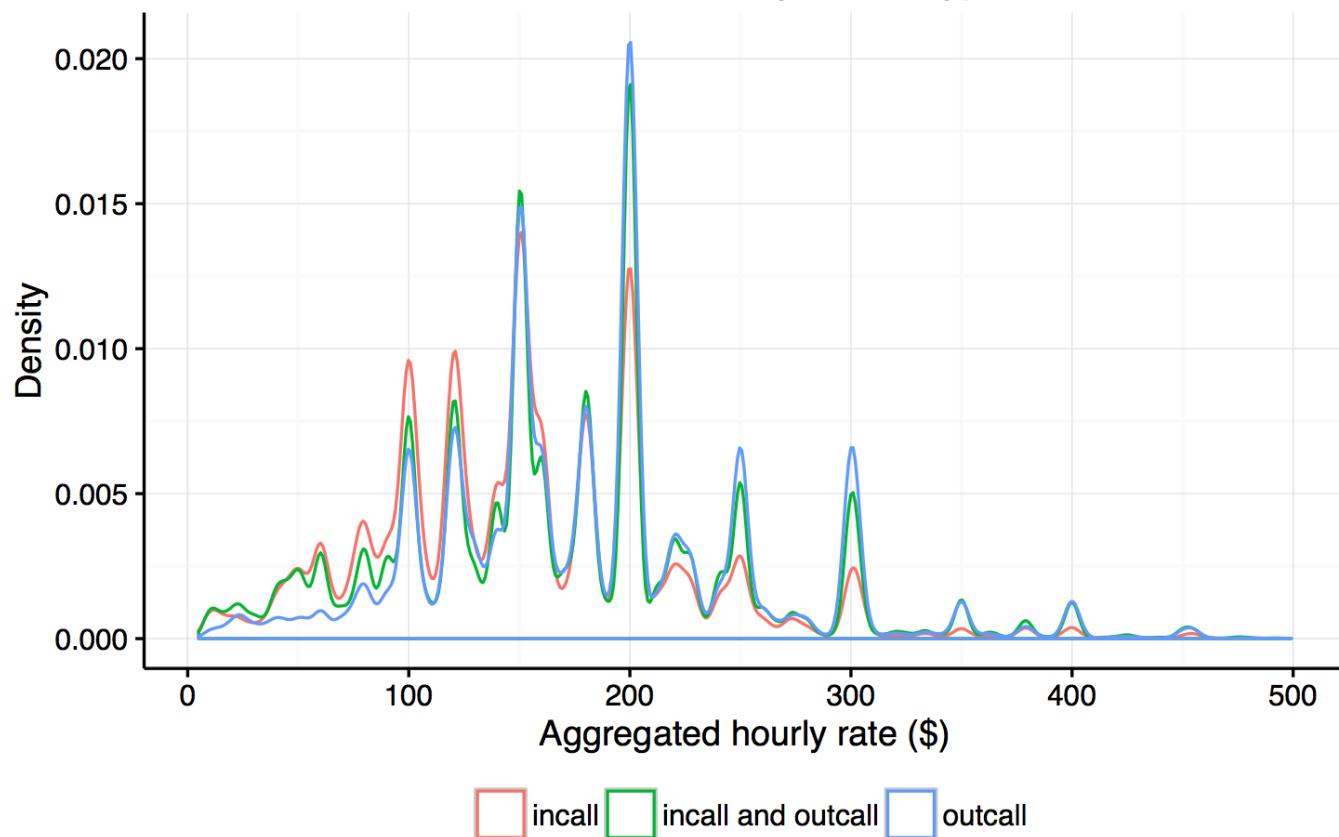
# Distribution of Hourly Prices



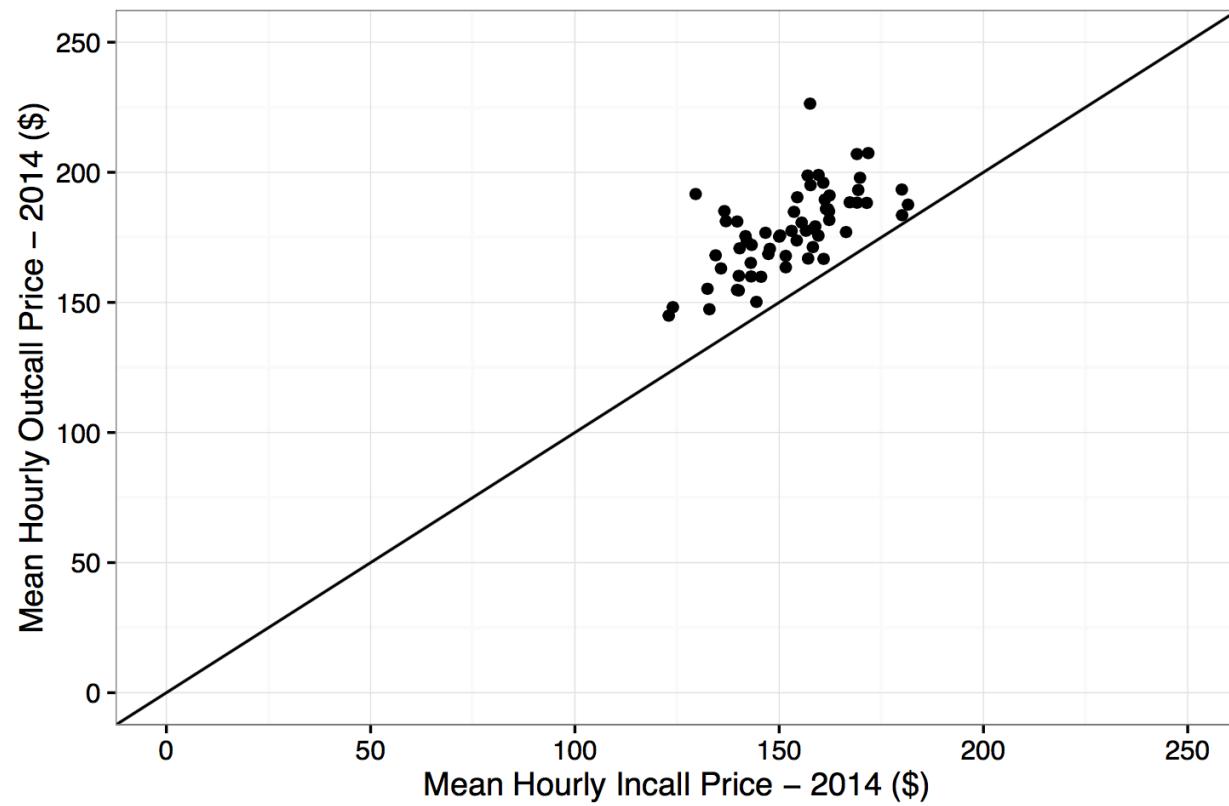
# Prices Vary by CBSA



# Prices Vary by Service Venue



# Many MSAs Have an Outcall Premium



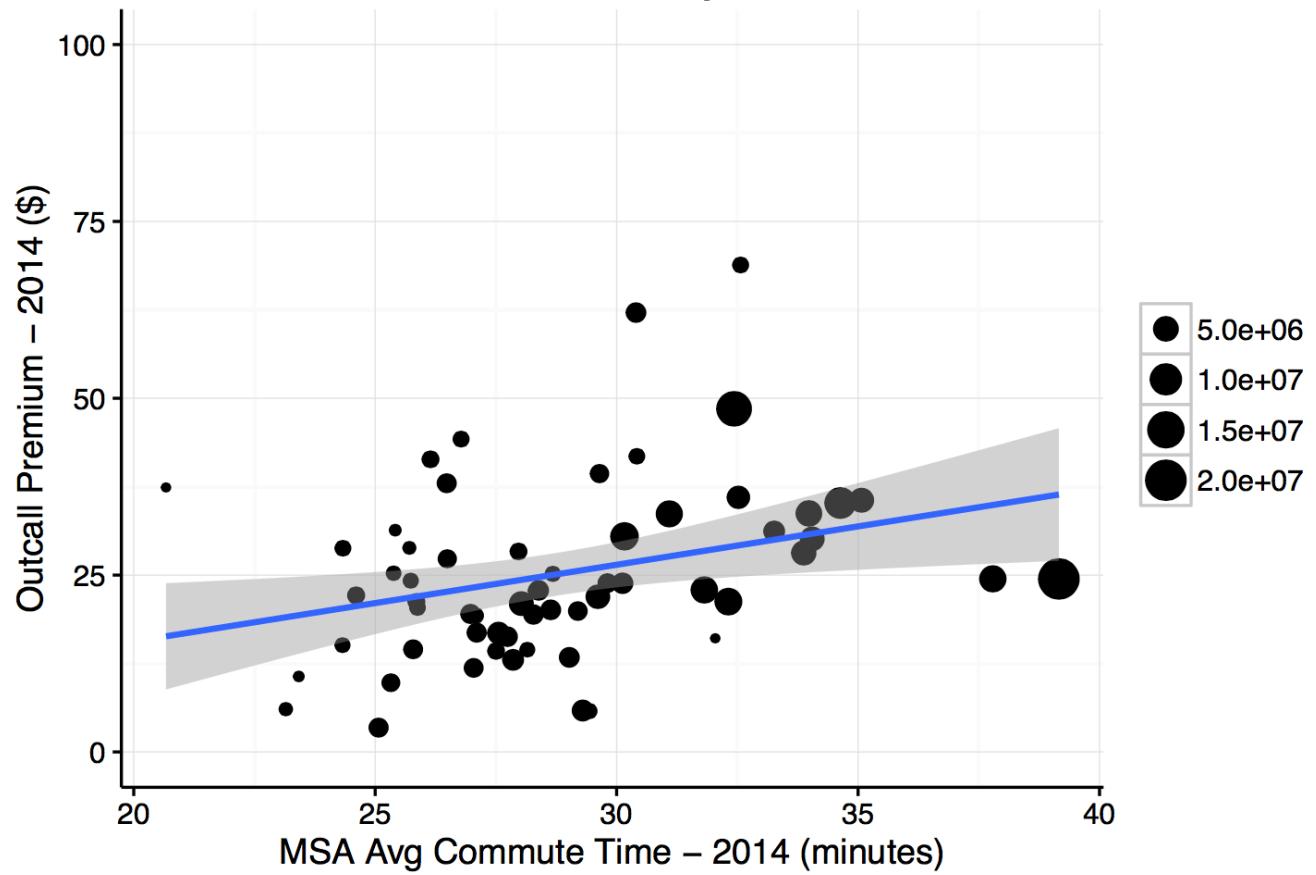
# Open Questions

- What factors drive the outcall premium?
- To what extent are prices driven by the licit labor market?
- *Metaquestion: How rational is pricing in sex work market?*
  
- Questions we have examined but will not discuss:
  - Are advertisements truthful about provider ages?
  - What is the price premium/discount for H/T-relevant factors?
    - Drug abuse
    - Physical abuse
    - English-fluent
    - Youthful

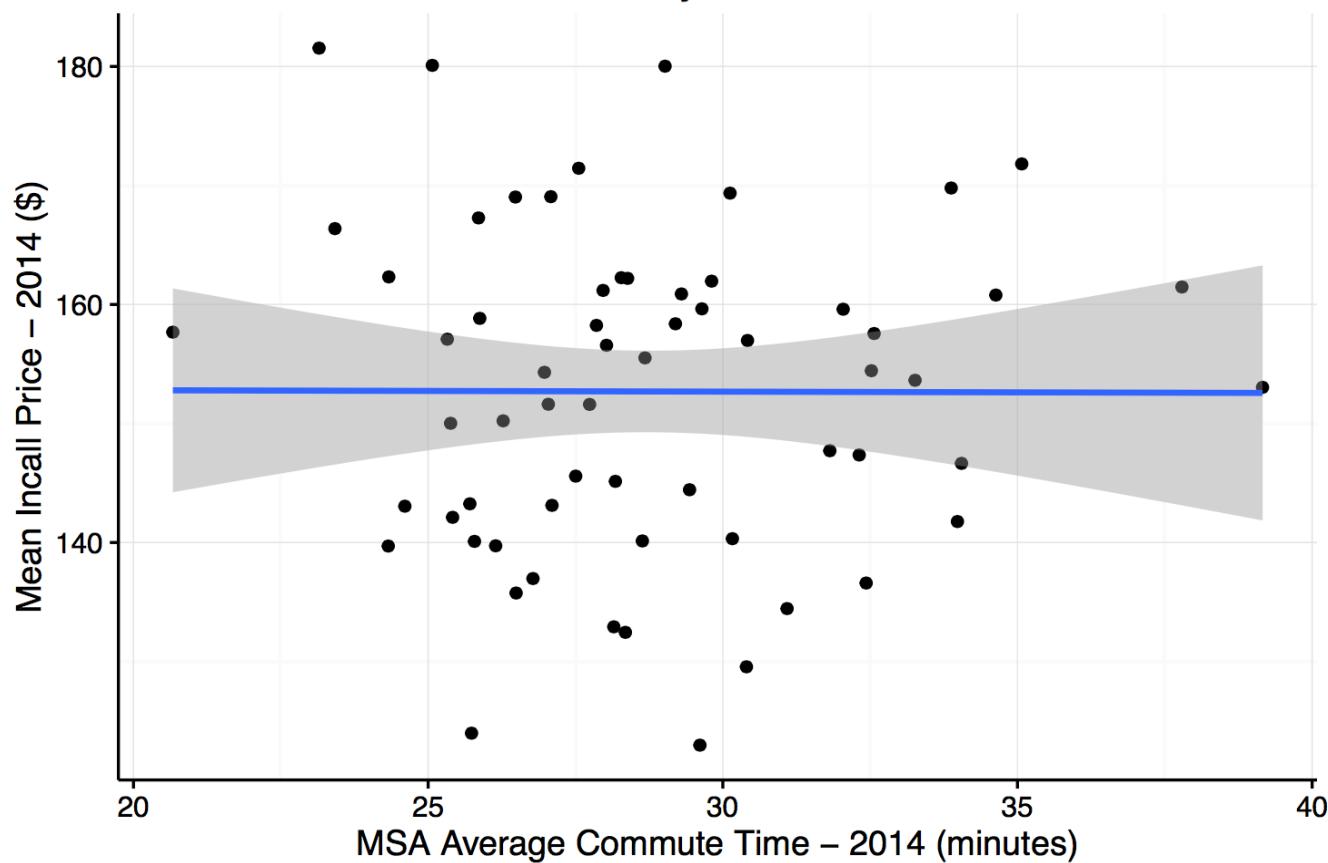
# External Datasets

- Risk of crime driving outcall premium?
  - Uniform Crime Reporting (UCR) for crime types and rates
  - Law Enforcement and Administrative Statistics (LEMAS) for # of police officers
  - National Incident Based Reporting System (NIBRS) for rate of arrest
- Risk of disease?
  - National Health Indicators Warehouse
- Commute times?
  - American Community Survey

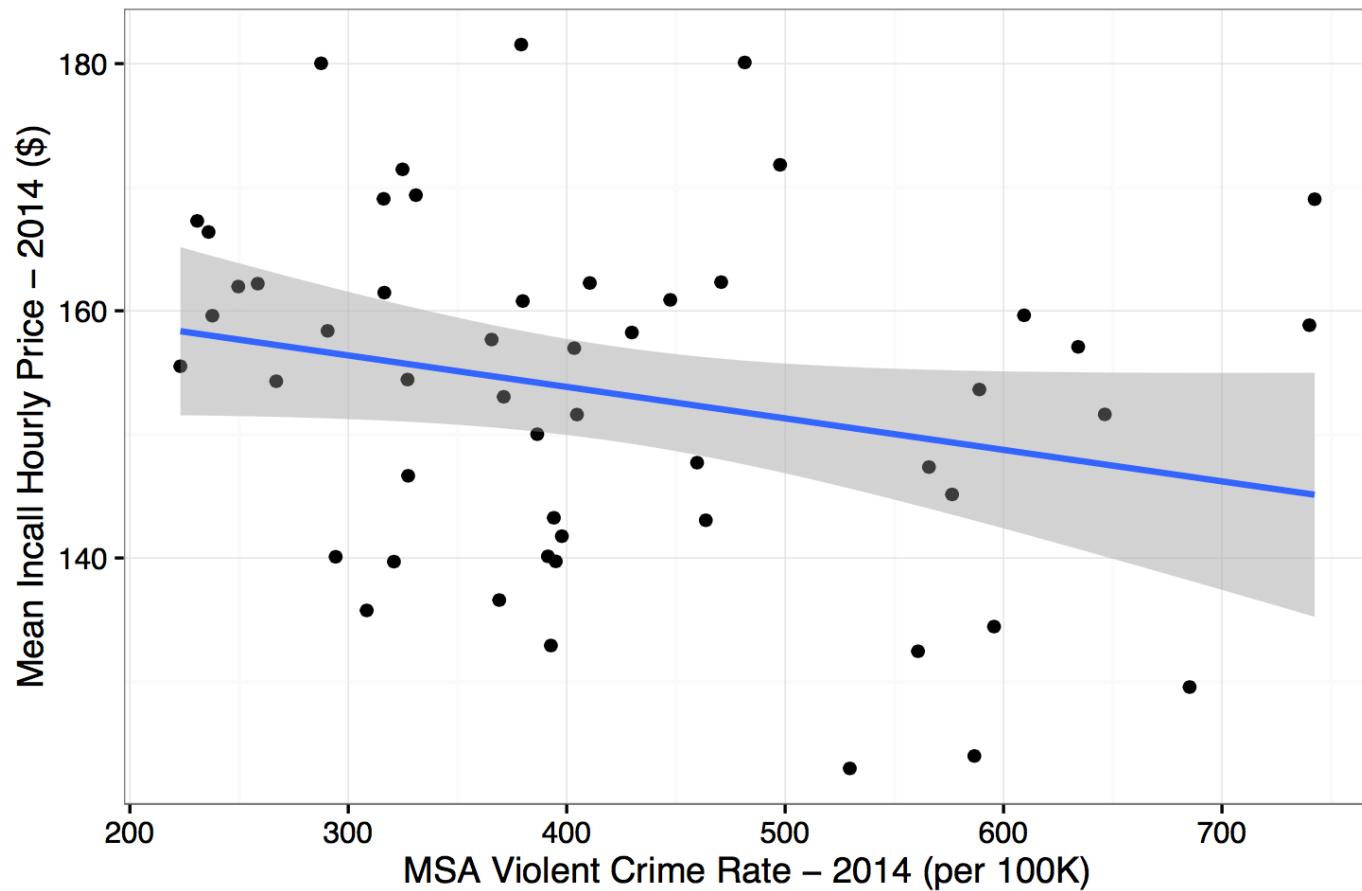
# Outcall Premium By Commute Time



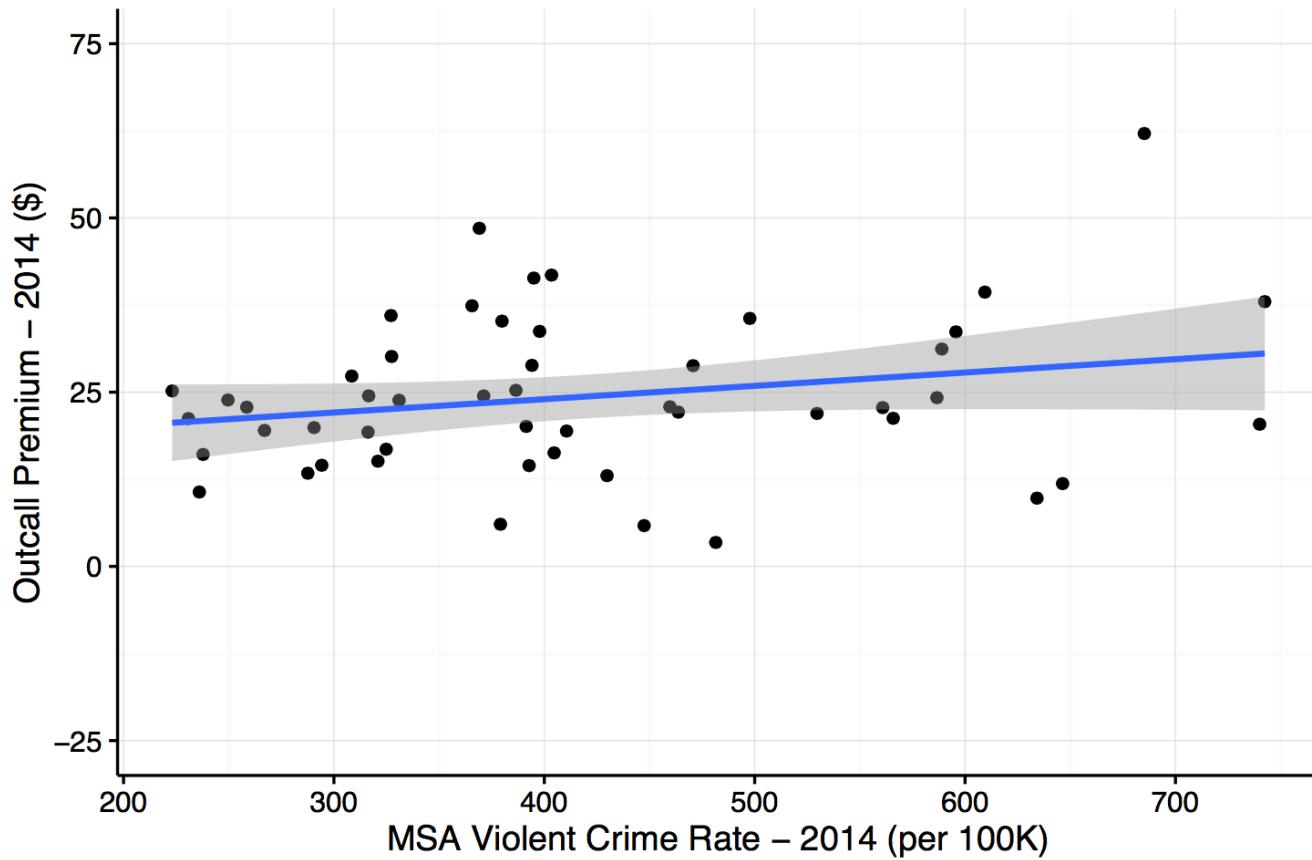
## Incall Price By Commute Time



# Incall Price by Violent Crime Rate



# Outcall Premium by Violent Crime Rate



# Regression-Based Analysis

- More detailed regression-based analysis examines more risks, controls for mitigating factors, examines combined risks
- Findings:
  - Commute and violence are very robust predictors of outcall premium
  - Roughly 2/3 of premium is commute; remainder violence

# Labor Market Interactions

- How do prices react to women's opportunities in broader market?
- If prices are insensitive to broader conditions, it might suggest a large amount of compelled participation
- Use Bartik wages [[Bartik 1991](#)] to measure labor market demand apart from local supply
- Build regression model to predict price in a given MSA and month
  - Control for posting venue and commute distance
  - Control for MSA-specific effects, seasonal time-of-year effects

# Labor Market Findings

- After controlling for MSA-specific effects, robust ***negative*** correlation between female opportunity and hourlong price
  - \$1 in overall wage increase predicts \$0.43 **decrease** in hourly rate
  - Why??
- Divide market into low-, mid-, high-end prices (<25%, 25-75; >75%)
- Negative correlation between female opportunity and # mid-price ads
  - Predicts 7.1% drop in first-time advertisers at mid-price (signif at 0.01)
  - 1\$ in overall wage increase predicts 3% drop in overall participation
- But, positive correlation between female opp'ty and # low-price ads
  - Almost identical rise in participation; 3.6% increase in first-time ads

# Labor Market Findings

- Findings are consistent with idea that as labor market improves,
    - Mid-price workers exit sex work market to participate in licit economy
    - Low-price workers enter sex work market, due to increased competition by higher-price workers
- (See [Bartlett, 2013], on minimum wage, low-wage opportunity, and immigration rates)

# Related Work

- Social science:
  - Counts of prostitution ads on Craigslist [Cunningham and Kendall, 2011]
  - Price and age from 40K posts on one high-end site [Edlund, Engelberg, Parsons 2009]
  - Price, services, other data from 190K profiles on site where escorts fill preexisting form [Delap 2014]
  - 998 forms from a single site [Moffatt, Peters 2004]
- DeepDive:
  - Overview: [PhD. Thesis, Zhang, 2014], [Extracting Databases from Dark Data With DeepDive, SIGMOD 2016]