The structure of political media coverage as revealed by quoting patterns:

**QUOting the POTUS**

*Vlad Niculae (me)  Cornell (MPI-SWS)  Stanford*
*Caroline Suen  Stanford*
*Justine Zhang  Cornell (Stanford, MPI-SWS)  Cornell (MPI-SWS)*
*Cristian Danescu-Niculescu-Mizil  Cornell (MPI-SWS)*
*Jure Leskovec  Stanford*
The press claims to be fair and balanced.

America’s trust in media is at an all time low.

(Source: Gallup poll, Sept. 17 2014)
“Media Bias” is subjective

(Out of the ones who say it’s biased.)
Source: Gallup poll, Sept. 17 2014
“Media Bias” is subjective multidimensional hard to quantify

[Dalton et al, 1998] [Peake, 2007]

[Schiffer, 2006]


[Ho et al, 2008]

Predetermined dimension

Democrat/Republican

Democrat/Republican

Liberal/Conservative

& other theoretically possible dimensions...

Proxy to slant

Endorsement

Supreme Court Justices
We want a way to address media bias:

Large Scale, 
Unsupervised, 
Without fixing a bias dimension.

Insight: Quoting patterns.
Selection Bias

No space to report everything.

Outlets must make conscious choices.

“Fair world:” are the choices based only on newsworthiness?

Or is there something else behind?
With or without this Congress, I will keep taking actions that help the economy grow. But I can do a whole lot more with your help. Because when we act together, there's nothing the United States of America can't achieve.

Barack Obama, 2012 State Of The Union address.
With or without this Congress, I will keep taking actions that help the economy grow. But I can do a whole lot more with your help. Because when we act together, there's nothing the United States of America can't achieve.
With or without this Congress, I will keep taking actions that help the economy grow. But I can do a whole lot more with your help. Because when we act together, there's nothing the United States of America can't achieve.

—CNN, Huffington Post
Are quoting patterns *systematic*?  
Do they correspond to *intuition*?  
Can *language* characterize the bias?
Are quoting patterns *systematic*?
Do they correspond to *intuition*?
Can *language* characterize the bias?
Are quoting patterns *systematic*?

**Quoting graph**

**The New York Times**

**SALON**

**FOX NEWS.com**

275 online media outlets. (newspapers, blogs, ...)

54K Obama quotes in 6 years.

Edge = outlet reported quote.

230K edges.

\[ X = \text{The Adjacency Matrix} \]
Are quoting patterns *systematic*?

Can we predict quoting?

If we can predict quoting choices, it means there is a pattern.
Are quoting patterns systematic?

Can we predict quoting?

Leave out 250K entries of $X$. 
Are quoting patterns *systematic*?

Can we predict quoting?

Leave out **250K** entries of $X$.

<table>
<thead>
<tr>
<th>Matthews Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
</tr>
<tr>
<td>14%</td>
</tr>
<tr>
<td>28%</td>
</tr>
</tbody>
</table>

The New York Times

SALON

FOX NEWS
Are quoting patterns systematic?

Can we predict quoting?

Leave out 250K entries of $X$.
Are quoting patterns **systematic?**

Can we predict quoting?

Leave out **250K** entries of $X$.

In a “**fair world**” without bias:
- popular quotes get cited more,

![](chart.png)
Can we predict quoting?

Leave out 250K entries of $X$.

In a "fair world" without bias:
- popular quotes get cited more,
Are quoting patterns \textit{systematic}?

Can we predict quoting?

Leave out \textbf{250K} entries of $X$.

In a \textit{“fair world”} without bias:
- popular quotes get cited more,
- and big outlets quote more.

\begin{center}
\begin{tabular}{c|c|c|c}
& Quote Popularity & Matthews Correlation Coefficient & \\
\hline
Recall & 0\% & 14\% & 28\% \\
\hline
Precision & 0\% & 12.5\% & 25\% \\
\hline
\end{tabular}
\end{center}
Are quoting patterns \textit{systematic}?

Can we predict quoting?

Leave out \textbf{250K} entries of $X$.

In a \textit{"fair world"} without bias:
- popular quotes get cited more,
- and big outlets quote more.

<table>
<thead>
<tr>
<th></th>
<th>Quote Popularity</th>
<th>+ Outlet Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matthews Correlation Coefficient</td>
<td>0%</td>
<td>14%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>17%</td>
<td>12,5%</td>
</tr>
<tr>
<td>33%</td>
<td>25%</td>
</tr>
<tr>
<td>50%</td>
<td>37,5%</td>
</tr>
</tbody>
</table>

Leave out 250K entries of $X$. In a “fair world” without bias:
- popular quotes get cited more,
- and big outlets quote more.
Are quoting patterns **systematic**?

Can we predict quoting?

Matrix completion by low-rank factorization
Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by low-rank factorization

\[ X \approx \hat{X} = U \cdot V^T \]
Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by low-rank factorization

\[
X \approx \hat{X} = U \cdot V^T
\]

\[
\begin{pmatrix}
\text{outlet space}
\end{pmatrix}
\approx
\begin{pmatrix}
\text{quote space}
\end{pmatrix}
\]

\[
\begin{bmatrix}
\text{The New York Times}
\end{bmatrix}
\approx
\begin{bmatrix}
\text{SALON}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{FOX NEWS.com}
\end{bmatrix}
\]

Graph showing Matthews Correlation Coefficient with Recall and Precision:

- Quote Popularity
- Outlet Size

- Matthews Correlation Coefficient
  - 0%
  - 14%
  - 28%

- Recall
  - 0%
  - 17%
  - 33%
  - 50%

- Precision
  - 0%
  - 12.5%
  - 25%
  - 37.5%
  - 50%
Are quoting patterns **systematic**?

**Can we predict quoting?**

Matrix completion by low-rank factorization

\[ X \approx \hat{X} = U \cdot V^T \]

(quote space)

(outlet space)

---

**Matthews Correlation Coefficient**

<table>
<thead>
<tr>
<th>Quote Popularity + Outlet Size</th>
<th>0%</th>
<th>14%</th>
<th>28%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Matrix Completion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>0%</td>
<td>17%</td>
<td>33%</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>0%</td>
<td>12.5%</td>
<td>25%</td>
</tr>
</tbody>
</table>
Are quoting patterns *systematic*?  
Do they correspond to *intuition*?  
Can *language* characterize the bias?
Are quoting patterns *systematic*?

Do they correspond to intuition?

Can language characterize the bias?

Yes! There is systematicity beyond newsworthiness.
Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

Can *language* characterize the bias?
Do they correspond to intuition?

\[ X \approx \hat{X} = U \cdot V^T \]

(outlet space)

(quote space)
Do they correspond to intuition?

\[ X \approx X = U \cdot VT \]

(outlet space)

(quote space)
Do they correspond to intuition?

The first two latent dimensions:

\[ \hat{X} = \begin{pmatrix} \vdots \\ \end{pmatrix} \cdot V^T \]
Do they correspond to intuition?

The first two latent dimensions:

\[ \hat{X} = \mathbf{X} \cdot \mathbf{V}^T \]

(outlet space)
Do they correspond to *intuition*?

The first two latent dimensions:

\[
\hat{\mathbf{X}} = \mathbf{X} \cdot \mathbf{V}^T
\]
Do they correspond to *intuition*?

The first two latent dimensions:

\[ \hat{X} = \begin{pmatrix} \text{outlet space} \end{pmatrix} \cdot V^T \]
Do they correspond to **intuition**?

The first two latent dimensions:

\[
\hat{X} = (\mathbf{V}) \cdot VT
\]
Does it correspond to *intuition*?

At the top of dimension 2:

- WEASEL ZIPPERS
- THE PATRIOT POST
- HOT AIR

Free Republic
Lonely Conservative
Right Wing News
...

37
Does it correspond to intuition?

At the top of dimension 2:

Free Republic
Lonely Conservative
Right Wing News
...

NATIONAL SECURITY
Who Is Watching Obama's Game of Drones?
Obamas drone program lacks transparency and oversight
Does it correspond to *intuition*?

At the top of dimension 2:

Free Republic
Lonely Conservative
Right Wing News
...

Around zero on dimension 2:

NY Post
The Economist
LA Times
...
Does it correspond to **intuition**?

At the top of dimension 2:

- Free Republic
- Lonely Conservative
- Right Wing News
- ... 

Around zero on dimension 2:

- NY Post
- The Economist
- LA Times
- ... 

*Supreme Court appears split in hearing on historic gay-marriage cases*

*As Supreme Court justices broke along familiar ideological lines, Justice Anthony M. Kennedy, a frequent swing vote, asked tough questions of both sides.*
Does it correspond to intuition?

At the top of dimension 2:

Free Republic
Lonely Conservative
Right Wing News
...

Around zero on dimension 2:

NY Post
The Economist
LA Times
...

Around the bottom of dimension 2:

Sydney Morning Herald
Ottawa Citizen
The Globe And Mail
...
Does it correspond to *intuition*?

At the top of dimension 2:

Barack Obama calls for US 'soul searching' as Baltimore braces for more violence

Police and National Guard troops flood the streets to try to head off a second night of rioting

Violence and vigilantism as Baltimore erupts into rioting

Around the bottom of dimension 2:

Sydney Morning Herald
Ottawa Citizen
The Globe And Mail
Are quoting patterns *systematic*? Yes!
Do they correspond to *intuition*?
Can *language* characterize the bias?
Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

Can they characterize bias?

Kind of. It’s not only liberal/conservative.
Are quoting patterns systematic? Yes!
Do they correspond to intuition? Kind of.
Can language characterize the bias?
Can *language* characterize the bias?

\[ X \approx \hat{X} = U \cdot V^T \]

\[
\begin{pmatrix}
\text{outlet space}
\end{pmatrix}
\approx
\begin{pmatrix}
\text{quote space}
\end{pmatrix}
\]
Can *language* characterize the bias?

\[ X \approx \hat{X} = U \cdot V^T \]

\text{outlet space} \quad \text{quote space}
Can *language* characterize the bias?

\[ \hat{X} = U \cdot \left( \begin{array}{c} \text{quote space} \end{array} \right) \]

(The same axes!)
Can *language* characterize the bias?

If we don’t work even harder than we did in 2008, then we’re going to have a government that tells the American people, “you’re on your own.”

\[ \hat{X} = U \cdot \]
Can *language* characterize the bias?

\[ \hat{X} = U \cdot (\text{quote space}) \]

If we don’t work even harder than we did in 2008, then we’re going to have a government that tells the American people, “you’re on your own.”

The truth is, you could figure out on the back of an envelope how to get this done. The question is one of political will.
Can *language* characterize the bias?

If we don’t work even harder than we did in 2008, then we’re going to have a government that tells the American people, “you’re on your own.”

The truth is, you could figure out on the back of an envelope how to get this done. The question is one of political will.

By the end of the next year, all U.S. troops will be out of Iraq.
Can *language* characterize the bias?

Average sentiment of quotes:

0.02

0.01

0.00

-0.01

-0.02

(The same axes!)
Can *language* characterize the bias?

Average sentiment of quotes:

(The same axes!)
Can *language* characterize the bias?

**Average sentiment of quotes:**

**Amount of negation in quotes:**

(The same axes!)
Conclusion

Quoting patterns are systematic:
We can predict them.

Online visualization & data:
http://snap.stanford.edu/quotus/
Conclusion

Quoting patterns are systematic: We can predict them.

Selection bias is multidimensional.

Language plays a part: Sentiment and negation show alignment.

Online visualization & data: http://snap.stanford.edu/quotus/