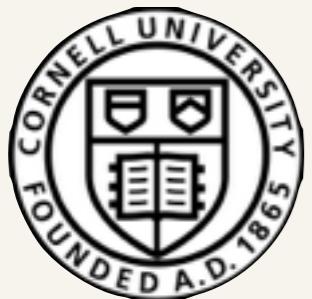


The structure of political media coverage
as revealed by quoting patterns:

QUOTing the POTUS



*equal
contribution* {

Cristian Danescu-Niculescu-Mizil

***Vlad Niculae** (me)

*Caroline Suen

*Justine Zhang

Jure Leskovec

Cornell (MPI-SWS)

Stanford

Cornell (Stanford, MPI-SWS)

Cornell (MPI-SWS)

Stanford

The press claims to be
fair and balanced.

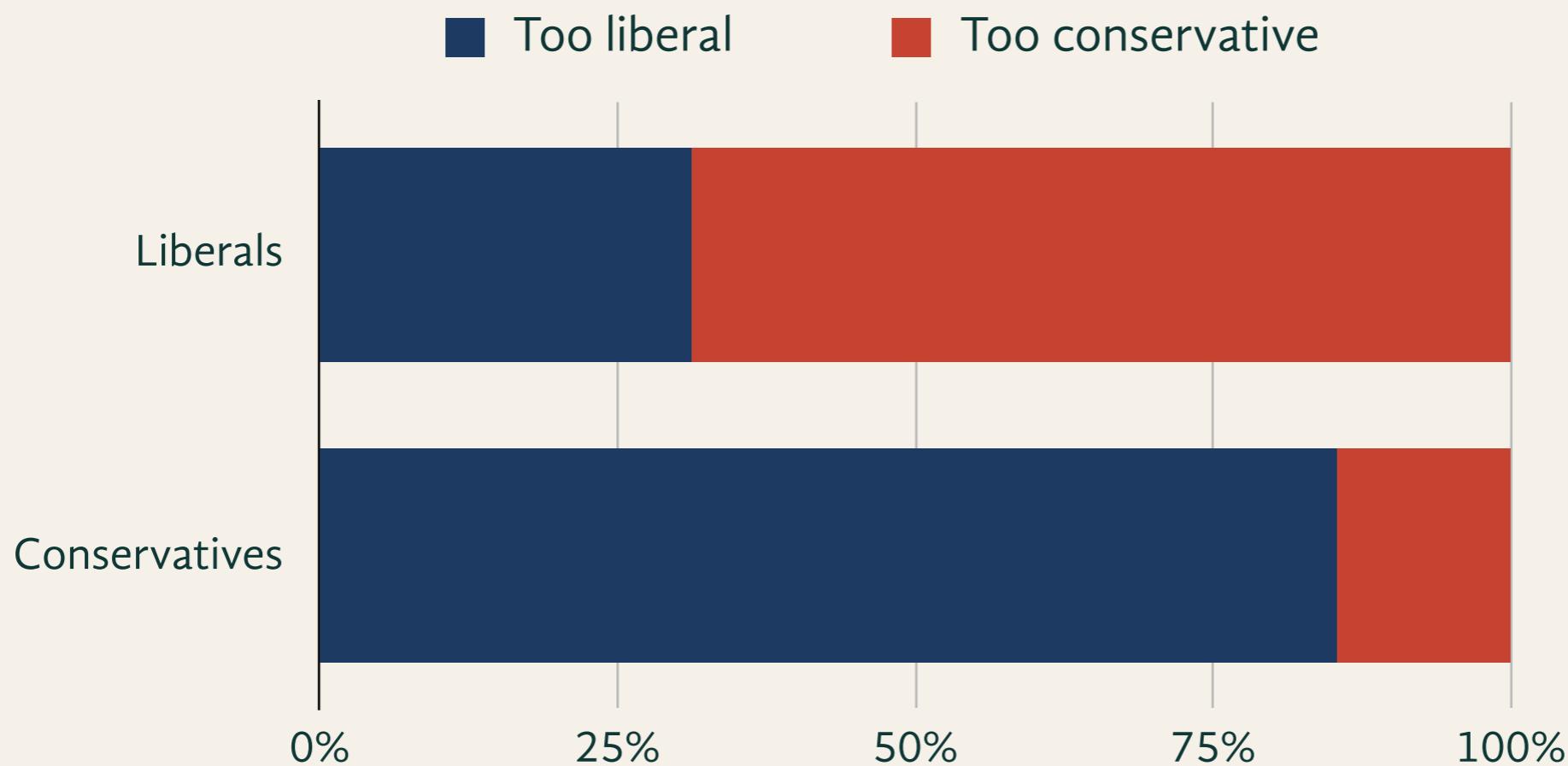
America's trust
in media is at
an all time low.

(Source: Gallup poll, Sept. 17 2014)

Tea Party Anti-Tax March, Sep 12, 2010.
Capital Mall, DC
Photo Credit: [Bob Simmons](#)
license CC-BY-NC-ND 2.0



“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]

[Gentzkow et al, 2010]
[Groseclose et al, 2005]
[Lin et al, 2011]

[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?





Photo Credit: [DonkeyHotey](#)
license CC-BY 2.0

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.

—Right Wing News

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*?

How do the patterns manifest in *language*?

Are quoting patterns *systematic*?

Do they correspond to *intuition*?

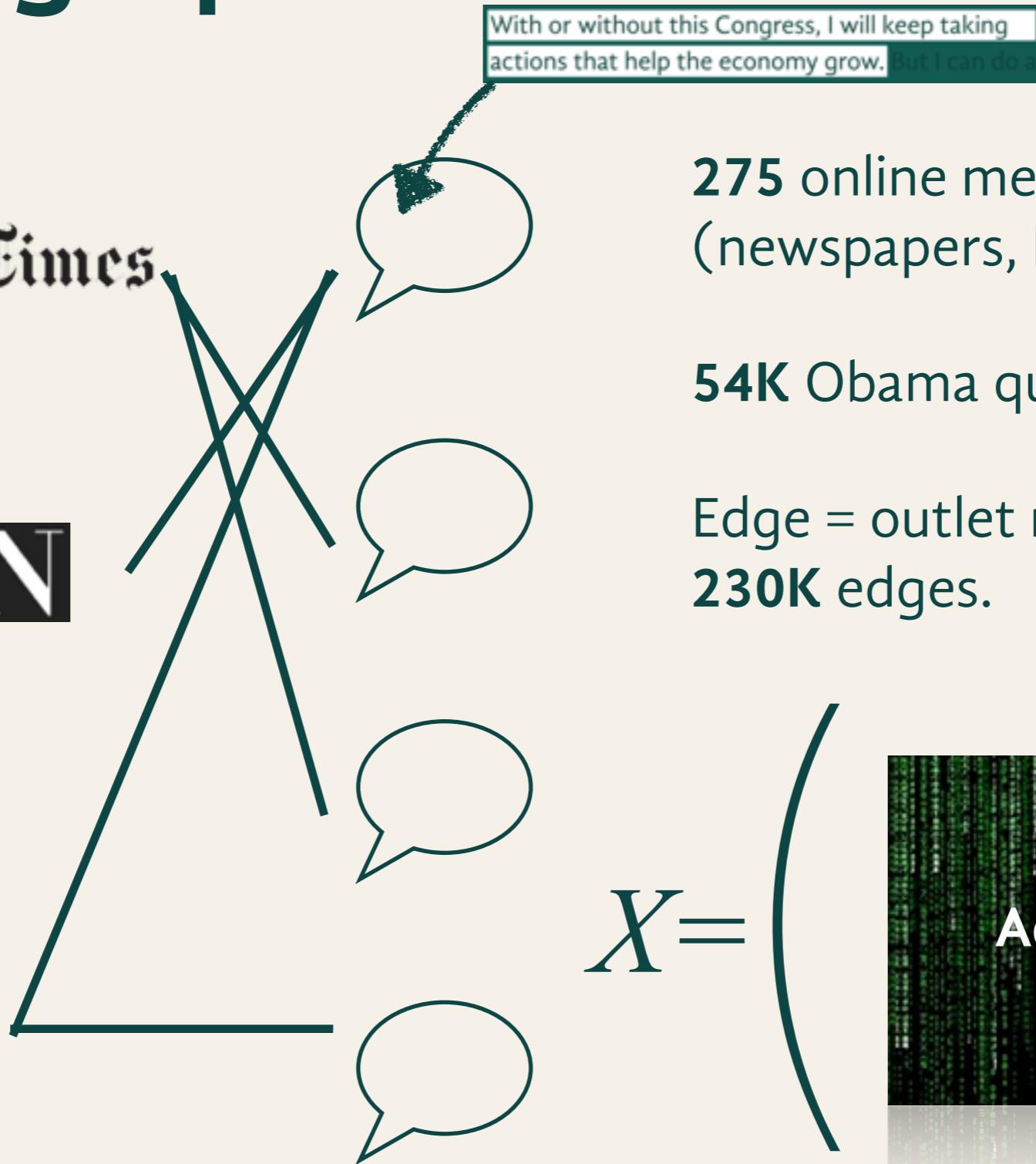
How do the patterns manifest in *language*?

Are quoting patterns *systematic*?

Quoting graph

The New York Times

SALON



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

54K Obama quotes in 6 years.

Edge = outlet reported quote.
230K edges.

$$X = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

The
Adjacency
Matrix

A large green matrix with binary code (0s and 1s) filling most of the space. In the center, the text "The Adjacency Matrix" is written in white. The matrix is enclosed in a large black bracket on the right side.

Are quoting patterns *systematic*?

Can we predict quoting?

If we can predict quoting choices,
it means there is a pattern.

The New York Times

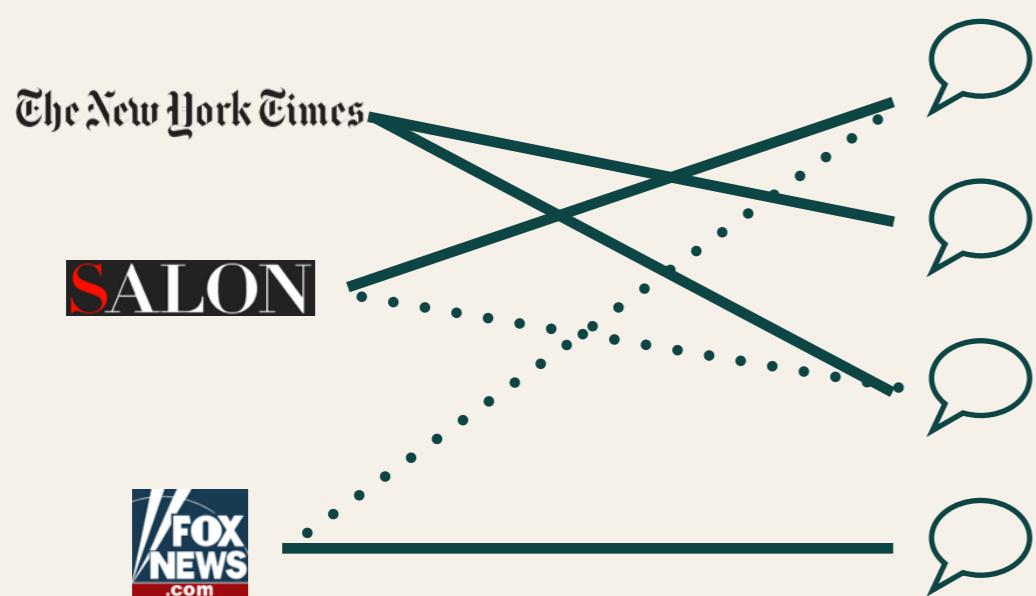
SALON



Are quoting patterns *systematic*?

Can we predict quoting?

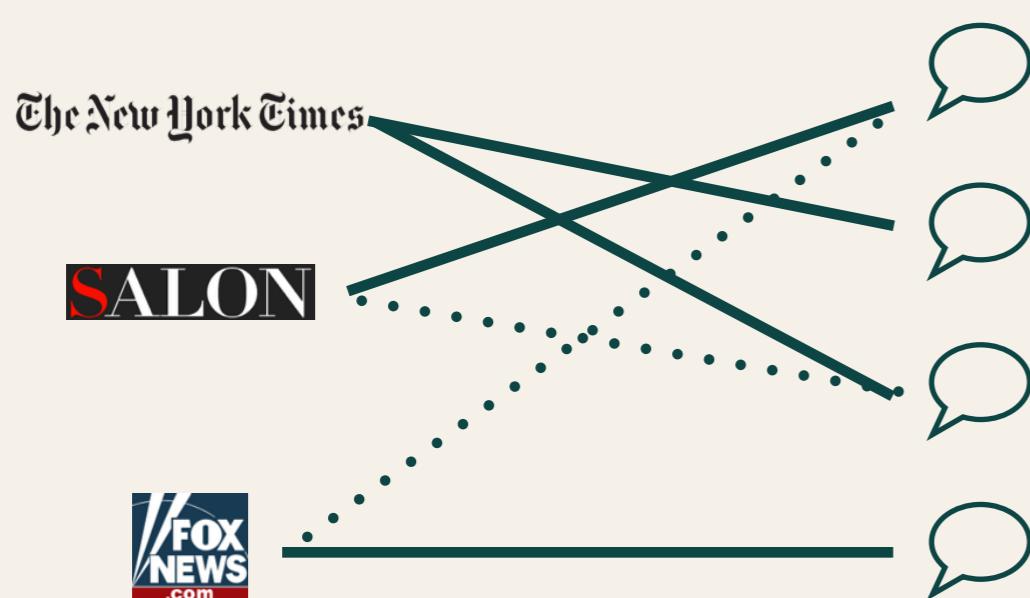
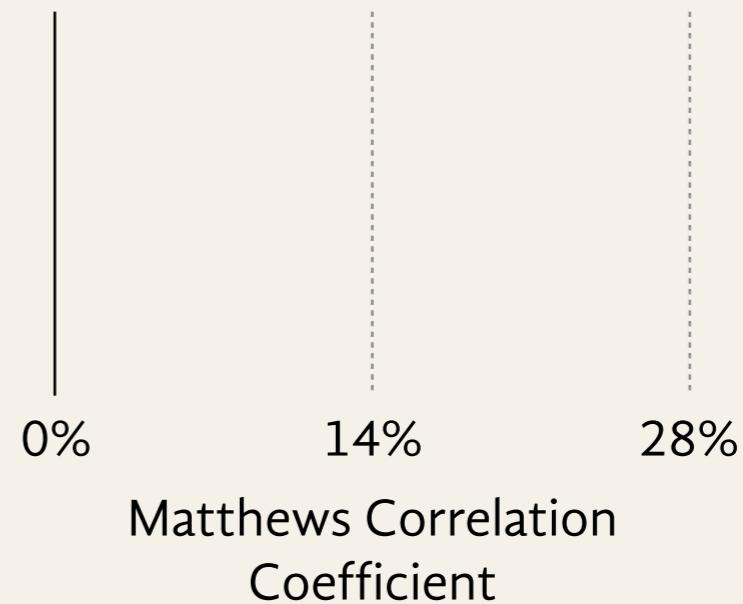
Leave out 250K entries of X .



Are quoting patterns *systematic*?

Can we predict quoting?

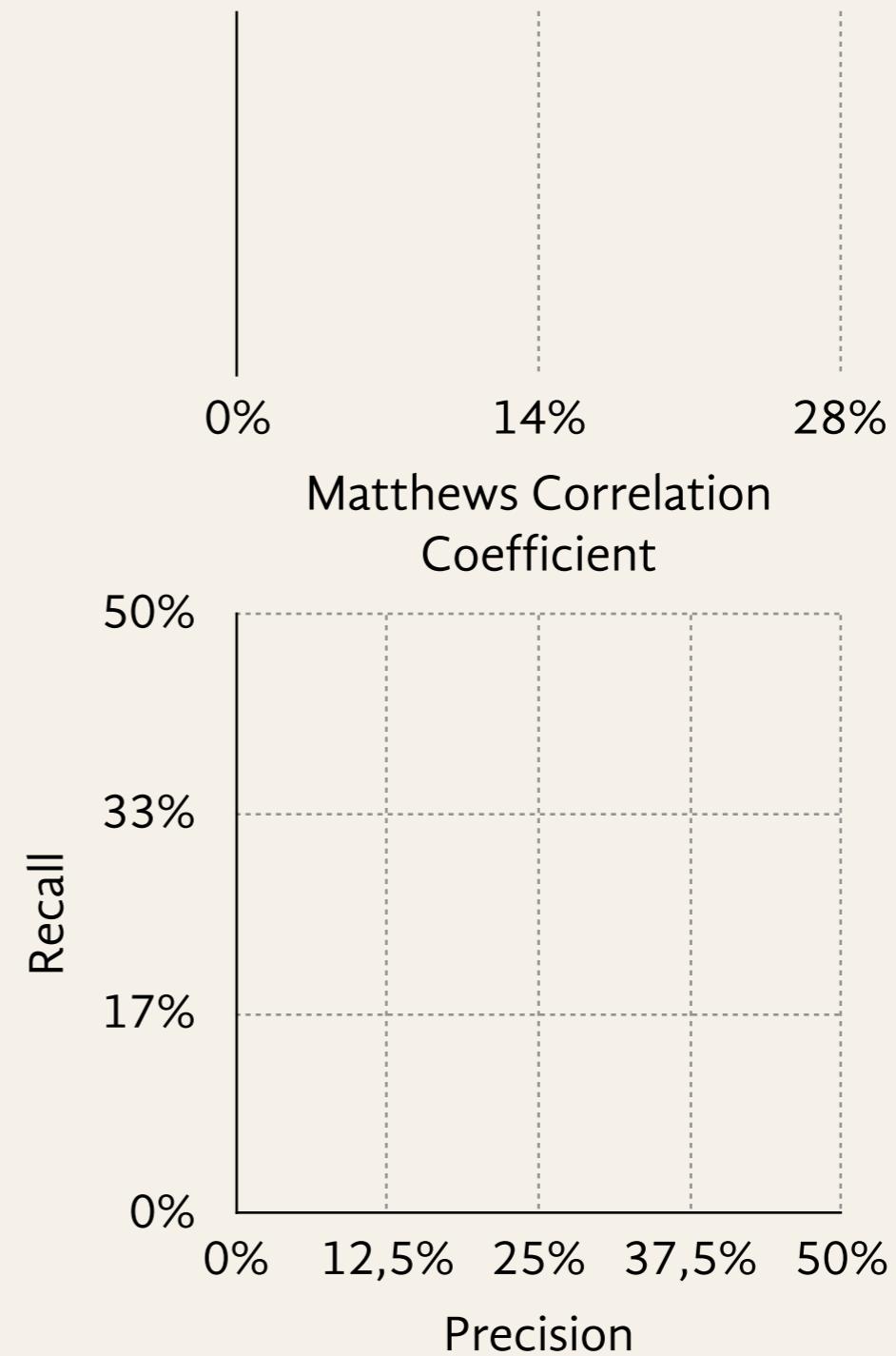
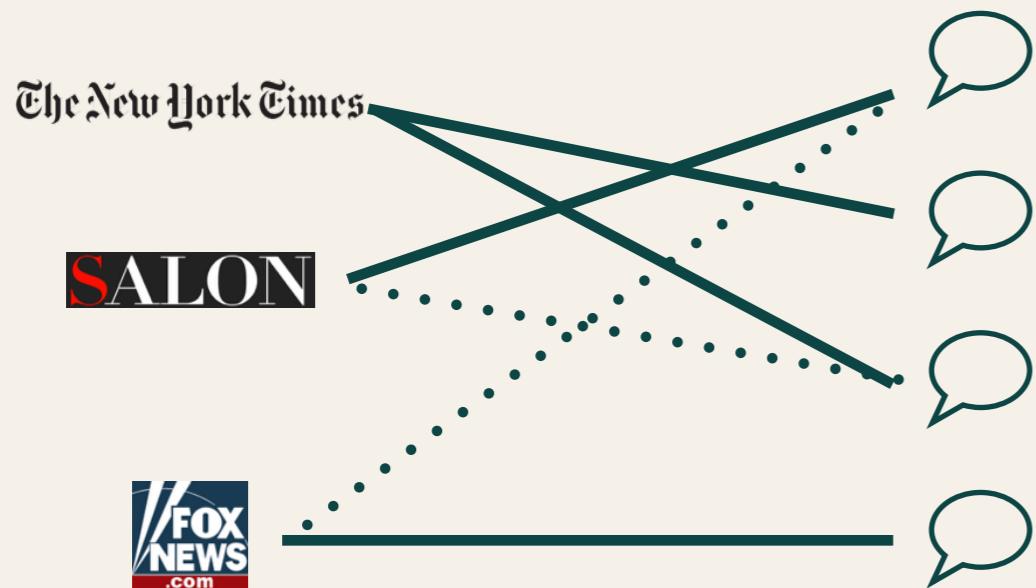
Leave out 250K entries of X .



Are quoting patterns *systematic*?

Can we predict quoting?

Leave out 250K entries of X .



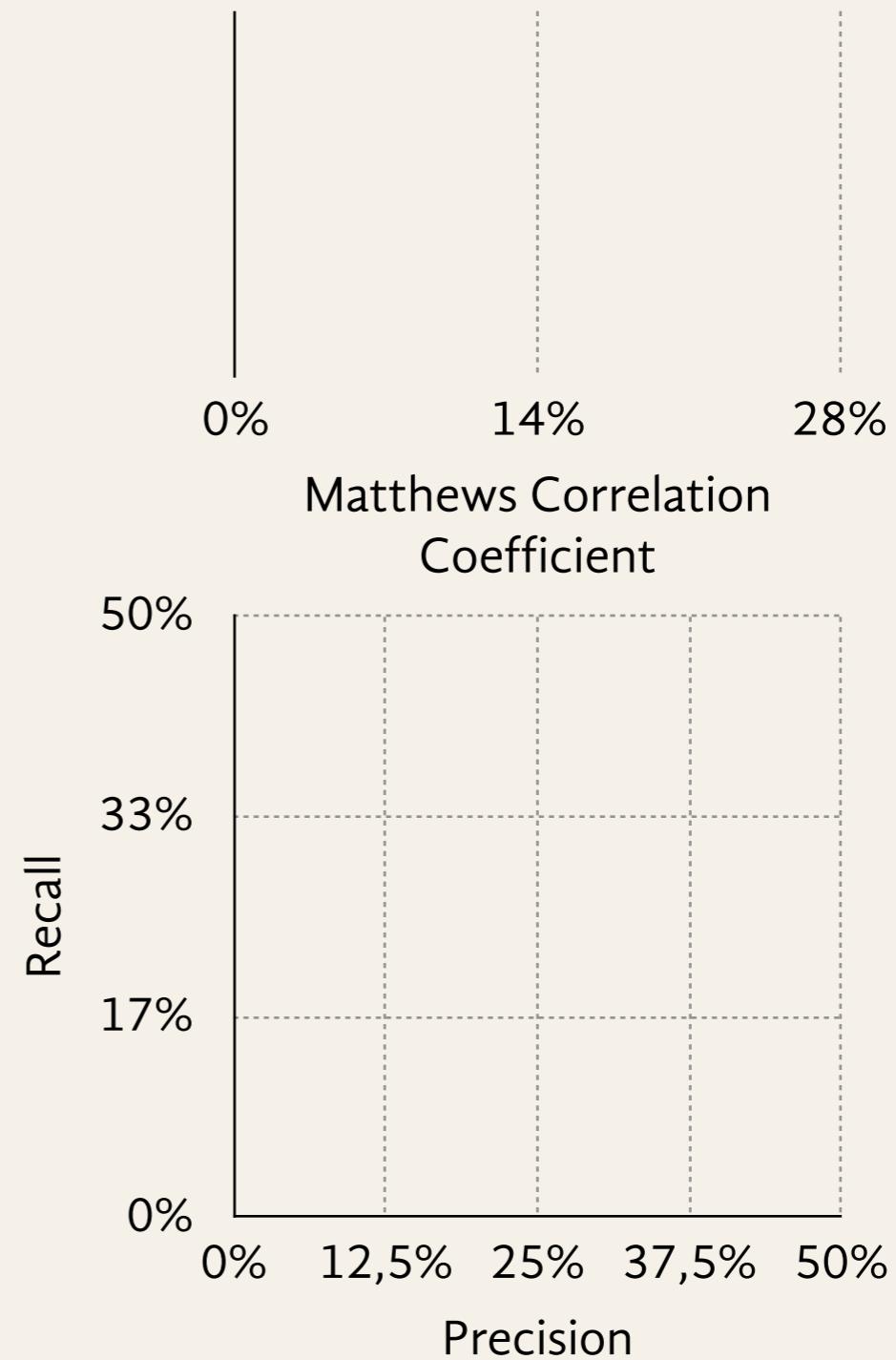
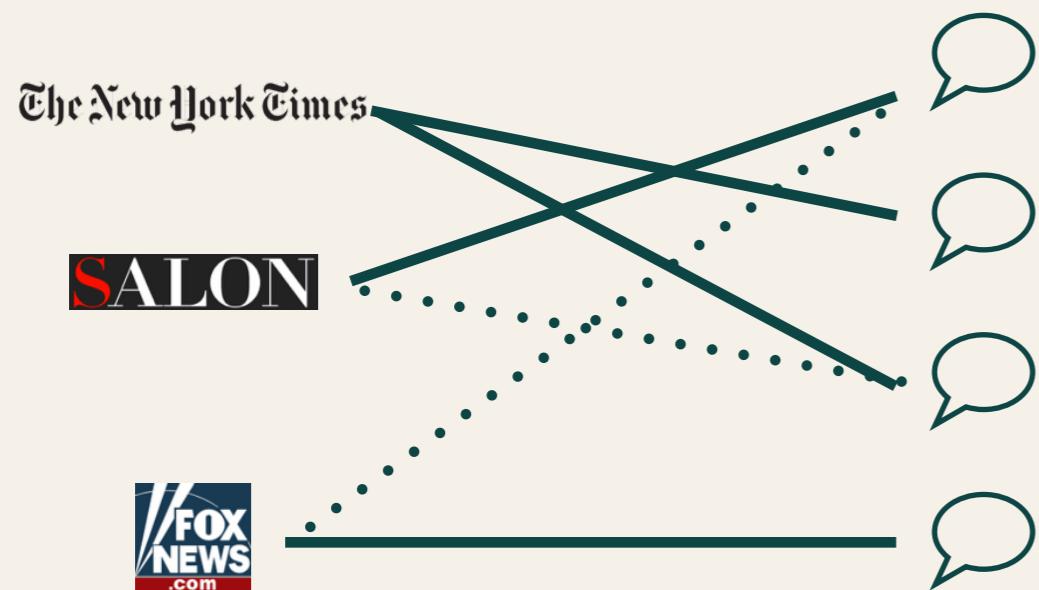
Are quoting patterns *systematic*?

Can we predict quoting?

Leave out 250K entries of X .

In a “fair world” without bias:

- popular quotes get cited more,



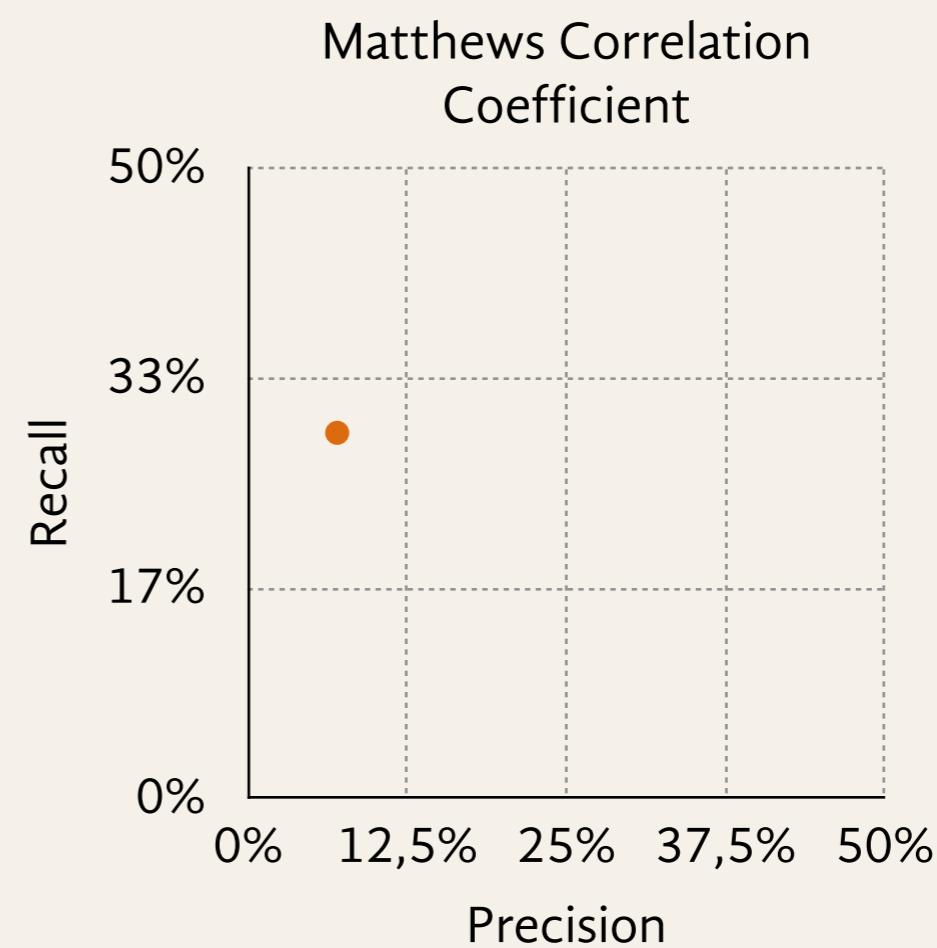
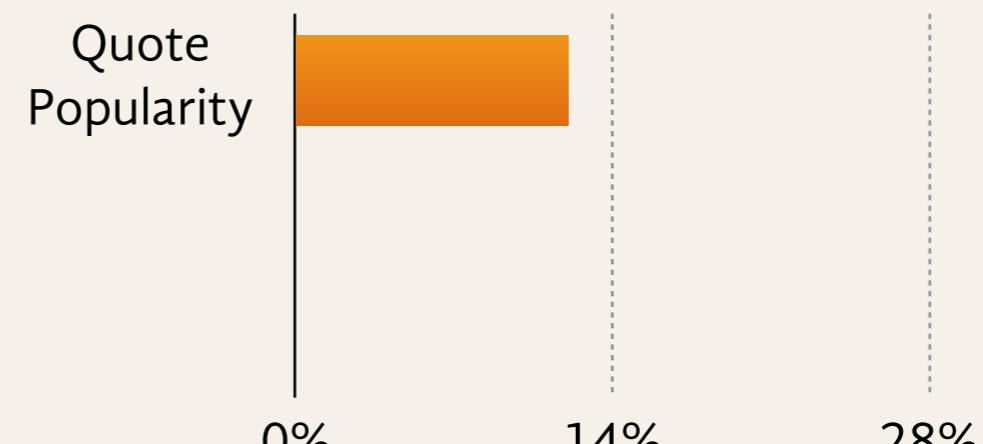
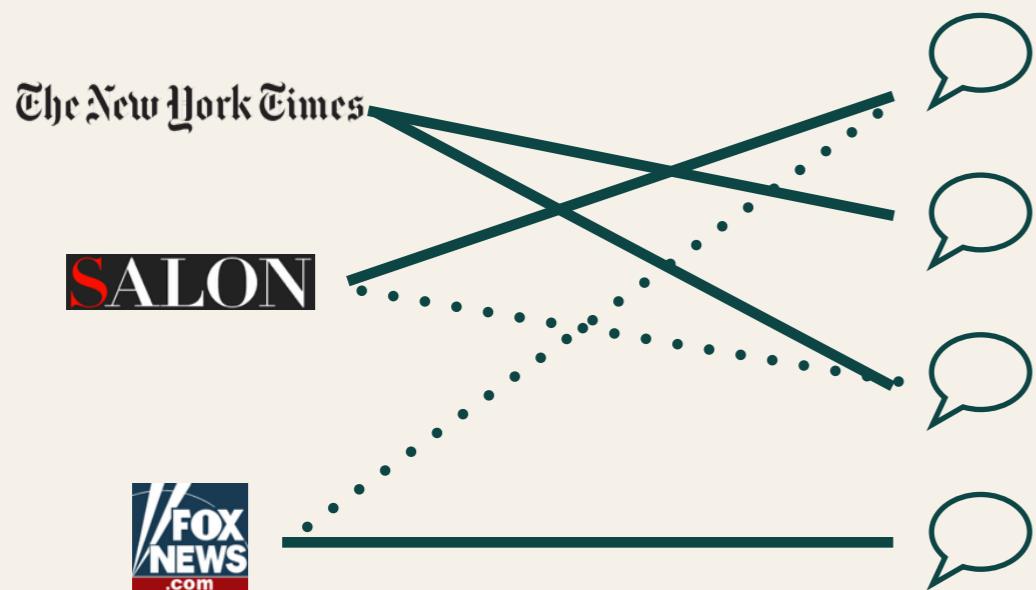
Are quoting patterns *systematic*?

Can we predict quoting?

Leave out 250K entries of X .

In a “fair world” without bias:

- popular quotes get cited more,



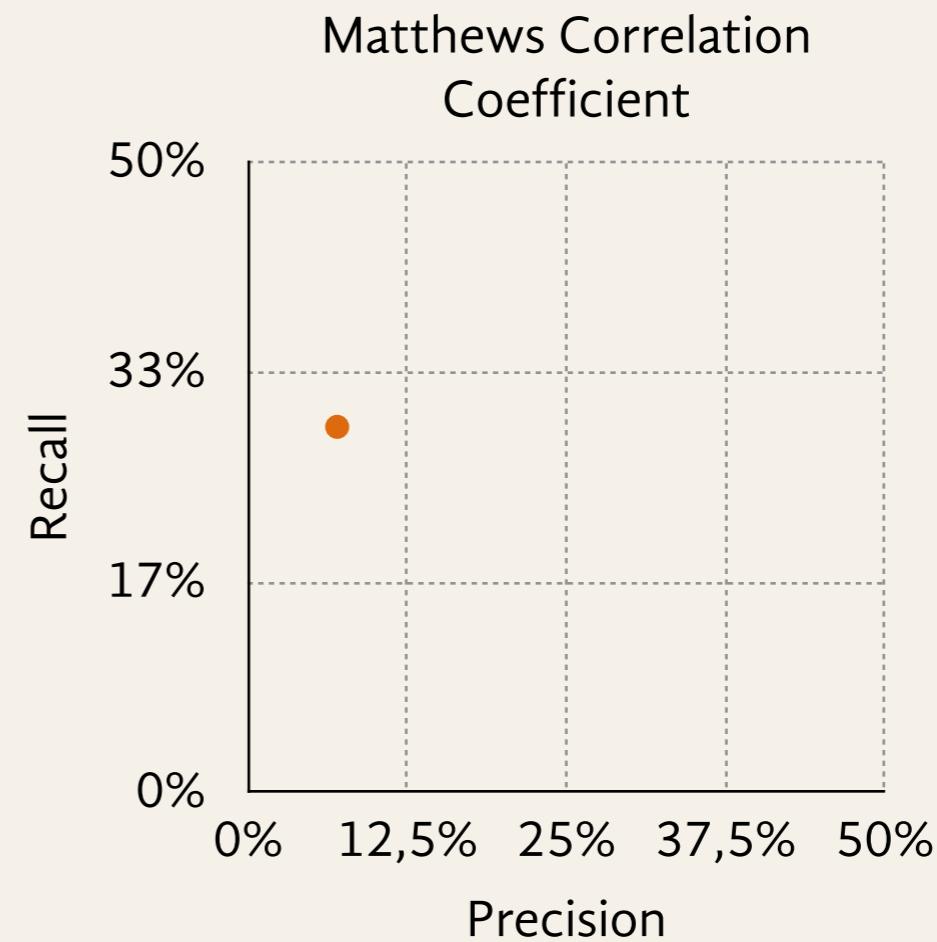
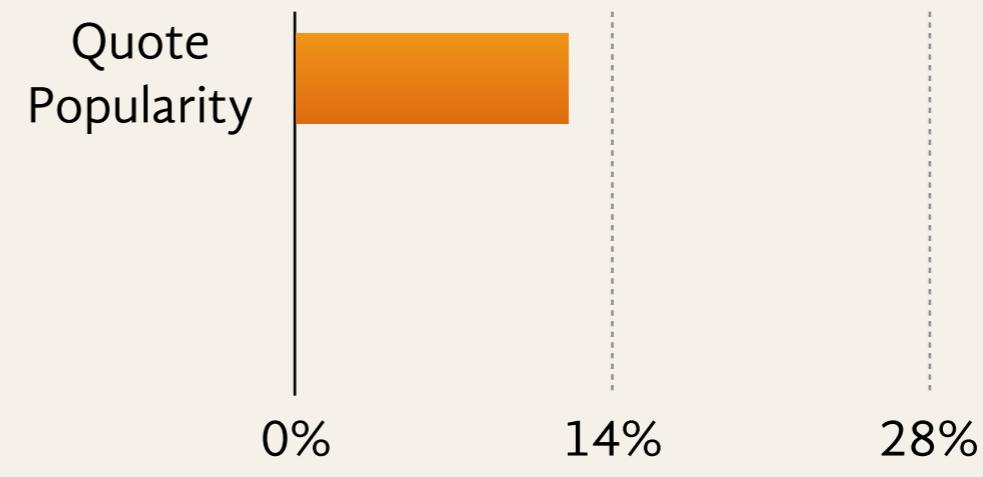
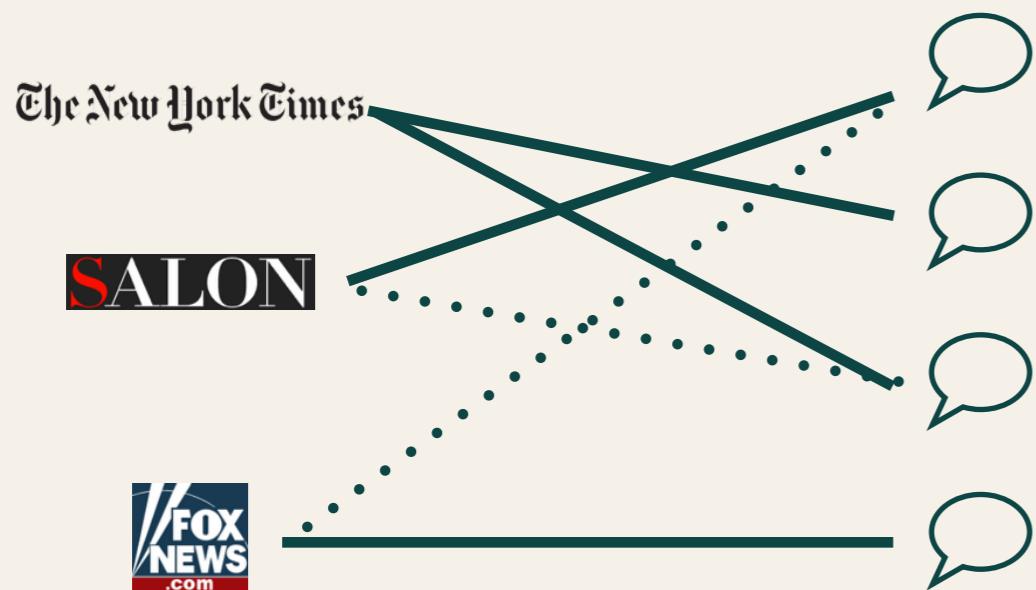
Are quoting patterns *systematic*?

Can we predict quoting?

Leave out 250K entries of X .

In a “fair world” without bias:

- popular quotes get cited more,
- and big outlets quote more.



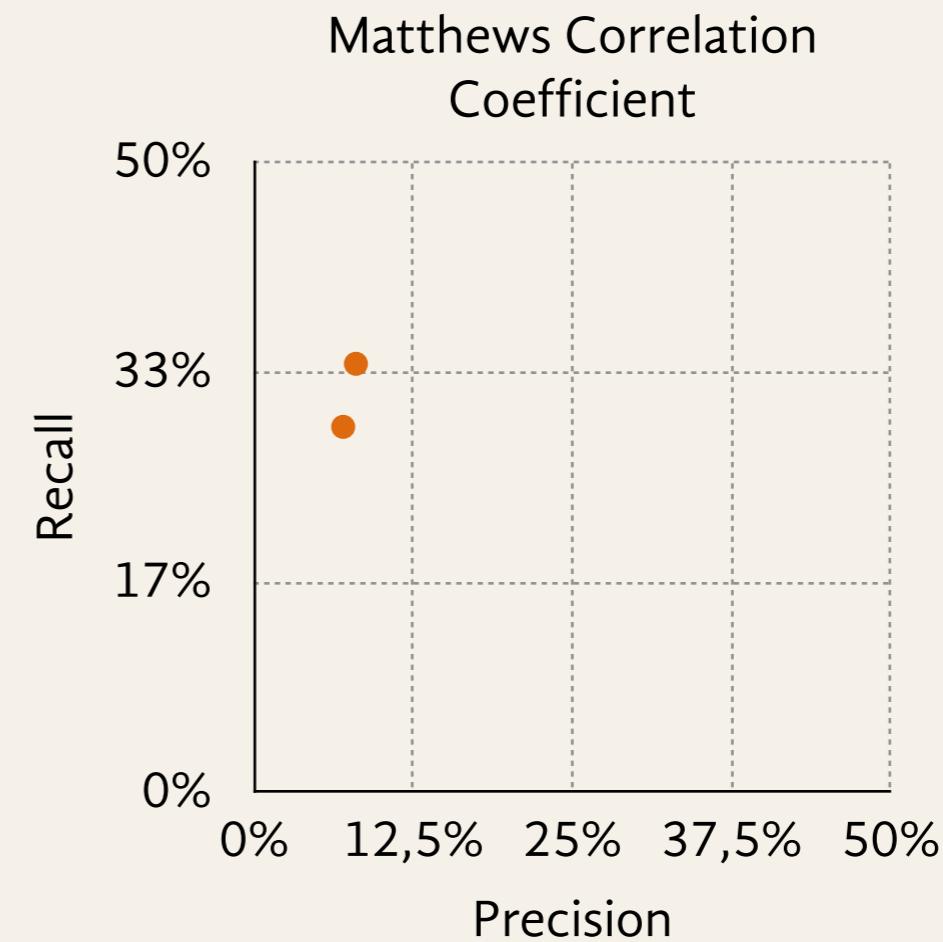
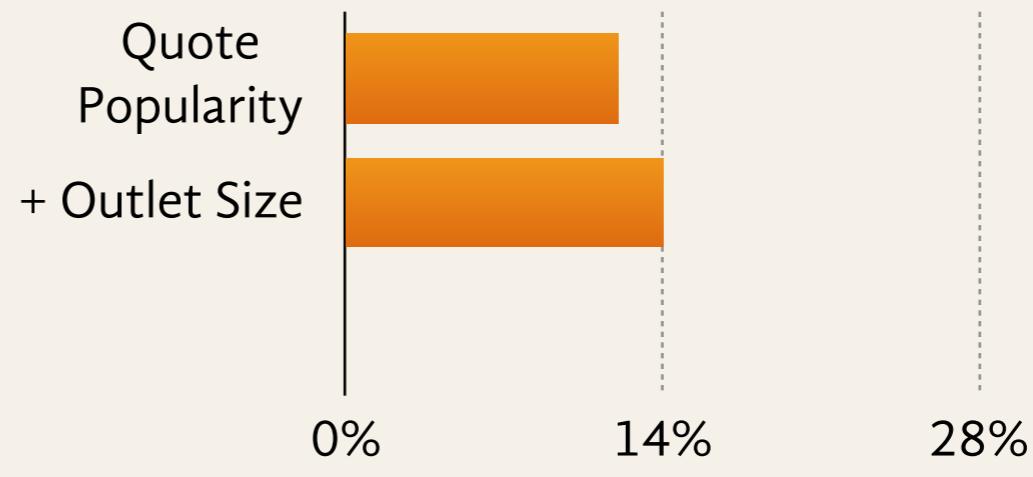
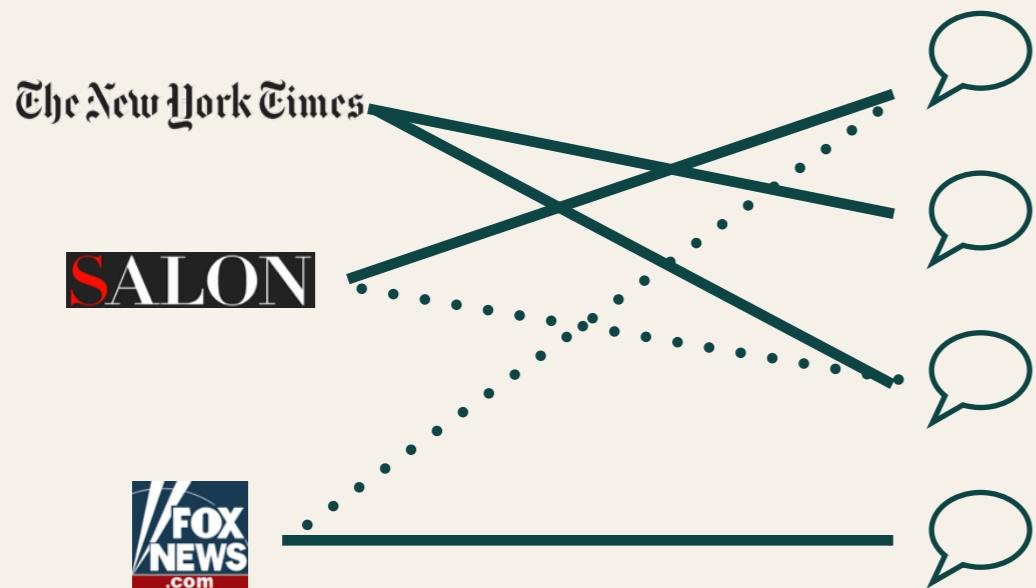
Are quoting patterns *systematic*?

Can we predict quoting?

Leave out 250K entries of X .

In a “fair world” without bias:

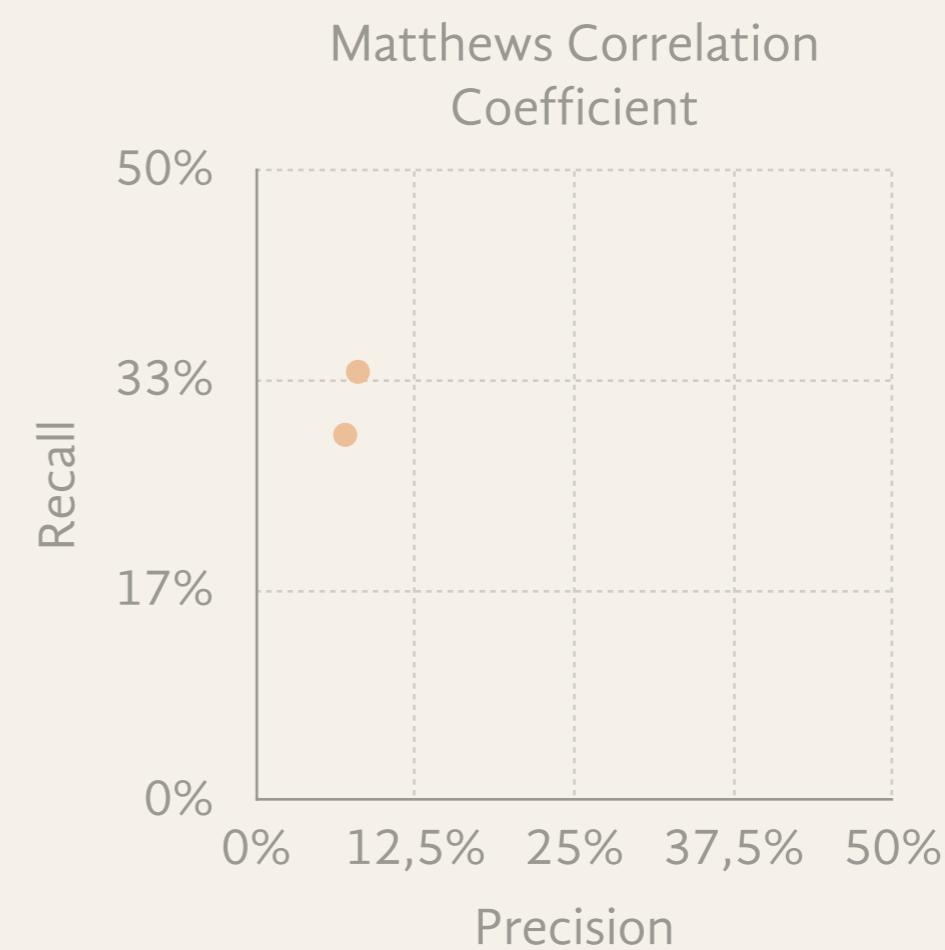
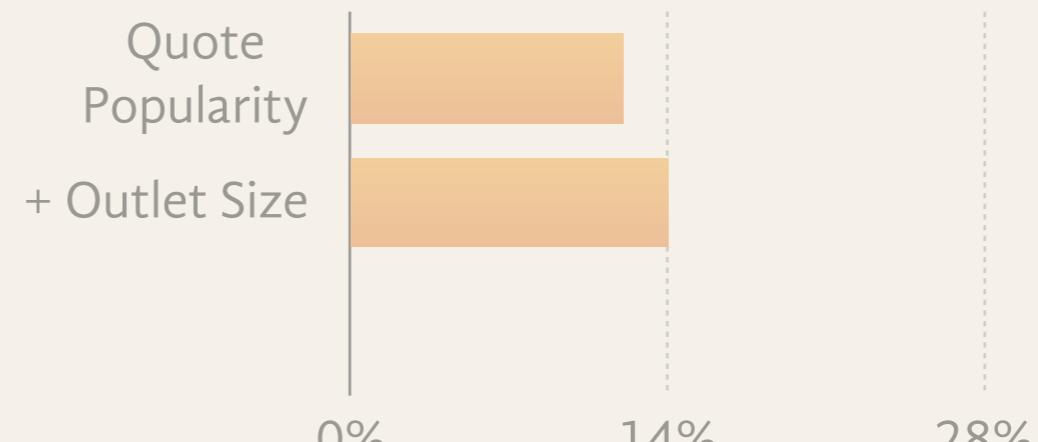
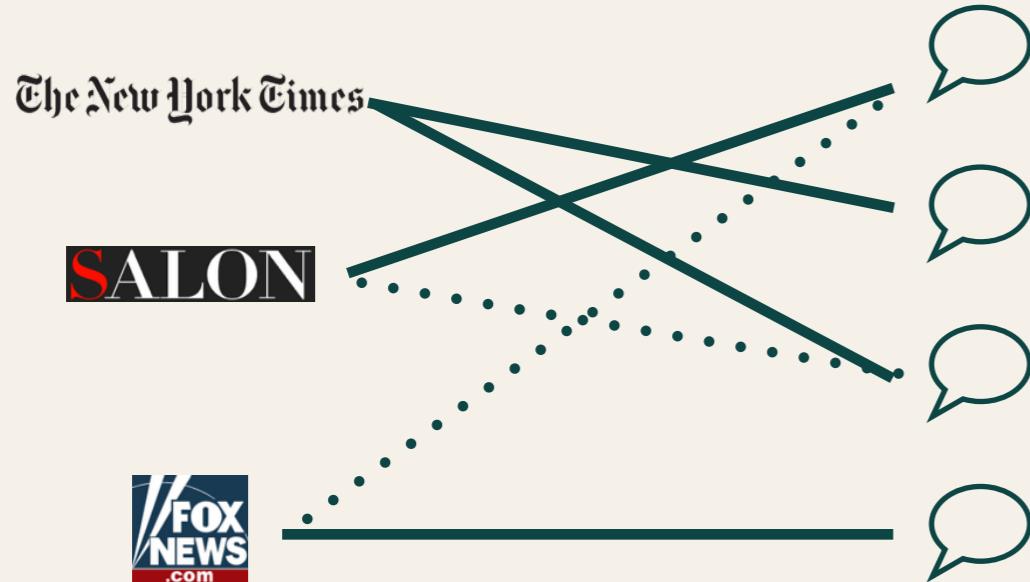
- popular quotes get cited more,
- and big outlets quote more.



Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

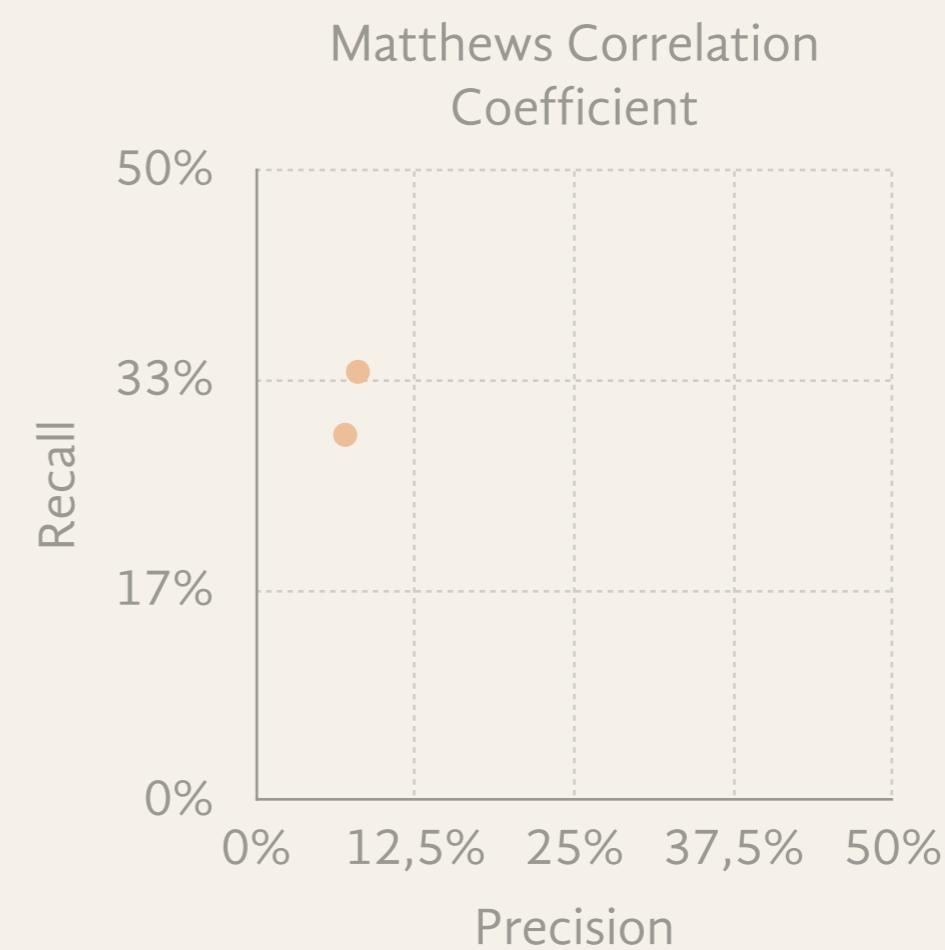
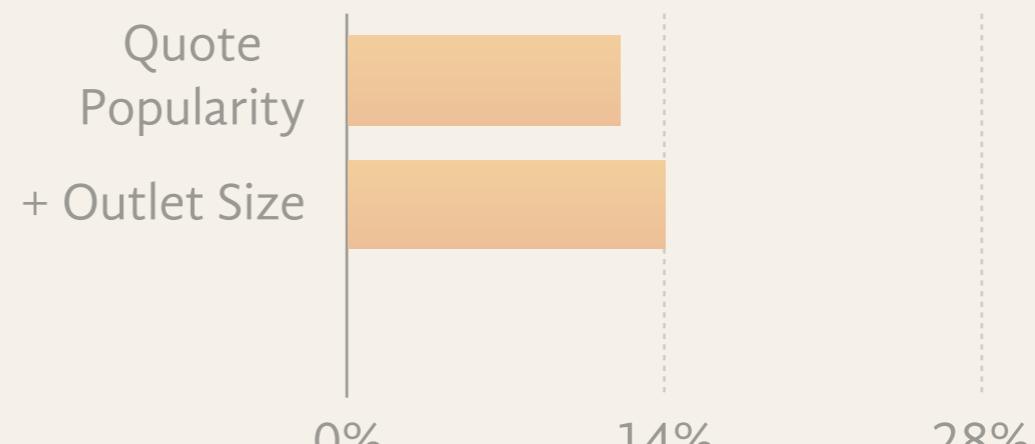
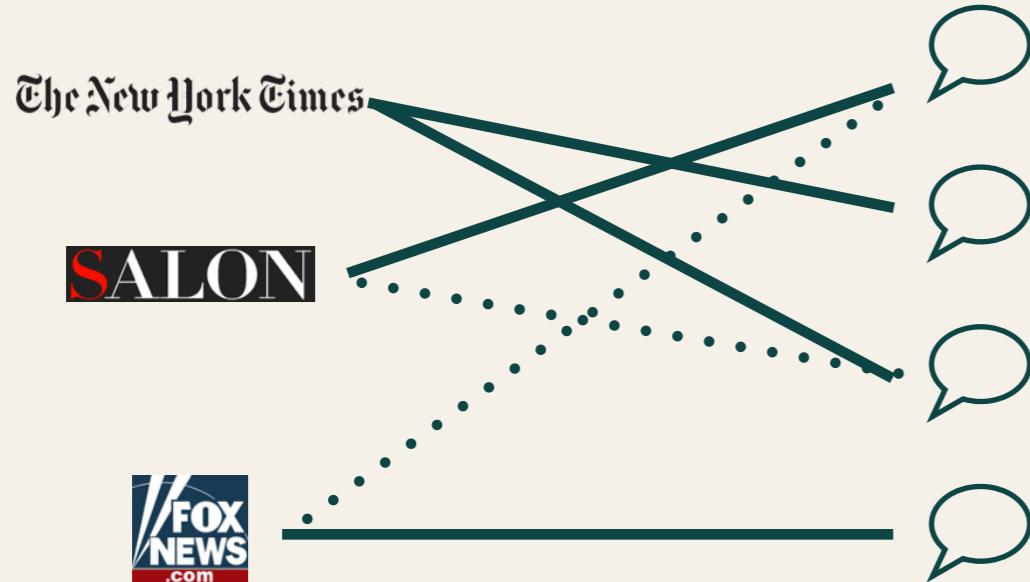


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

$$X \approx \hat{X} = U \cdot VT$$

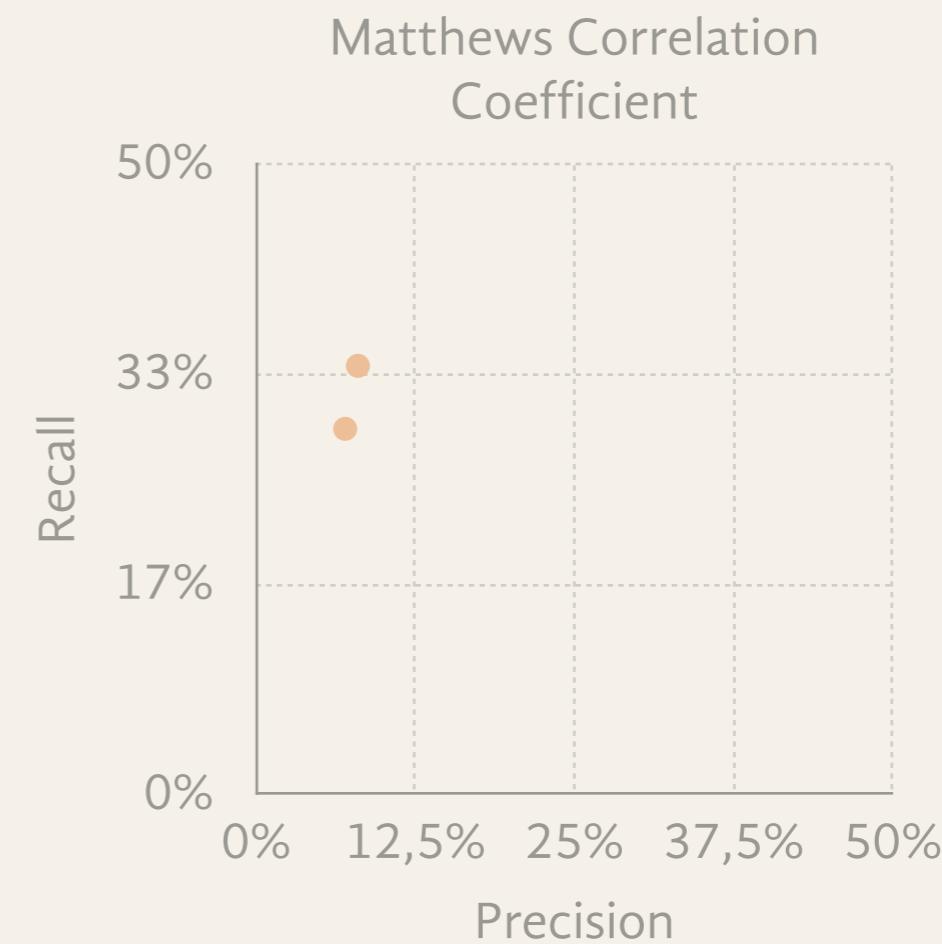
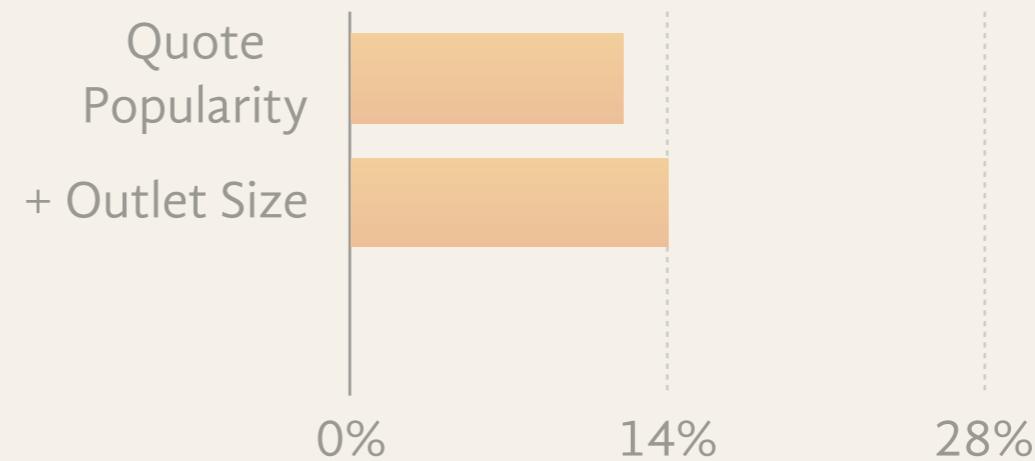
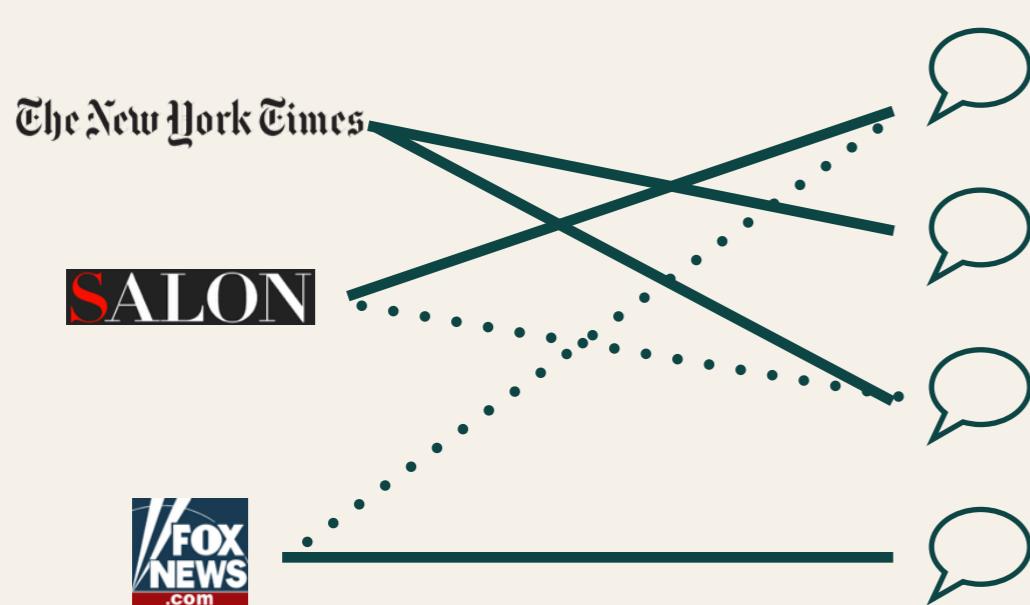
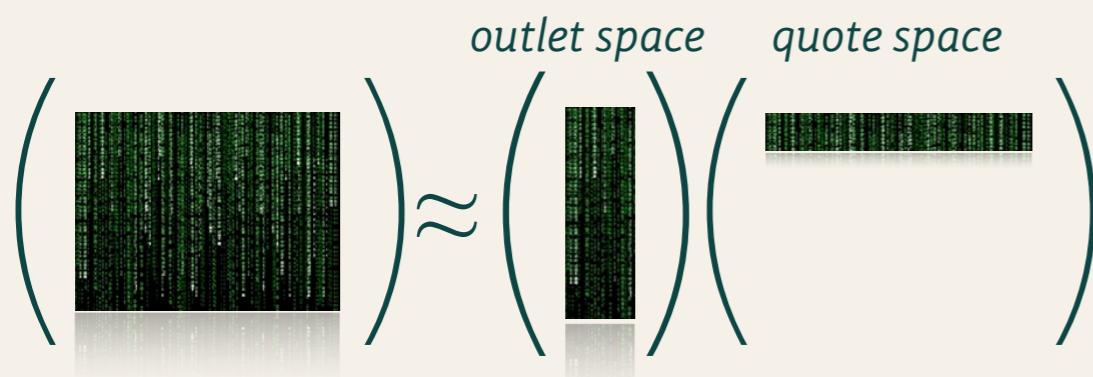


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

$$X \approx \hat{X} = U \cdot VT$$

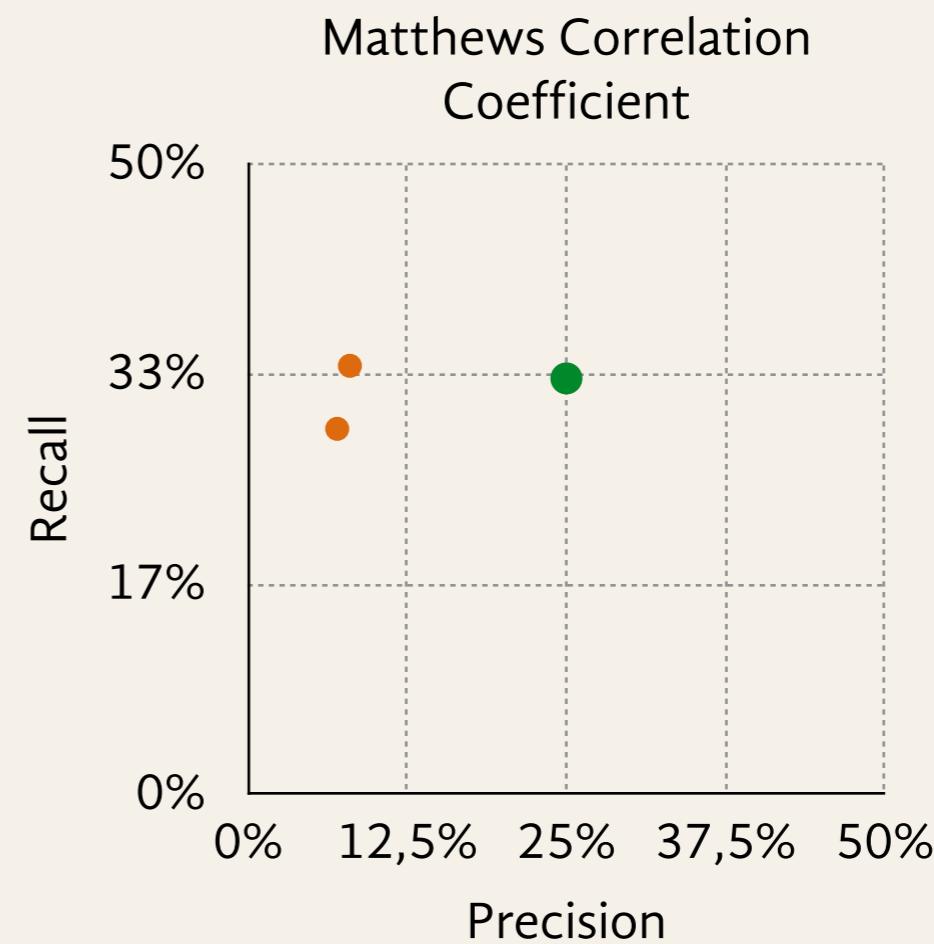
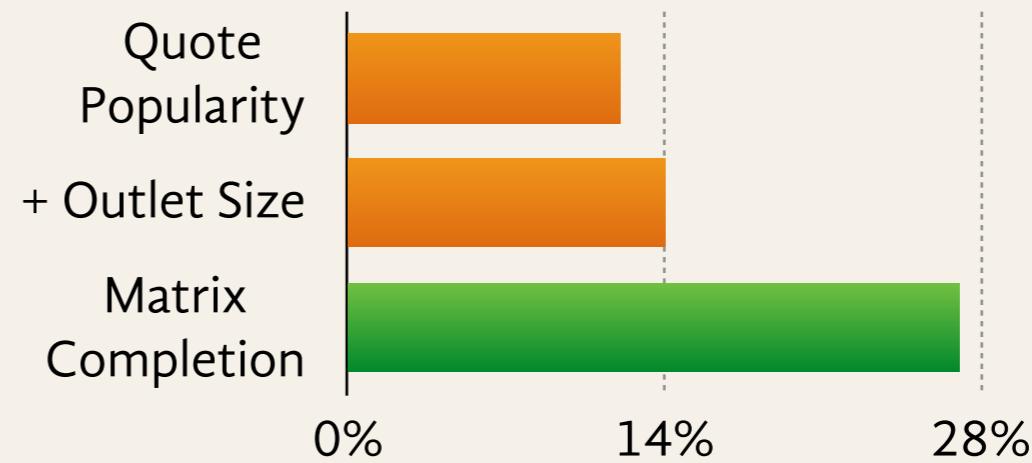
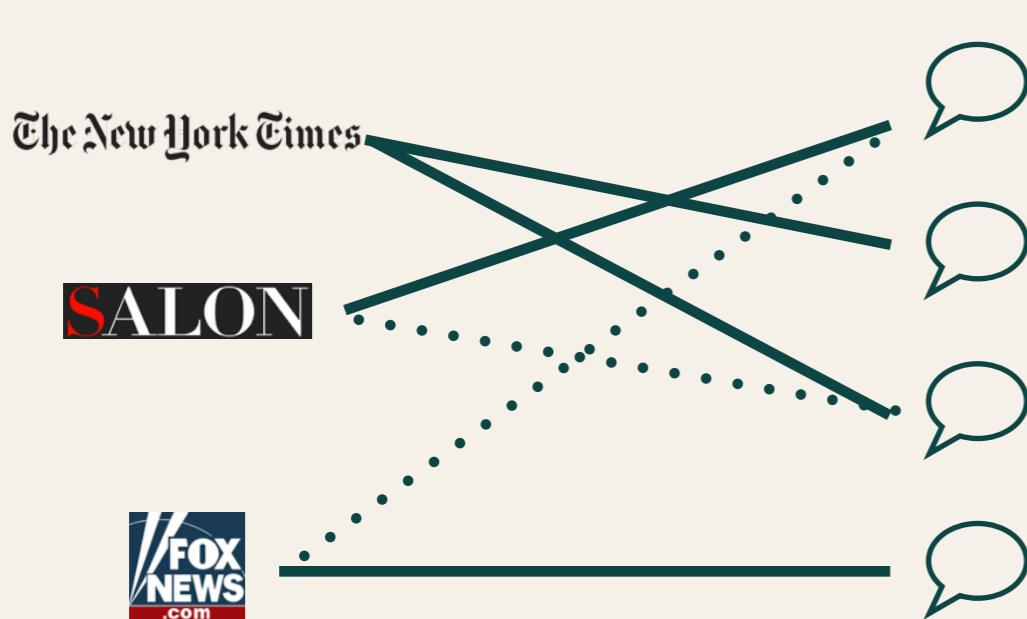
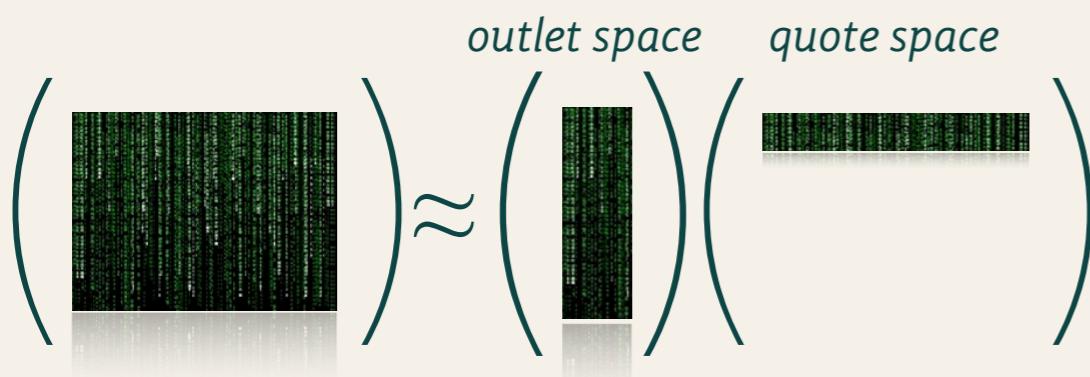


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by low-rank factorization

$$X \approx \hat{X} = U \cdot VT$$



Are quoting patterns *systematic*?

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Are quoting patterns *systematic*?

Do

Yes!

How do

**There is systematicity
beyond newsworthiness.**

on?

language?

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Do they correspond to *intuition*?

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

$$\left(\begin{array}{c} \text{matrix code} \\ \text{matrix code} \\ \text{matrix code} \end{array} \right) \approx \left(\begin{array}{c} \text{outlet space} \\ \text{matrix code} \end{array} \right) \left(\begin{array}{c} \text{quote space} \\ \text{matrix code} \end{array} \right)$$

Do they correspond to *intuition*?

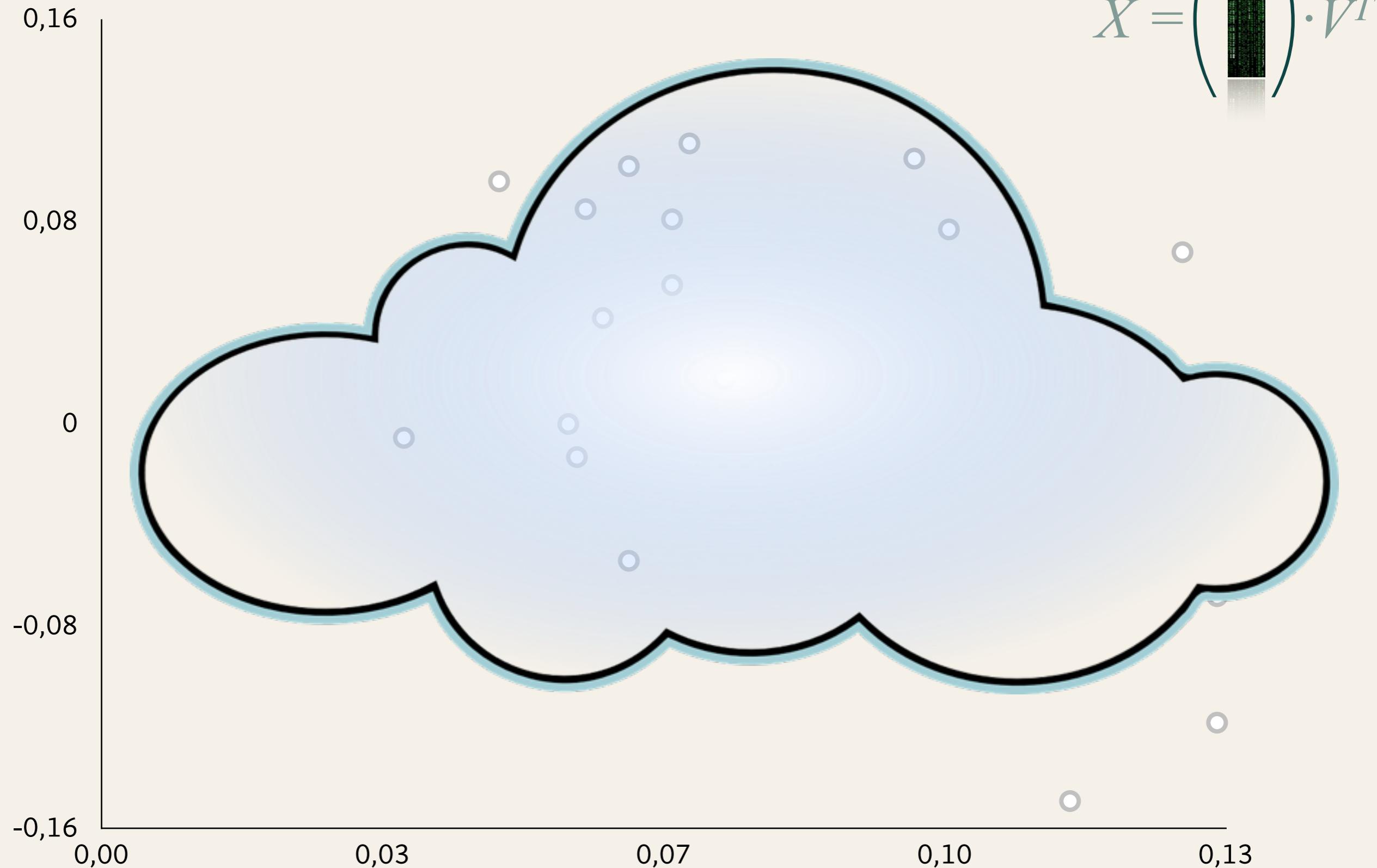
$$X \approx \hat{X} = U \cdot Y^T$$

outlet space *quote space*

The diagram shows a matrix X being approximated by \hat{X} . The decomposition is given by $X \approx \hat{X} = U \cdot Y^T$, where U represents the 'outlet space' and Y^T represents the 'quote space'. The matrix X is shown as a large grey rectangle. The approximation is indicated by a double-headed arrow between X and \hat{X} . The matrix U is represented by a green vertical rectangle containing binary code, and the matrix Y^T is represented by a grey horizontal rectangle containing binary code.

Do they correspond to *intuition*?

The first two latent dimensions:



Do they correspond to *intuition*?

The first two latent dimensions:

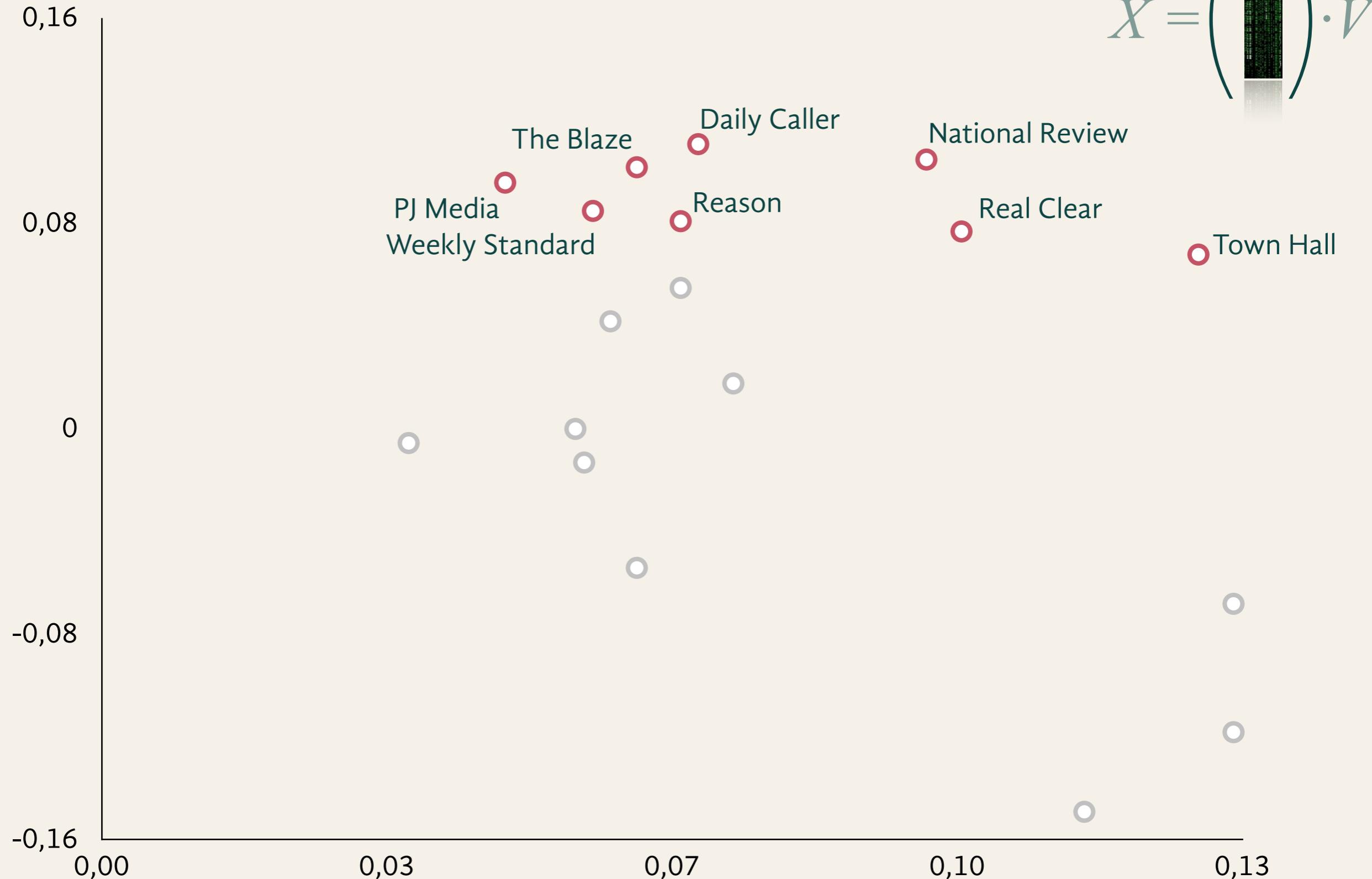


outlet space

$$\hat{X} = \begin{pmatrix} \text{matrix icon} \end{pmatrix} \cdot VT$$

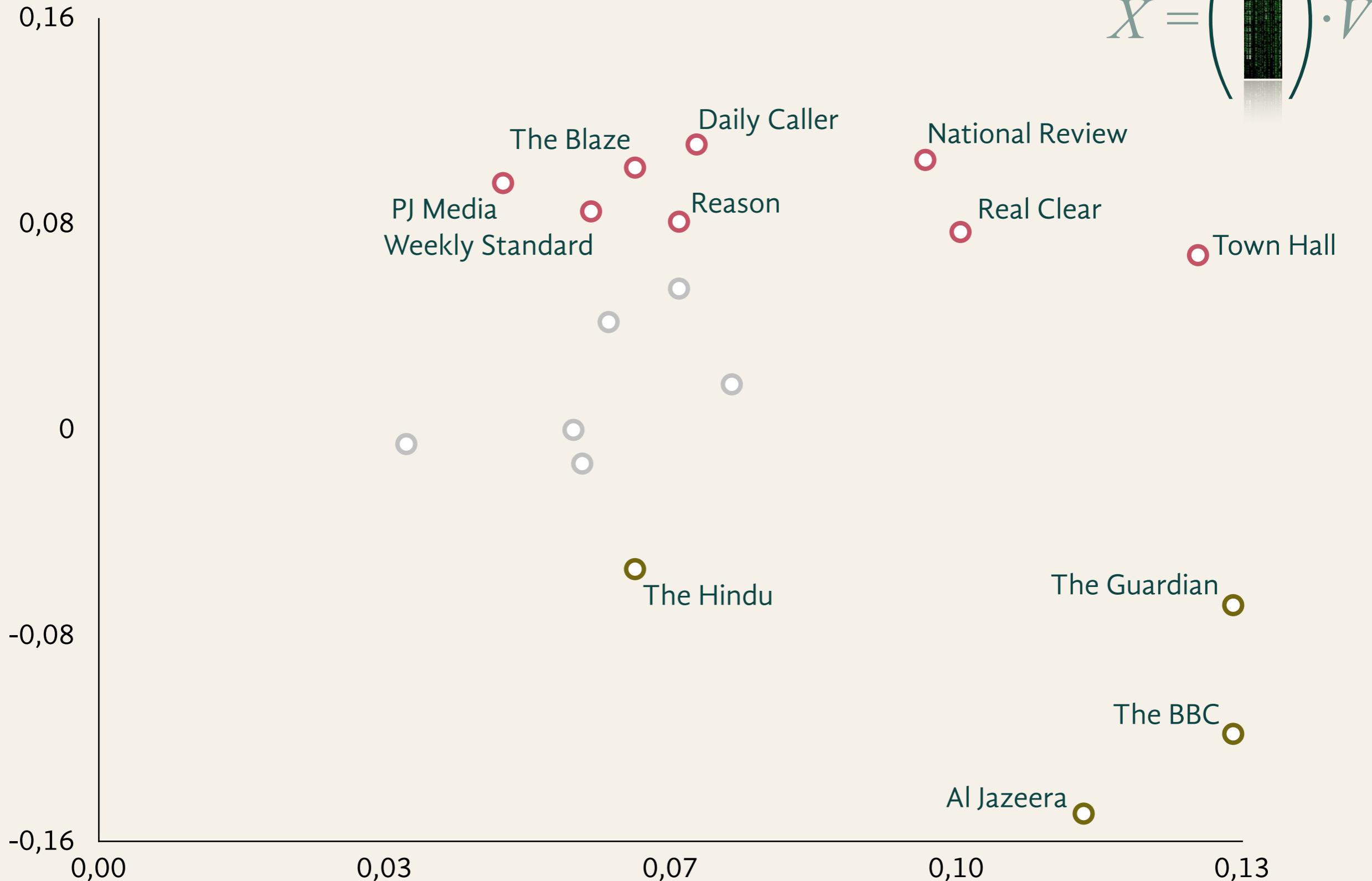
Do they correspond to *intuition*?

The first two latent dimensions:



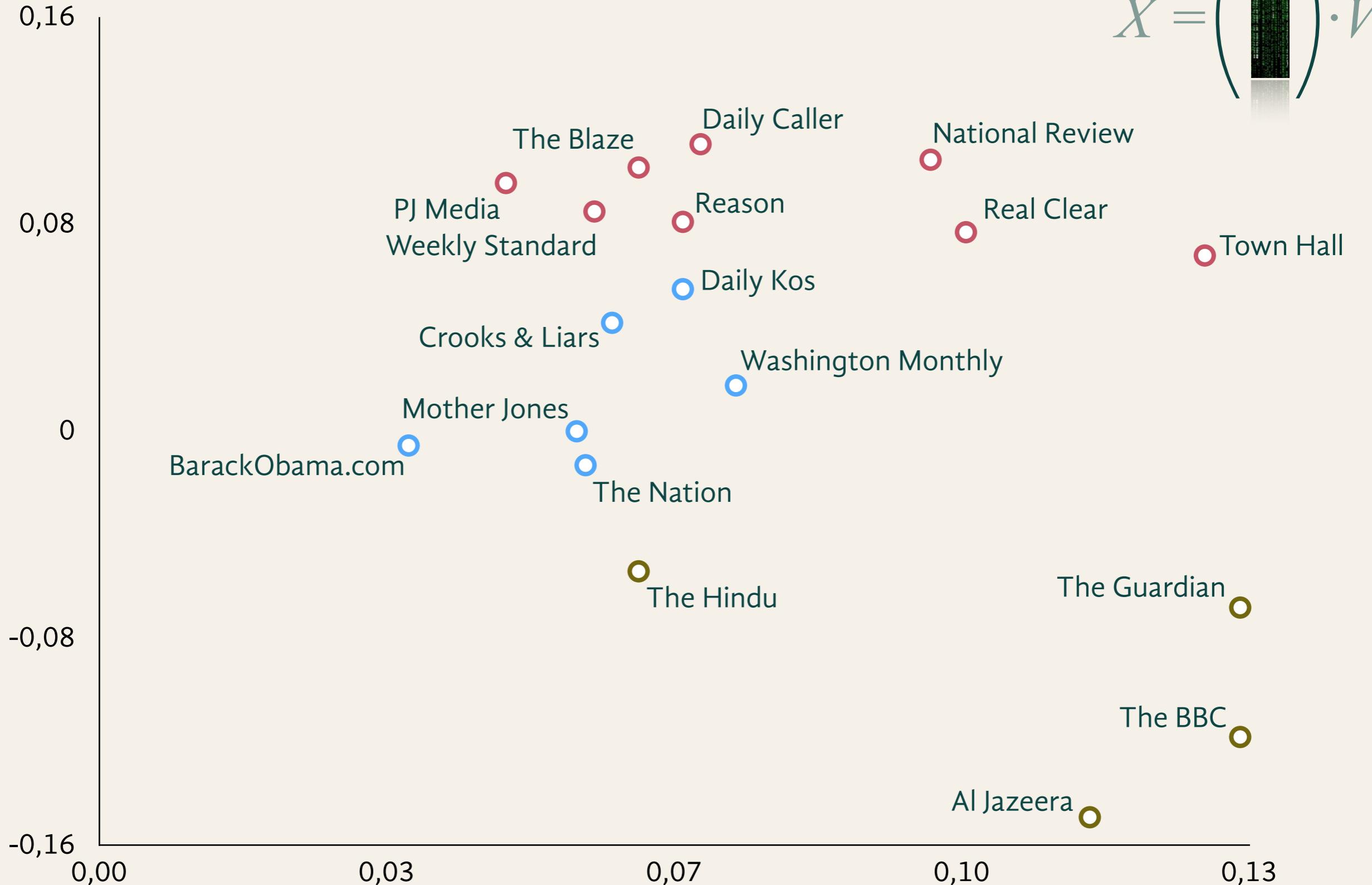
Do they correspond to *intuition*?

The first two latent dimensions:



Do they correspond to *intuition*?

The first two latent dimensions:



outlet space

$$\hat{X} = \begin{pmatrix} \text{matrix} \end{pmatrix} \cdot VT$$

Does it correspond to *intuition*?

At the top of dimension 2:



THE PATRIOT POST®
VOICE OF ESSENTIAL LIBERTY



Free Republic
The Gateway Pundit
Lonely Conservative
Right Wing News ...

Does it correspond to *intuition*?

At the top of dimension 2:



THE PATRIOT POST.
VOICE OF ESSENTIAL LIBERTY



Free Republic
The Gateway Pundit
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 ...

At the bottom of dimension 2:



National Post
The Age, SMH
Canada.com

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

How do they relate to political language?

Kind of.

It's not only
liberal/conservative.

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*? Kind of.

How do the patterns manifest in *language*?

How do the patterns manifest in *language*?

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

$$\left(\begin{array}{c} \text{matrix code} \\ \text{matrix code} \\ \text{matrix code} \end{array} \right) \approx \left(\begin{array}{c} \text{outlet space} \\ \text{matrix code} \end{array} \right) \left(\begin{array}{c} \text{quote space} \\ \text{matrix code} \end{array} \right)$$

How do the patterns manifest in *language*?

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

(matrix X) \approx *(matrix U)* \cdot *(matrix V^T)*

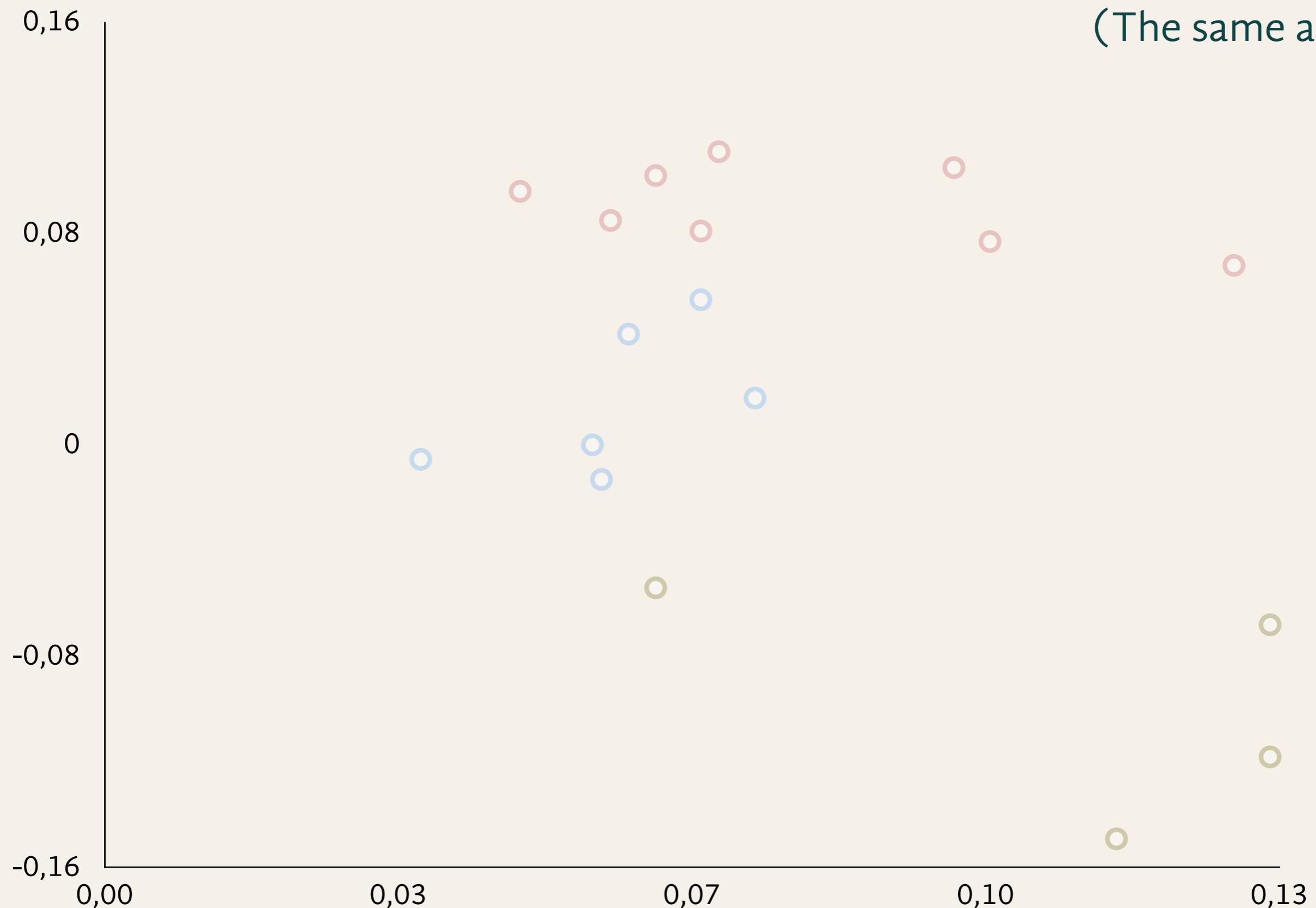
outlet space *quote space*

The diagram illustrates the matrix factorization $X \approx \hat{X} = U \cdot V^T$. Matrix X is represented by a large matrix with two columns. The first column is labeled "outlet space" and the second column is labeled "quote space". Matrix U is represented by a tall matrix with two columns. Matrix V^T is represented by a wide matrix with two rows. A red border highlights the V^T matrix.

How do the patterns manifest in *language*?

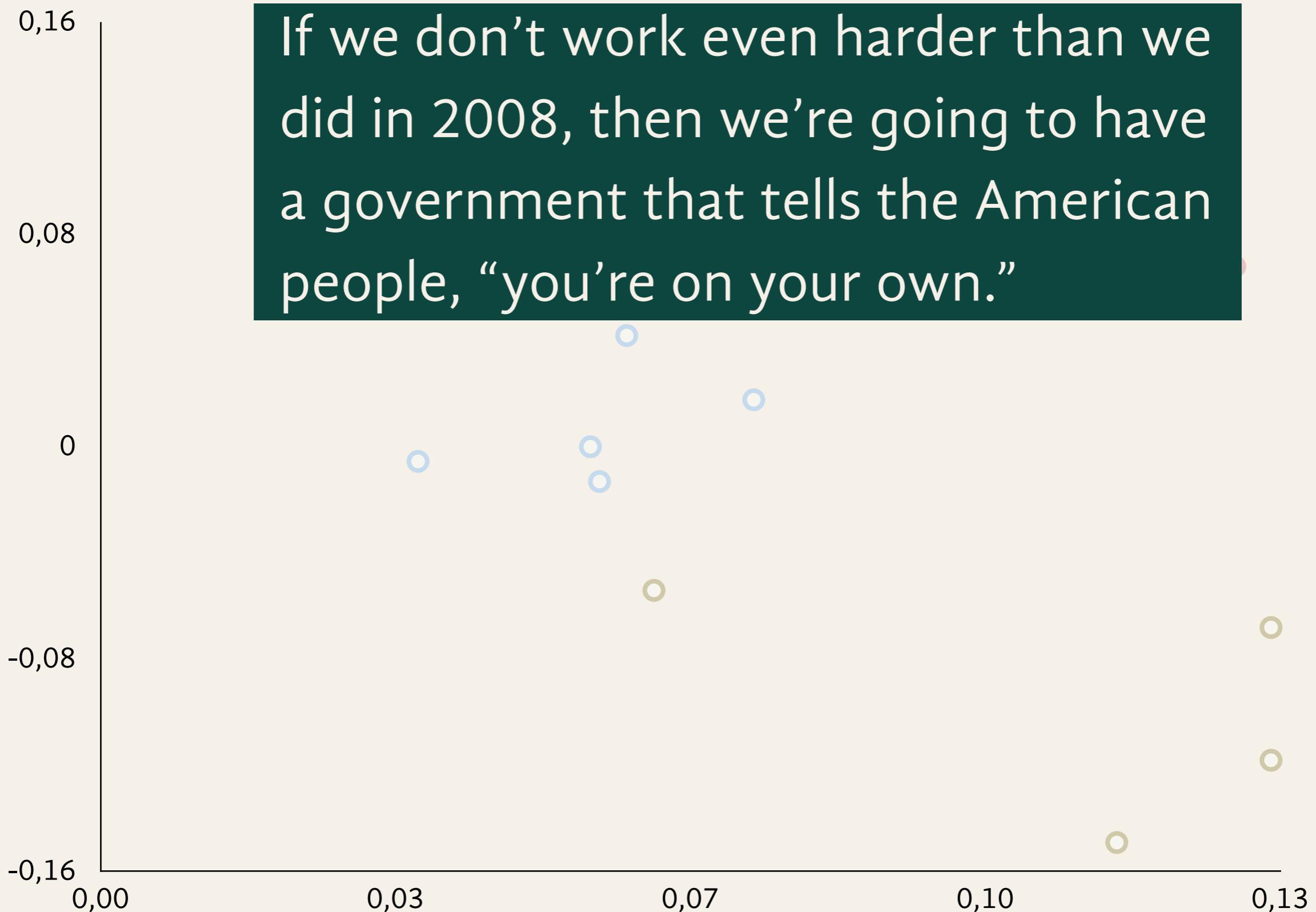
$$\hat{X} = U \cdot (\text{quote space})$$

(The same axes!)



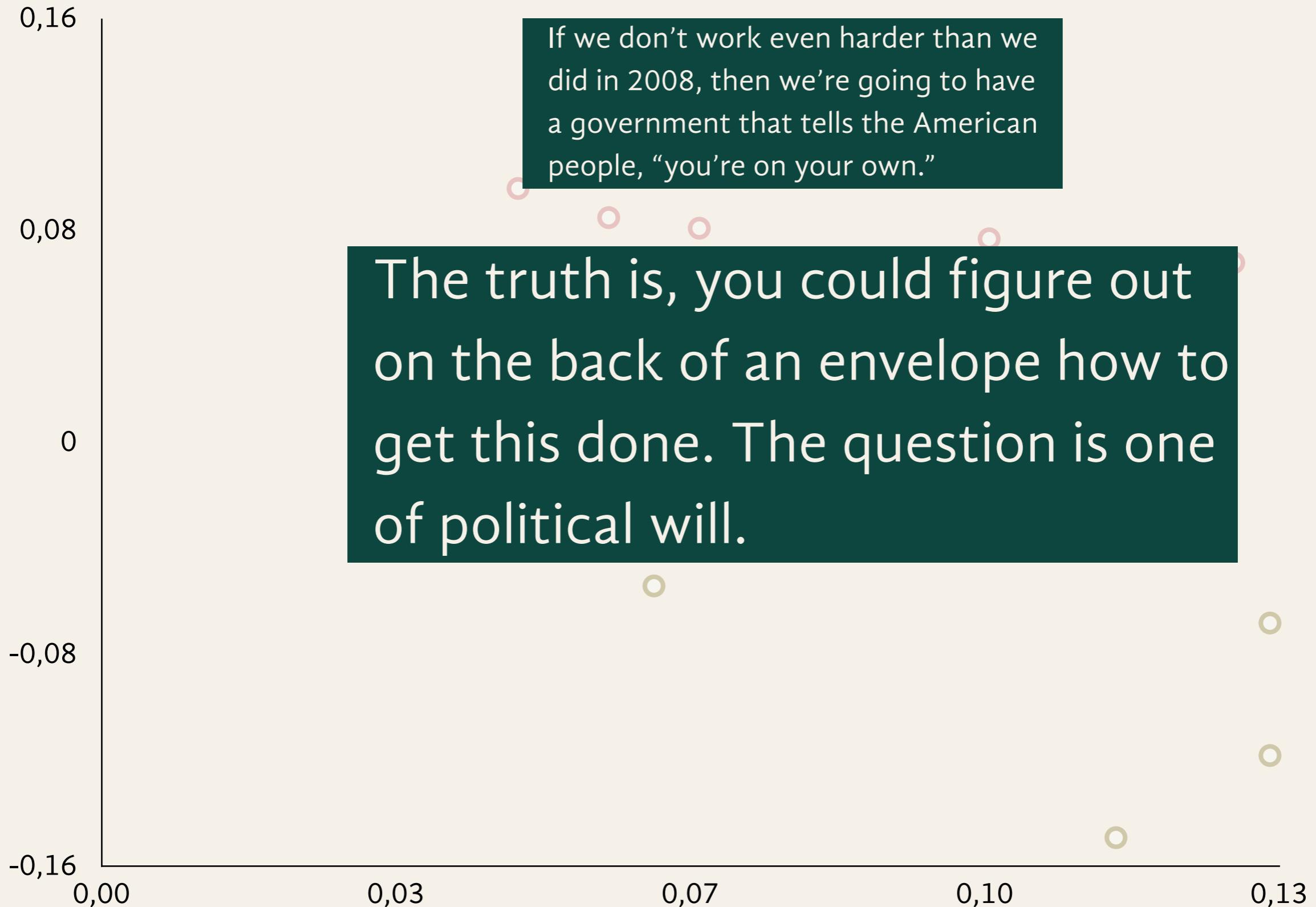
How do the patterns manifest in *language*?

$$\hat{X} = U \cdot (\begin{array}{c} \text{quote space} \\ \hline \text{[REDACTED]} \end{array})$$



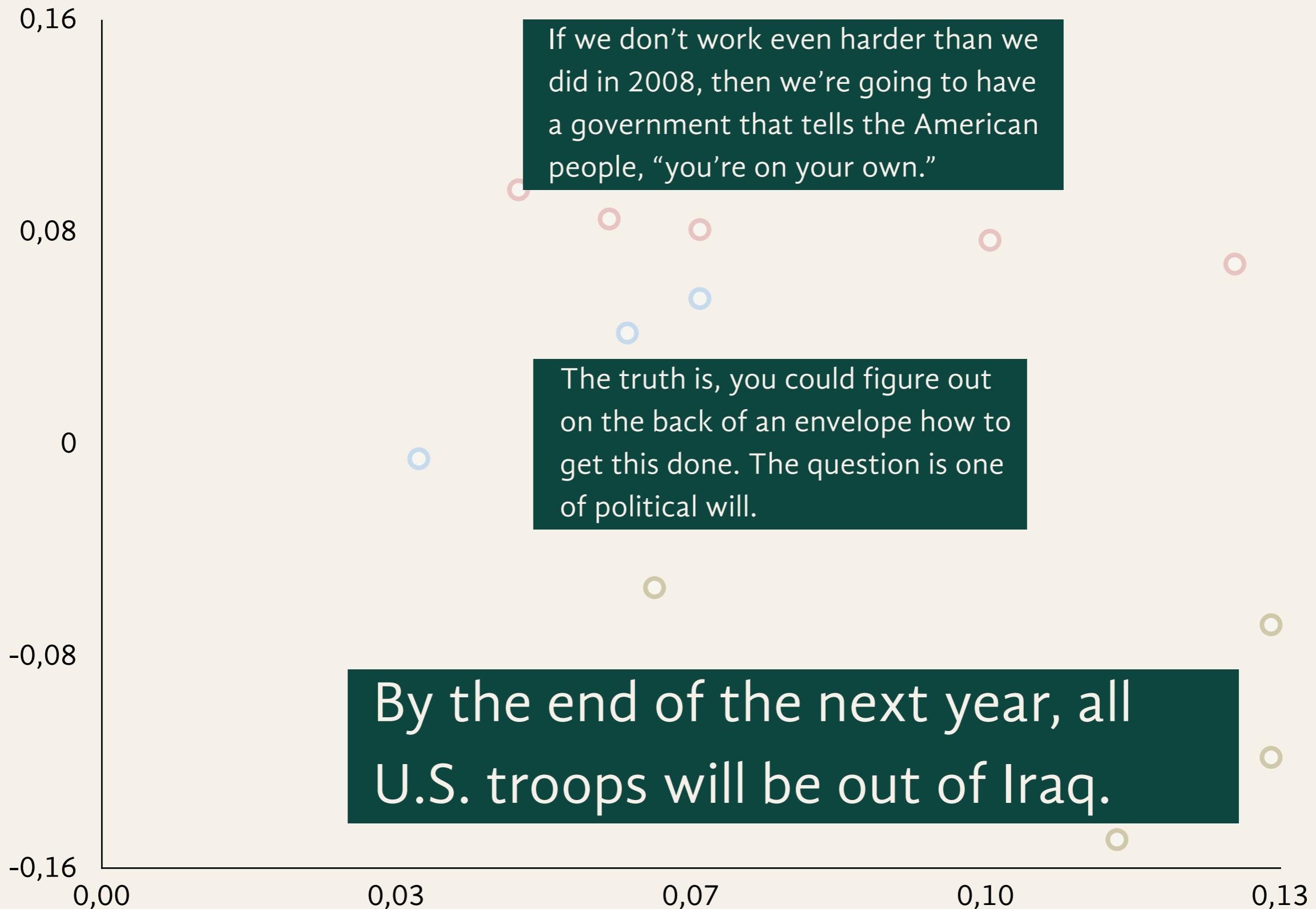
How do the patterns manifest in *language*?

$$\hat{X} = U \cdot (\begin{array}{c} \text{quote space} \\ \hline \text{[REDACTED]} \end{array})$$



How do the patterns manifest in *language*?

$$\hat{X} = U \cdot (\begin{array}{c} \text{quote space} \\ \hline \text{[REDACTED]} \end{array})$$



How do the patterns manifest in *language*?

Average sentiment of quotes:

0.02 0.006 0.008 0.01 0.012 0.014

0.01

0.00

-0.01

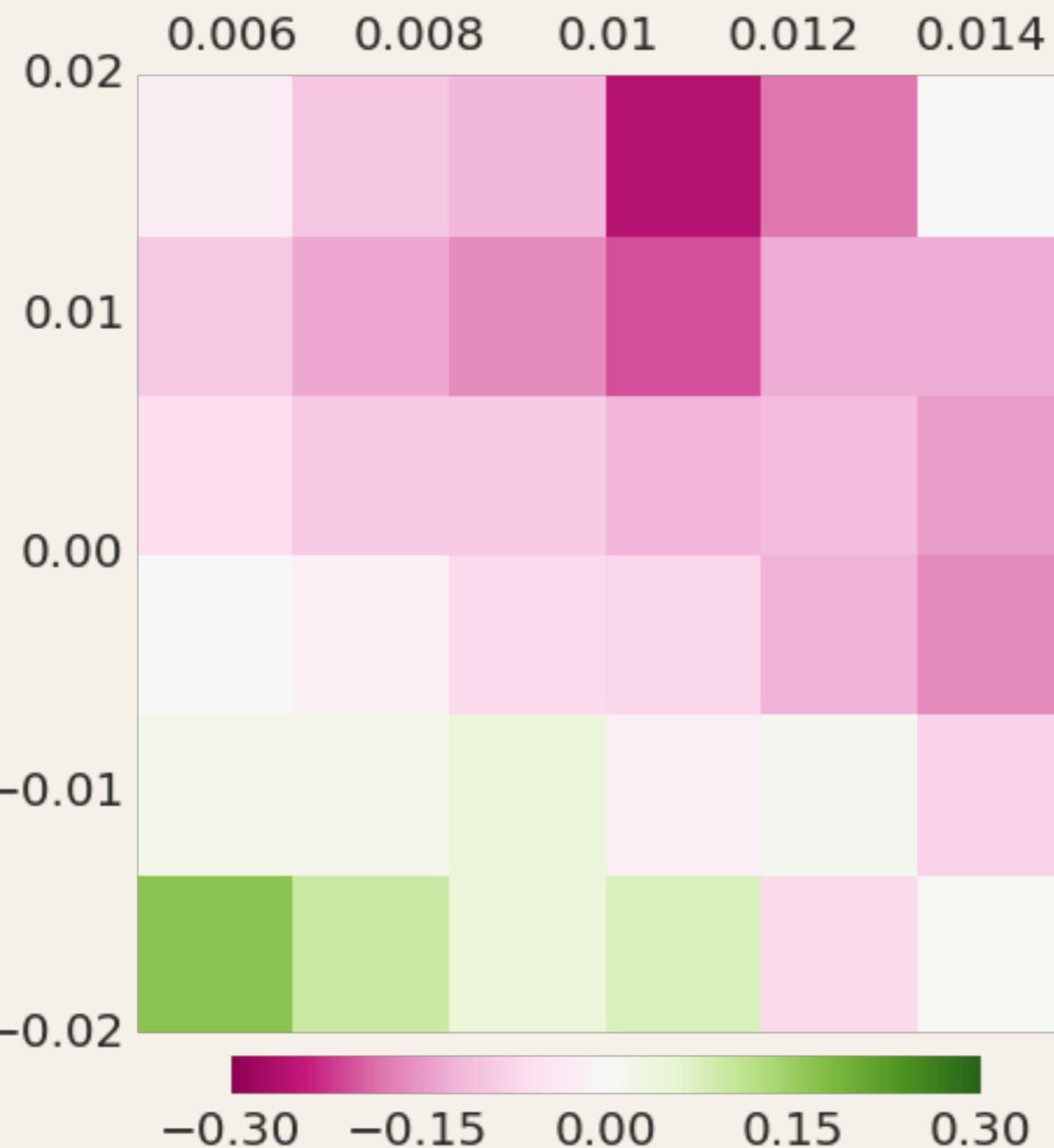
-0.02

-0.30 -0.15 0.00 0.15 0.30

(The same axes!)

How do the patterns manifest in *language*?

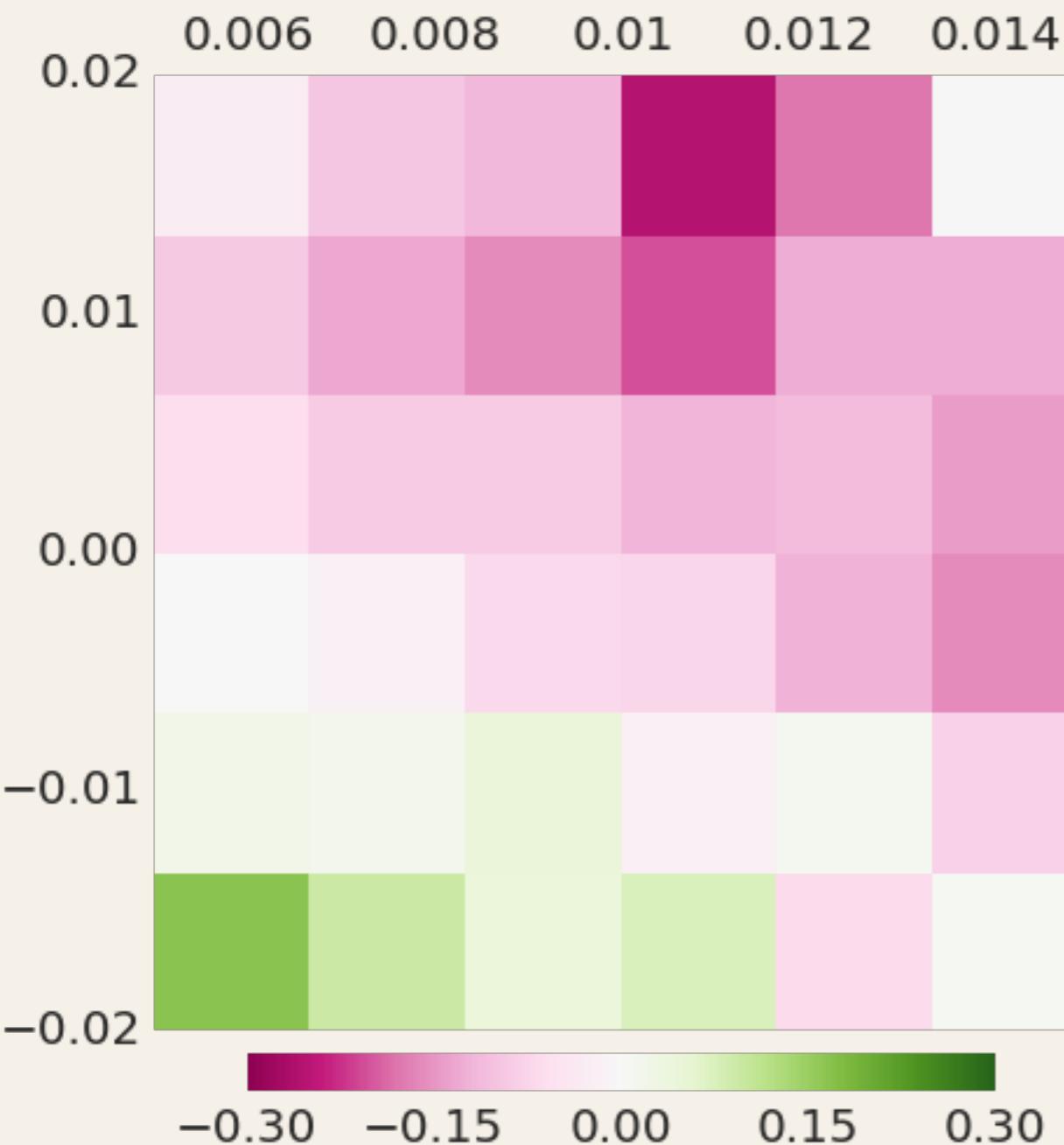
Average sentiment of quotes:



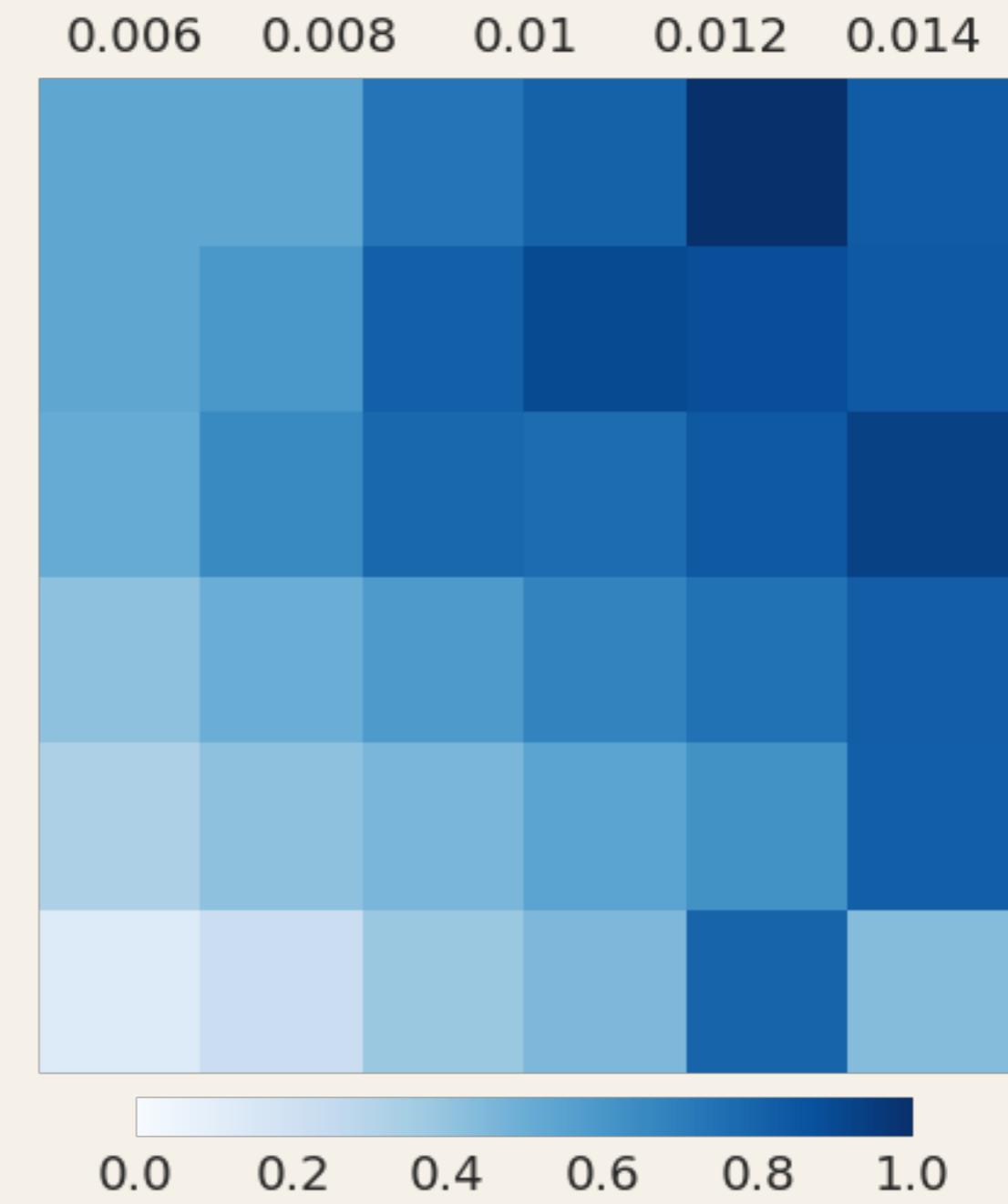
(The same axes!)

How do the patterns manifest in *language*?

Average sentiment of quotes:



Amount of negation in quotes:



(The same axes!)

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*? Kind of.

How do the patterns manifest in *language*?

Are they
Do they

Linguistic properties
map consistently
to the bias space.

? Yes!
Kind of.

How do the patterns manifest in *language*?

Conclusion

Quoting patterns are systematic:
We can predict them.

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

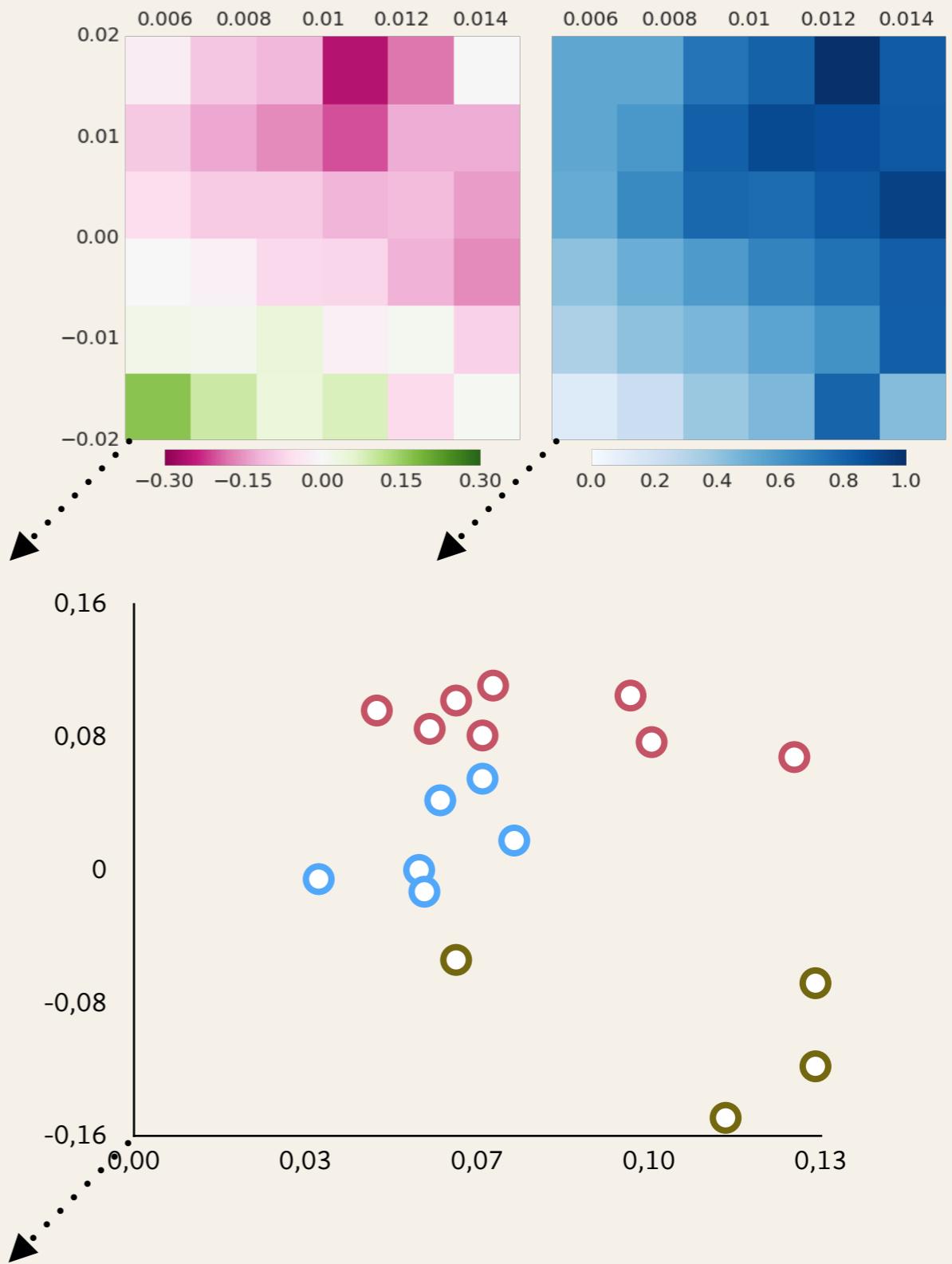
Conclusion

Quoting patterns are systematic:
We can predict them.

Selection bias is multidimensional.

Language plays a part:
Sentiment and negation show
consistent alignment.

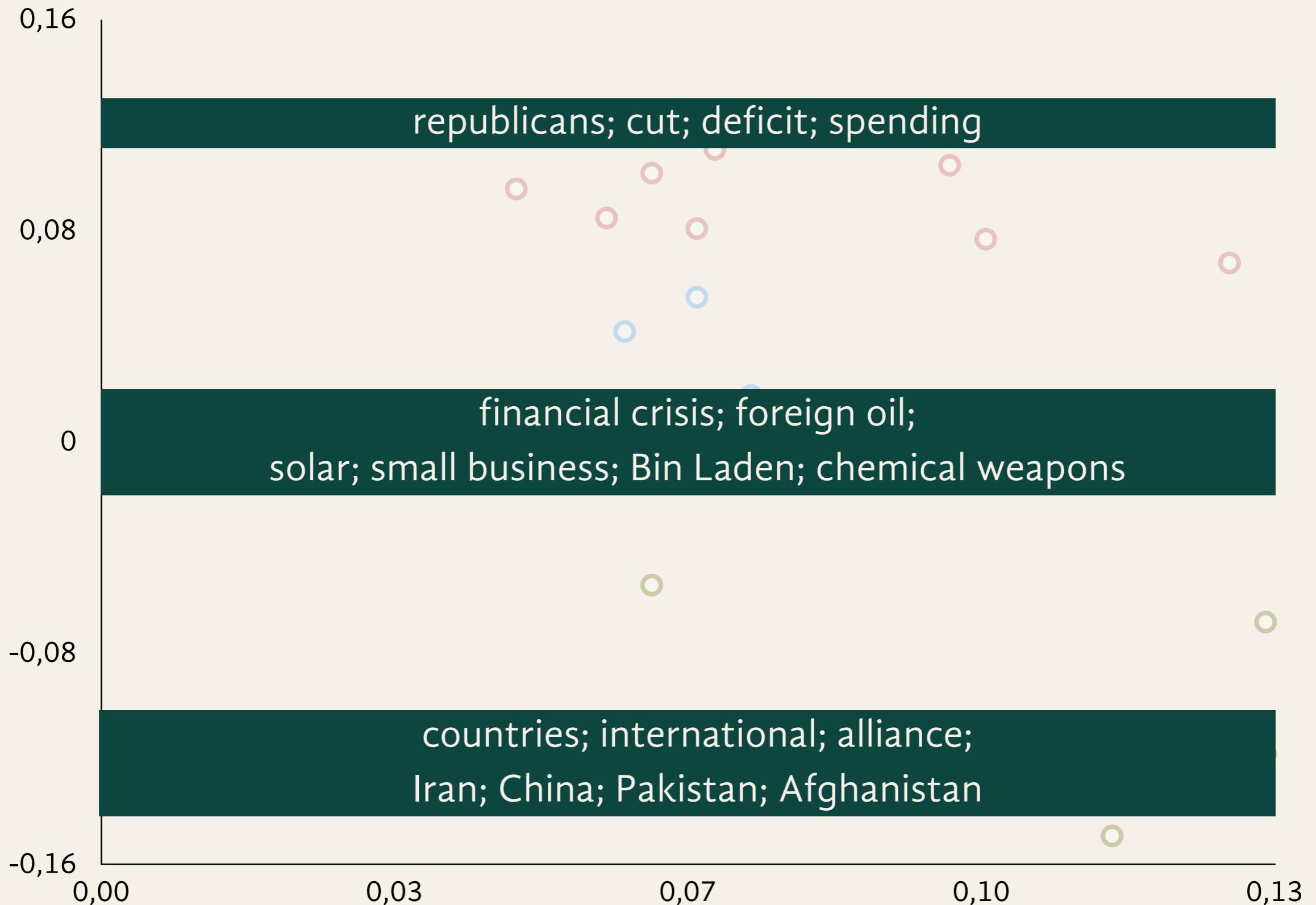
Online visualization & data:
<http://snap.stanford.edu/quotus/>



extra slides

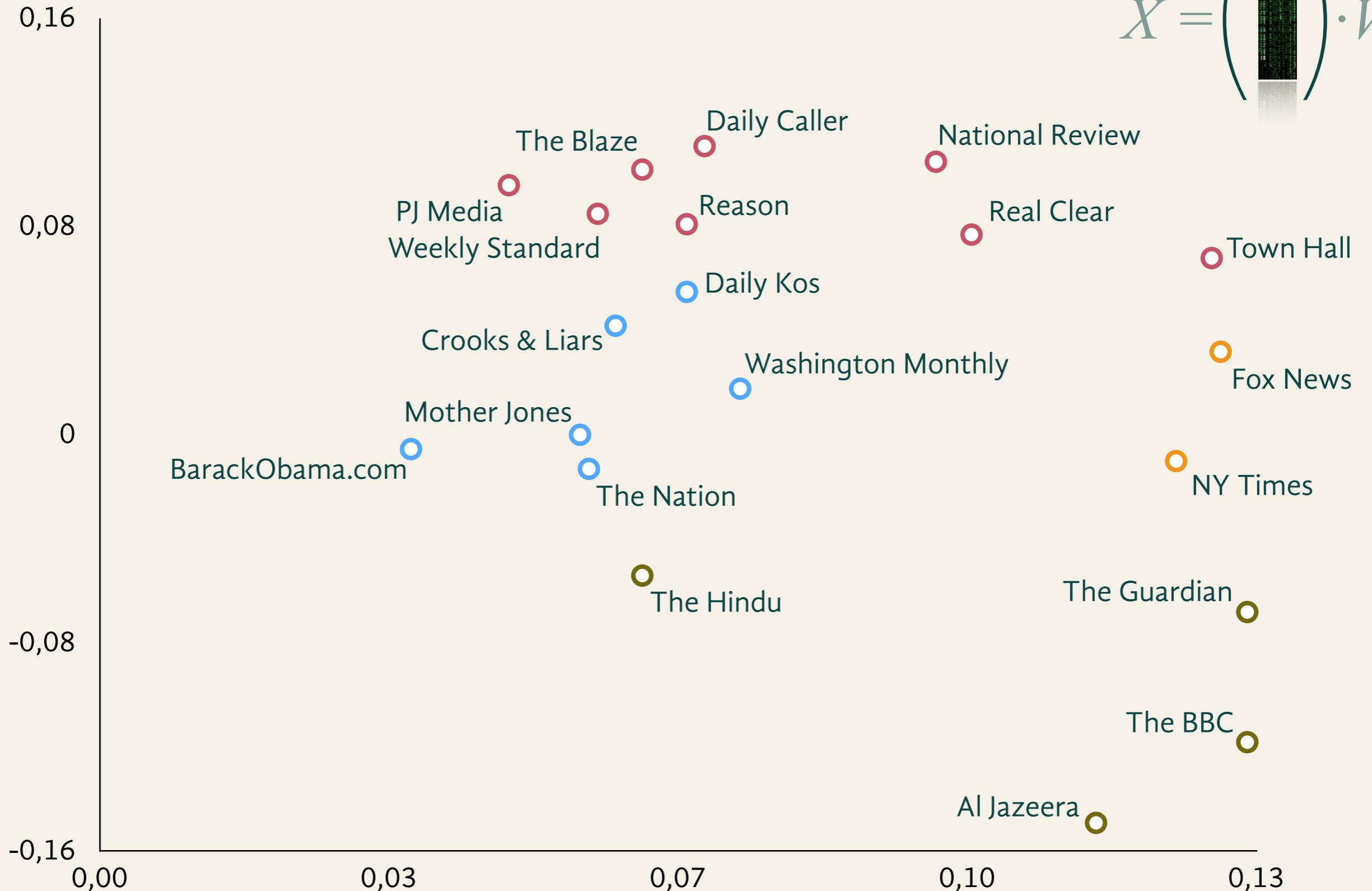
Can *language* characterize the bias?

$$\hat{X} = U \cdot \begin{pmatrix} \text{quote space} \\ \text{space} \end{pmatrix}$$



Do they correspond to *intuition*?

The first two latent dimensions:



Can *language* characterize the bias?

$$\hat{X} = U \cdot \left(\begin{array}{c} \text{quote space} \\ \hline \dots \end{array} \right)$$

