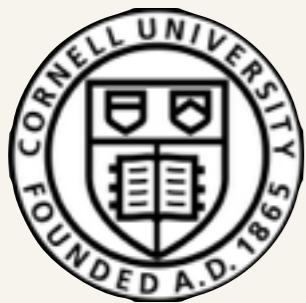


The structure of **political media coverage**
as revealed by **quoting patterns**:



QUOTing the POTUS



equal contribution {

* Vlad Niculae (me)	<i>Cornell (MPI-SWS)</i>
*Caroline Suen	<i>Stanford</i>
*Justine Zhang	<i>Cornell (Stanford, MPI-SWS)</i>
Cristian Danescu-Niculescu-Mizil	<i>Cornell (MPI-SWS)</i>
Jure Leskovec	<i>Stanford</i>

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.

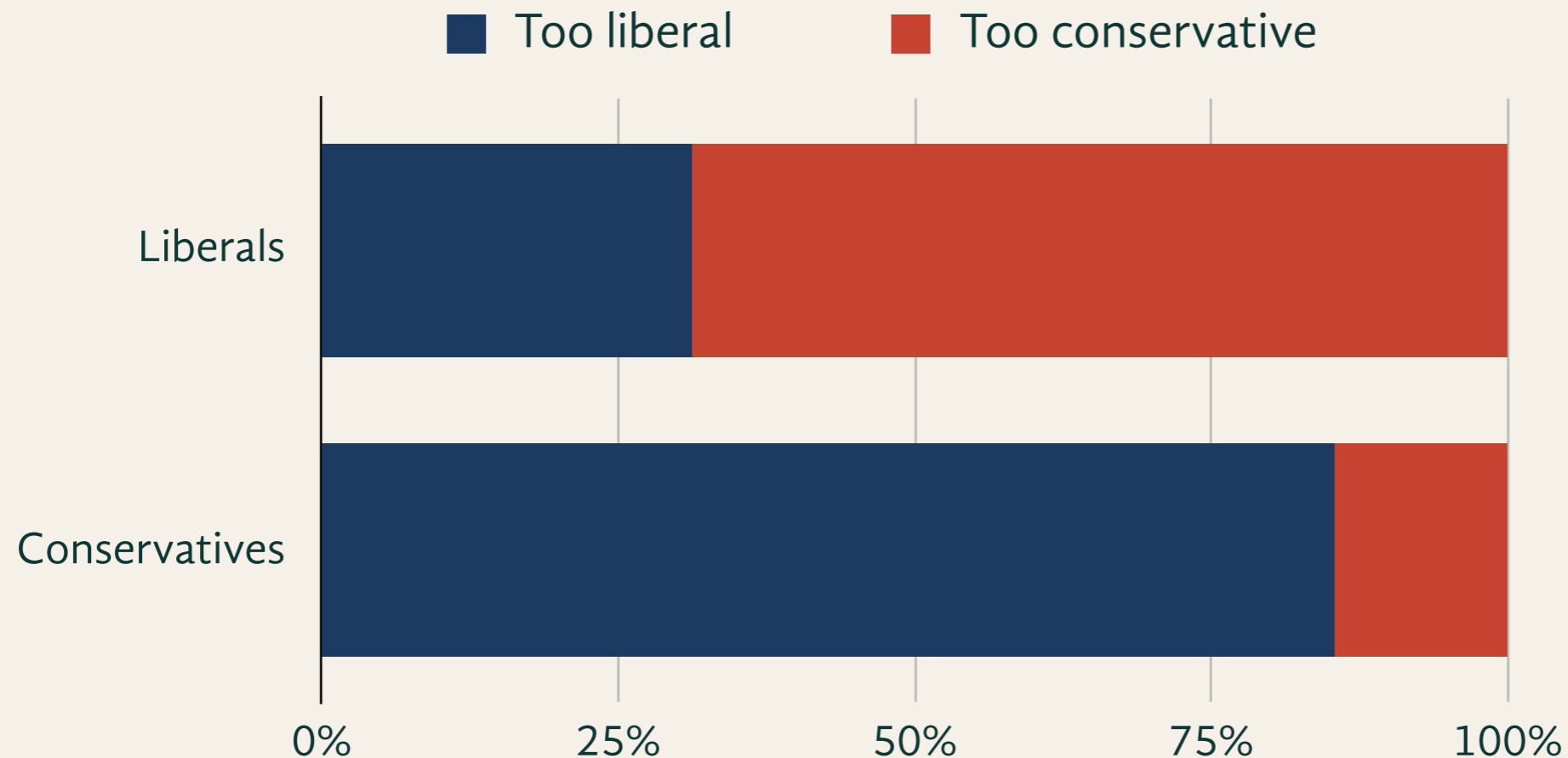
Capital Mall, DC

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



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*?

How do the patterns manifest in *language*?

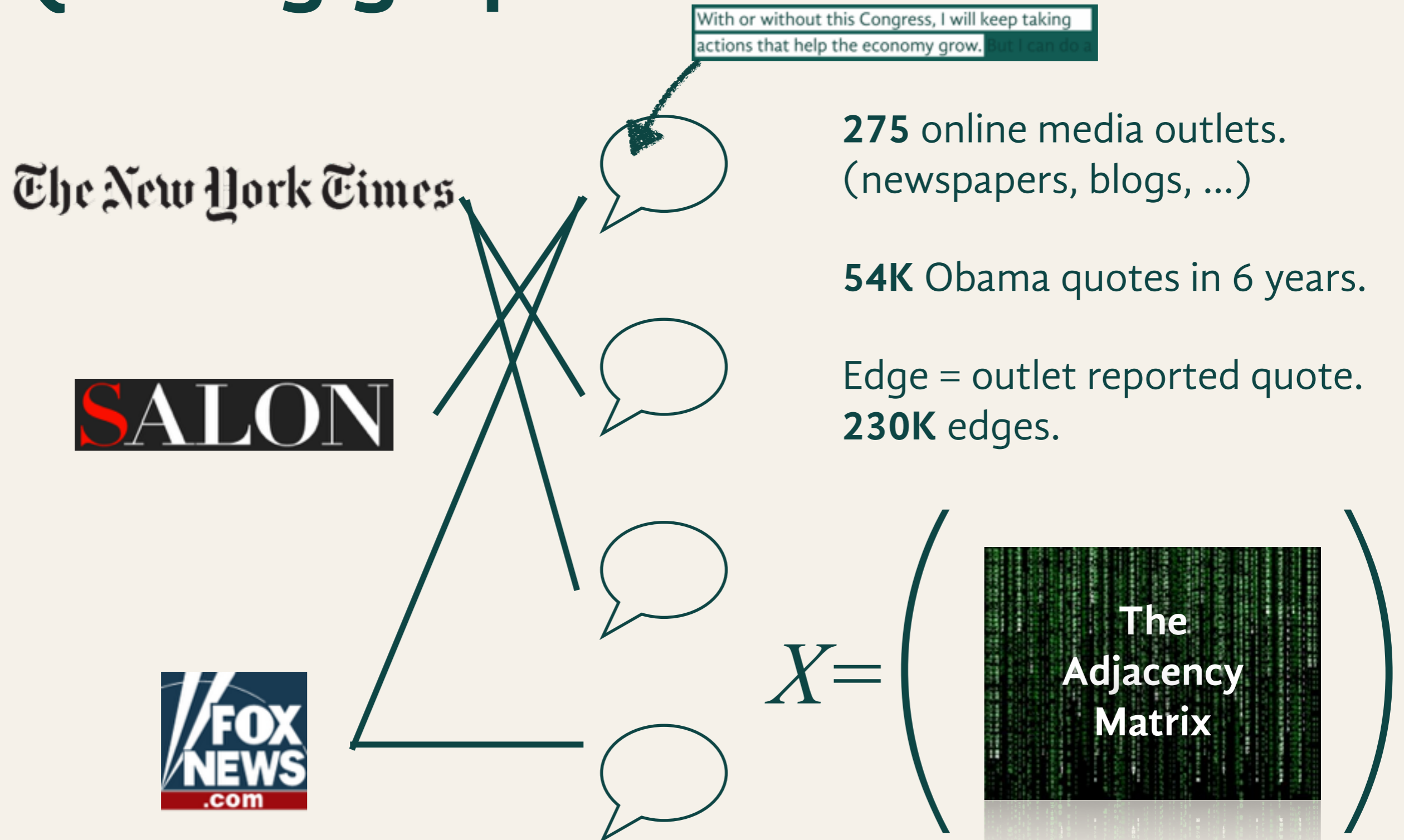
Are quoting patterns *systematic*?

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Are quoting patterns *systematic*?

Quoting graph



Are quoting patterns *systematic*?

Can we predict quoting?

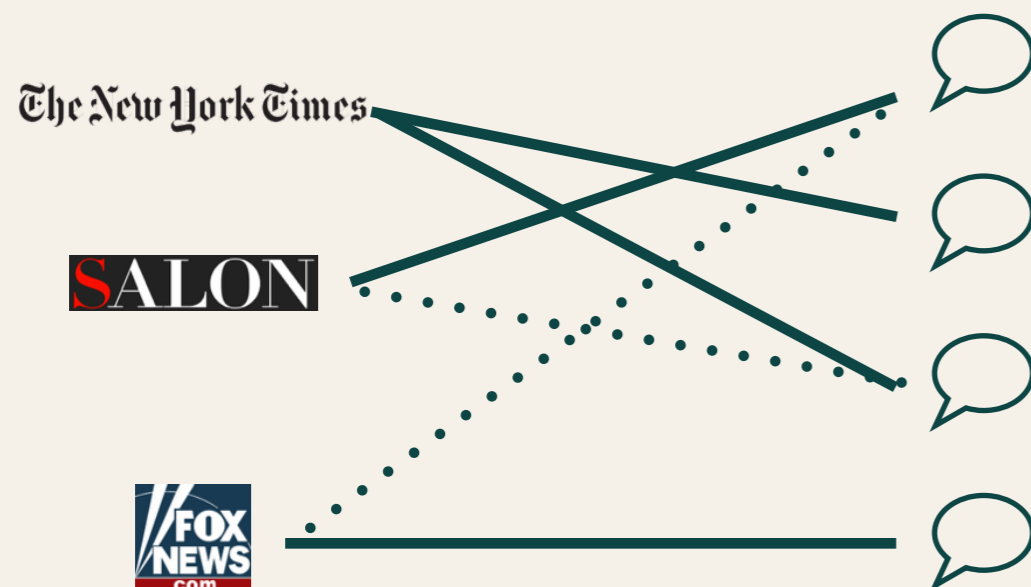
If we can predict quoting choices, it means there is a pattern.



Are quoting patterns *systematic*?

Can we predict quoting?

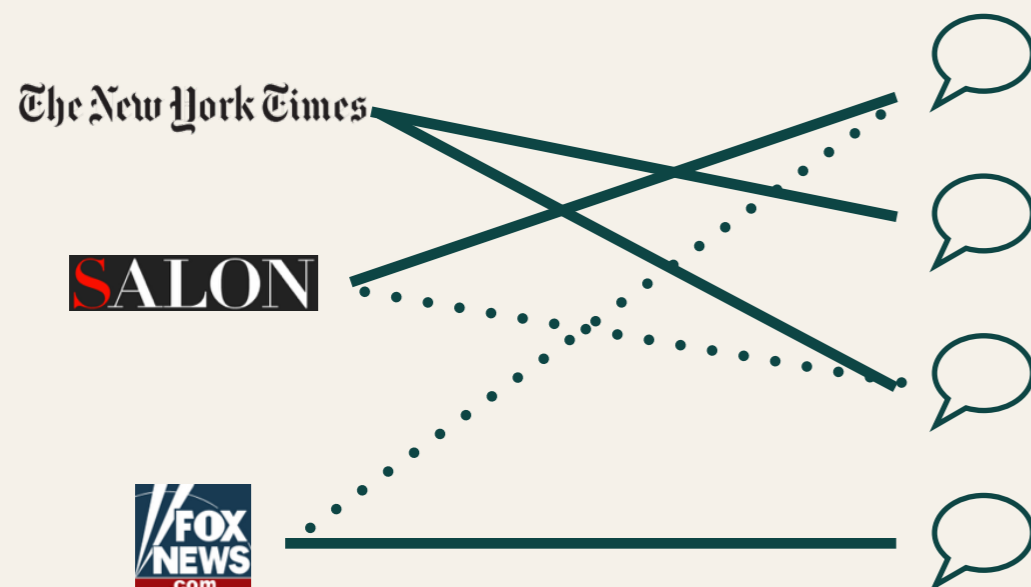
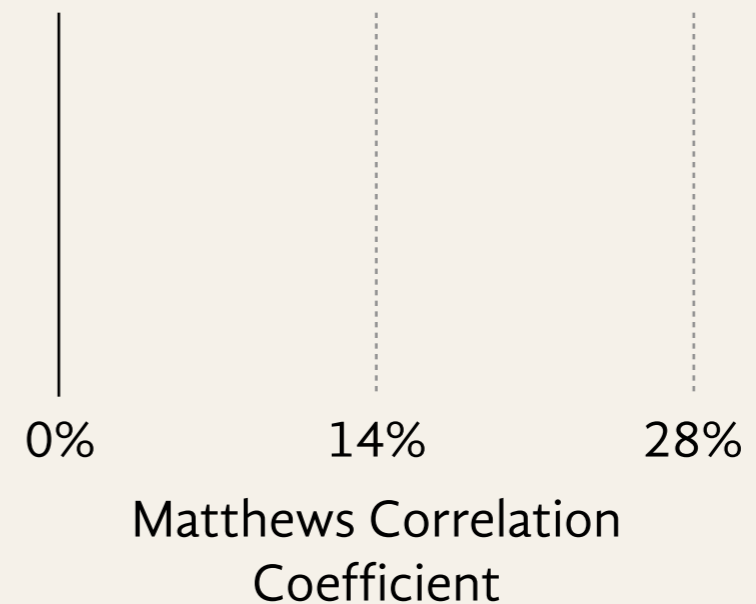
Leave out **250K** entries of X .



Are quoting patterns *systematic*?

Can we predict quoting?

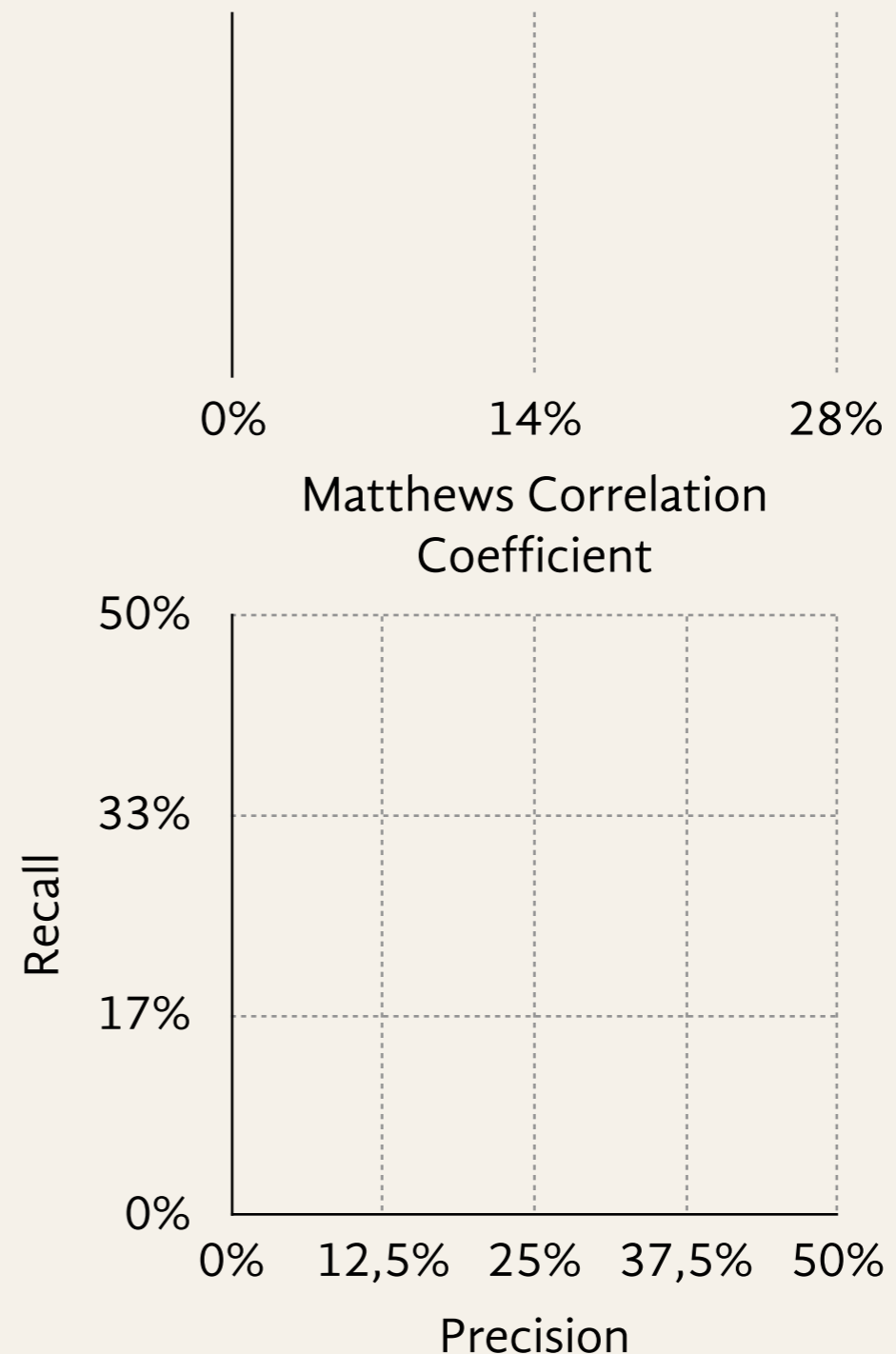
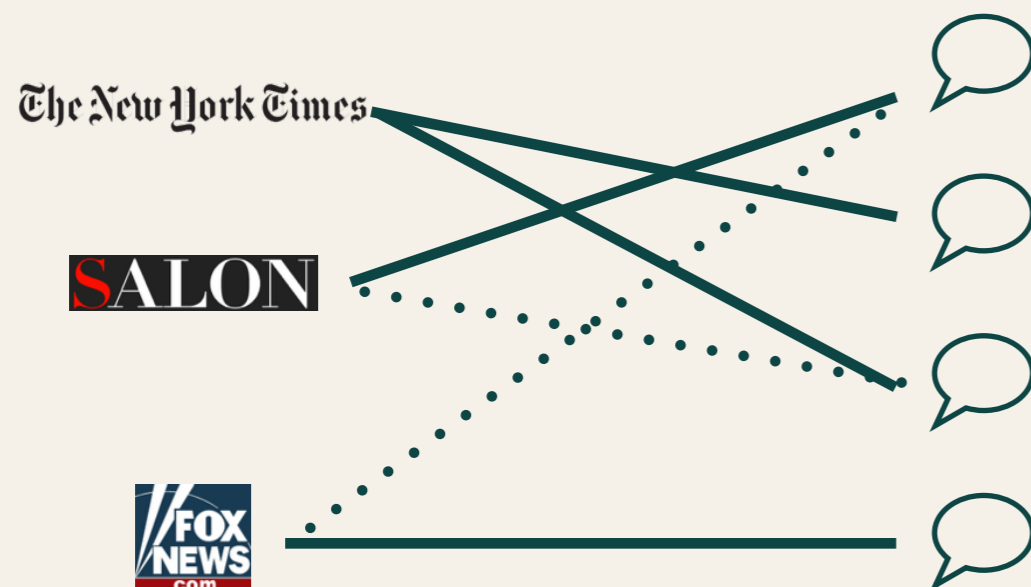
Leave out **250K** entries of X .



Are quoting patterns *systematic*?

Can we predict quoting?

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



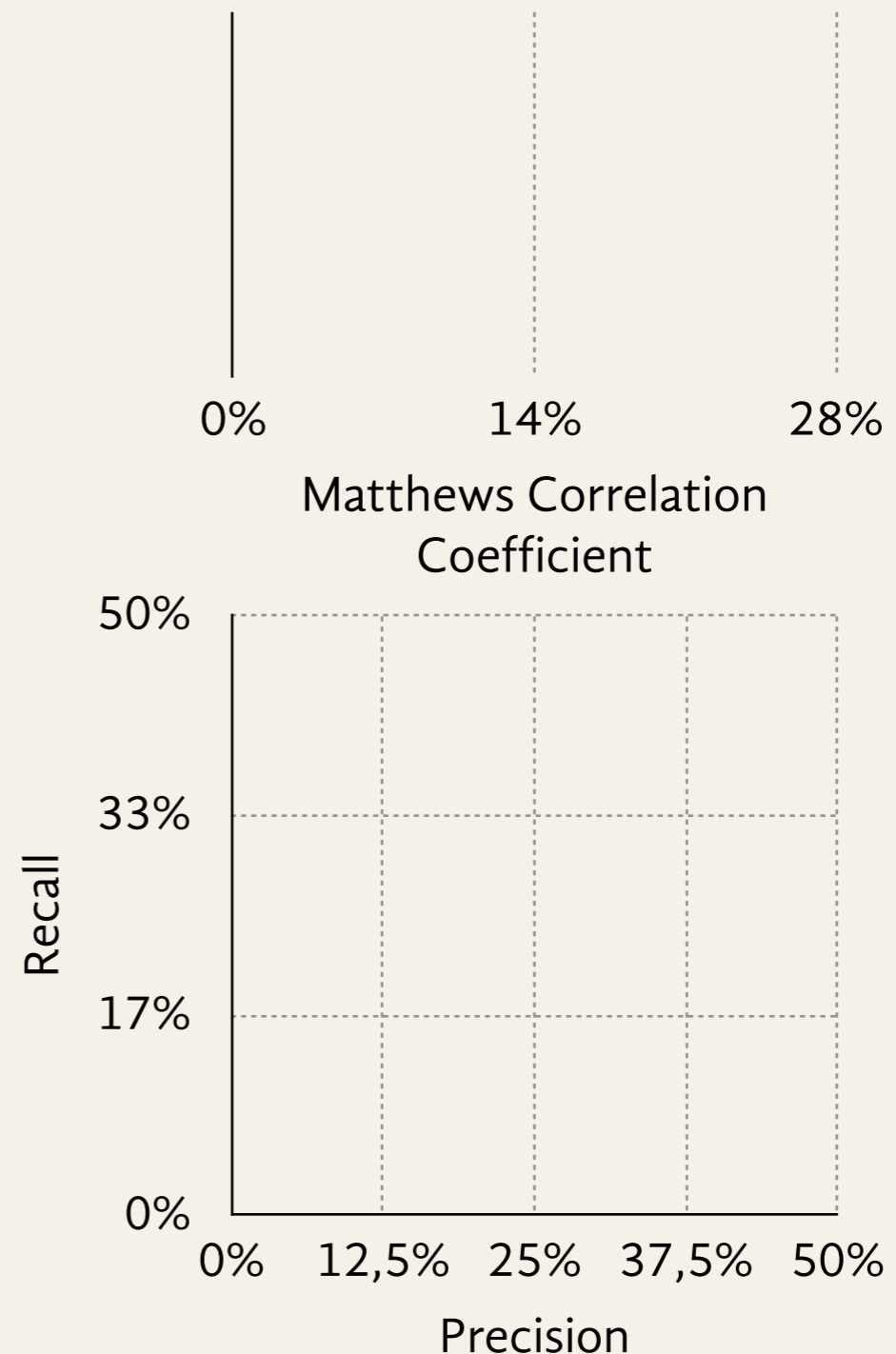
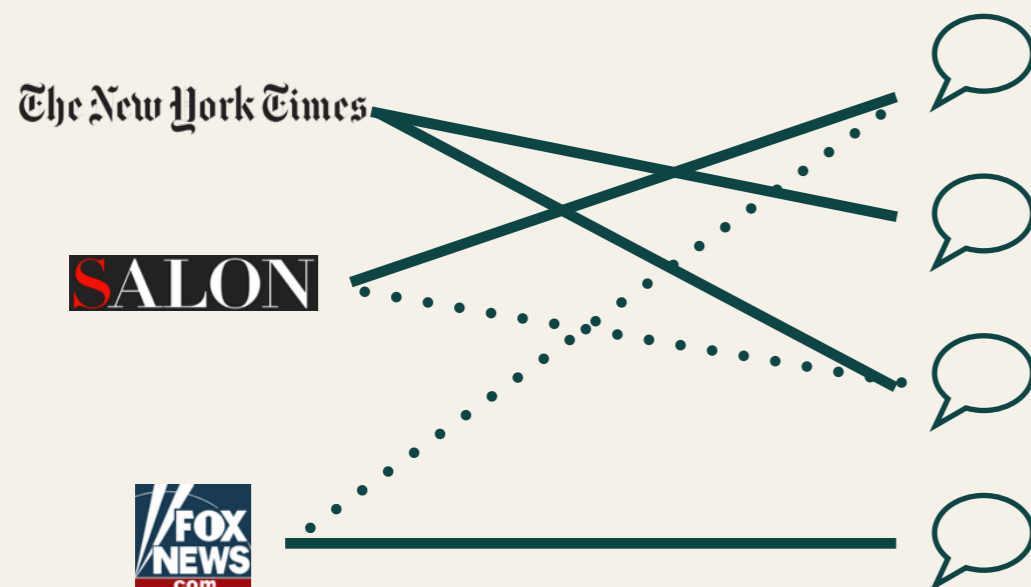
Are quoting patterns *systematic*?

Can we predict quoting?

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

In a “**fair world**” without bias:

- popular quotes get cited more,



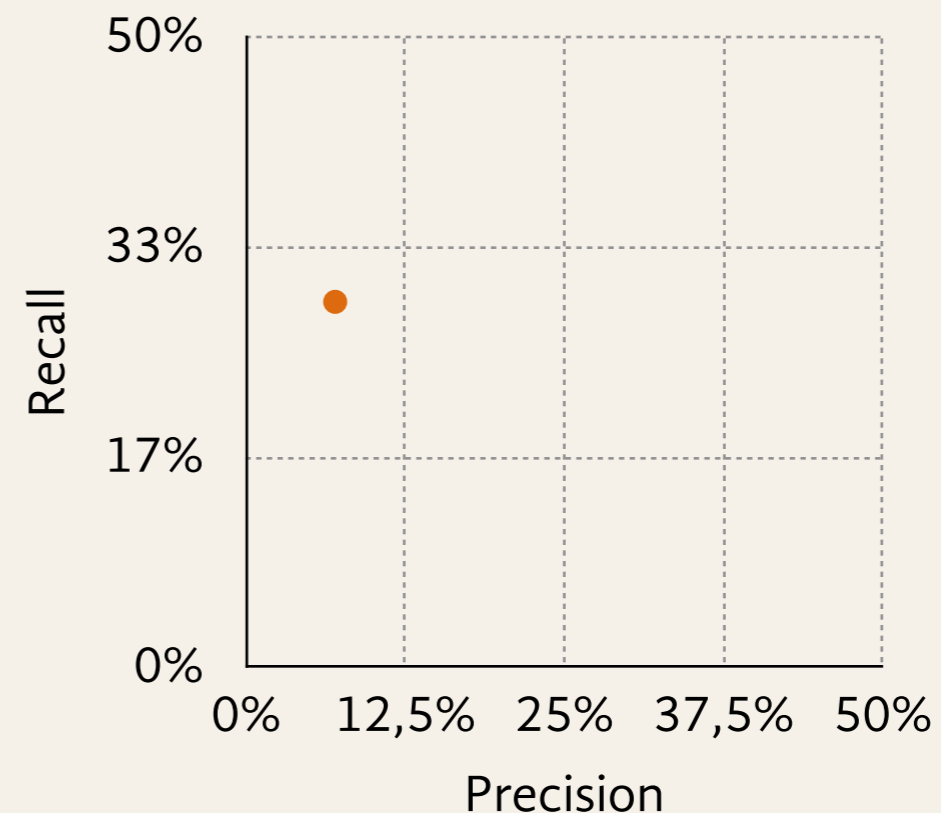
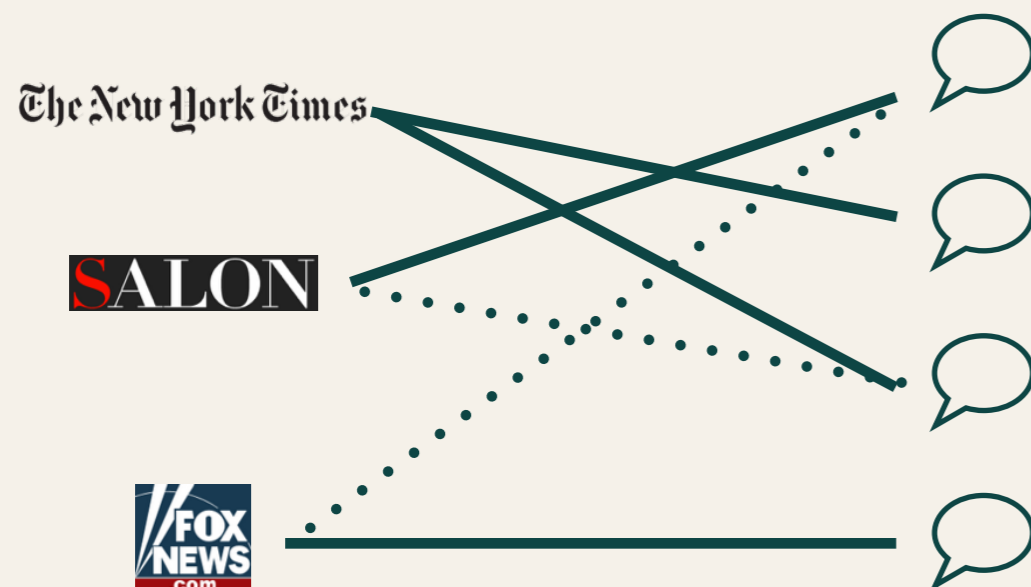
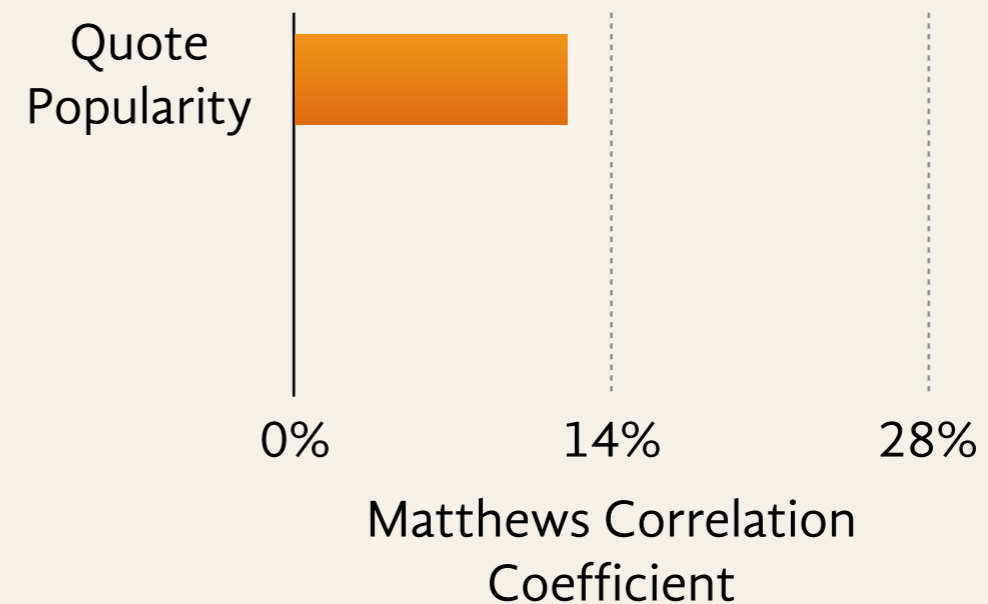
Are quoting patterns *systematic*?

Can we predict quoting?

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

In a “**fair world**” without bias:

- popular quotes get cited more,



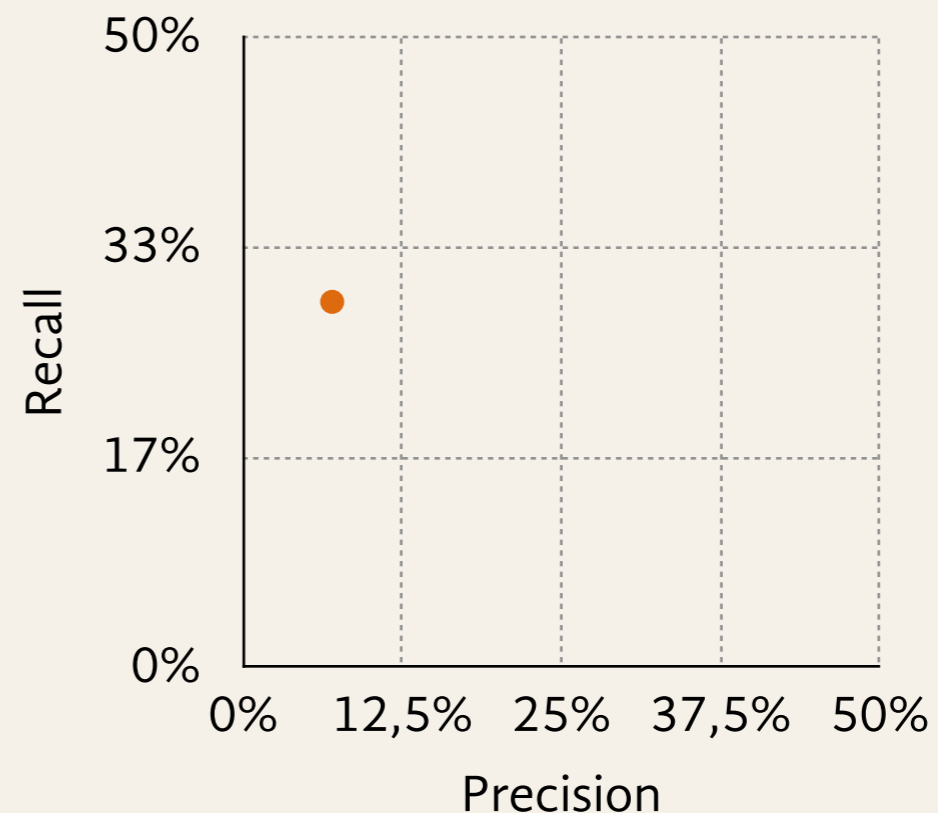
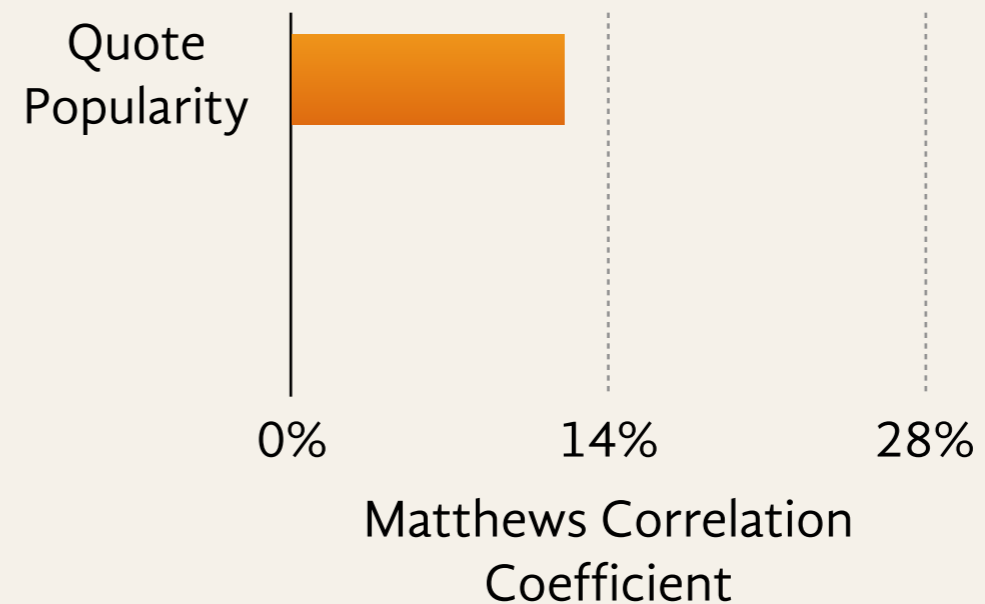
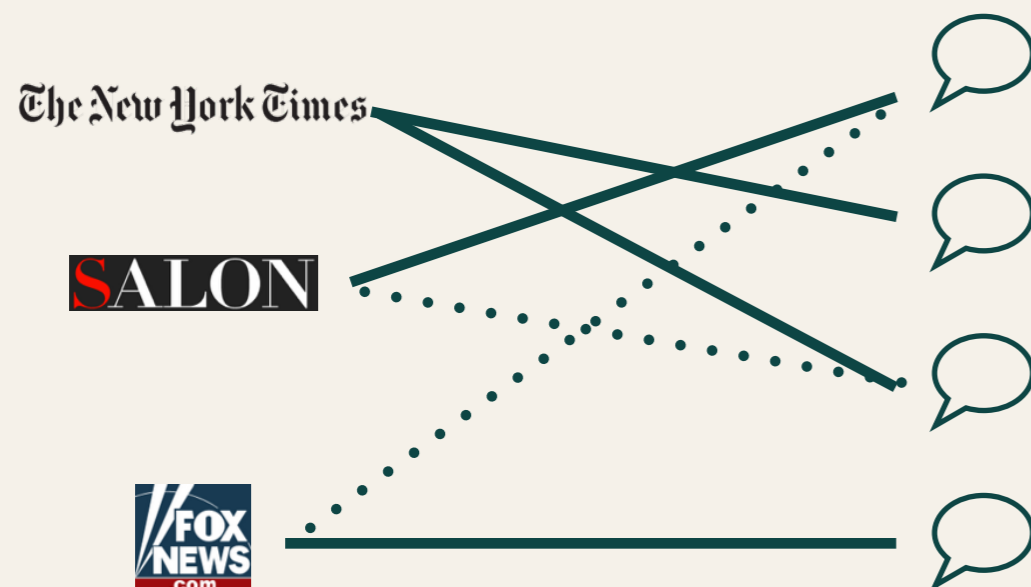
Are quoting patterns *systematic*?

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.



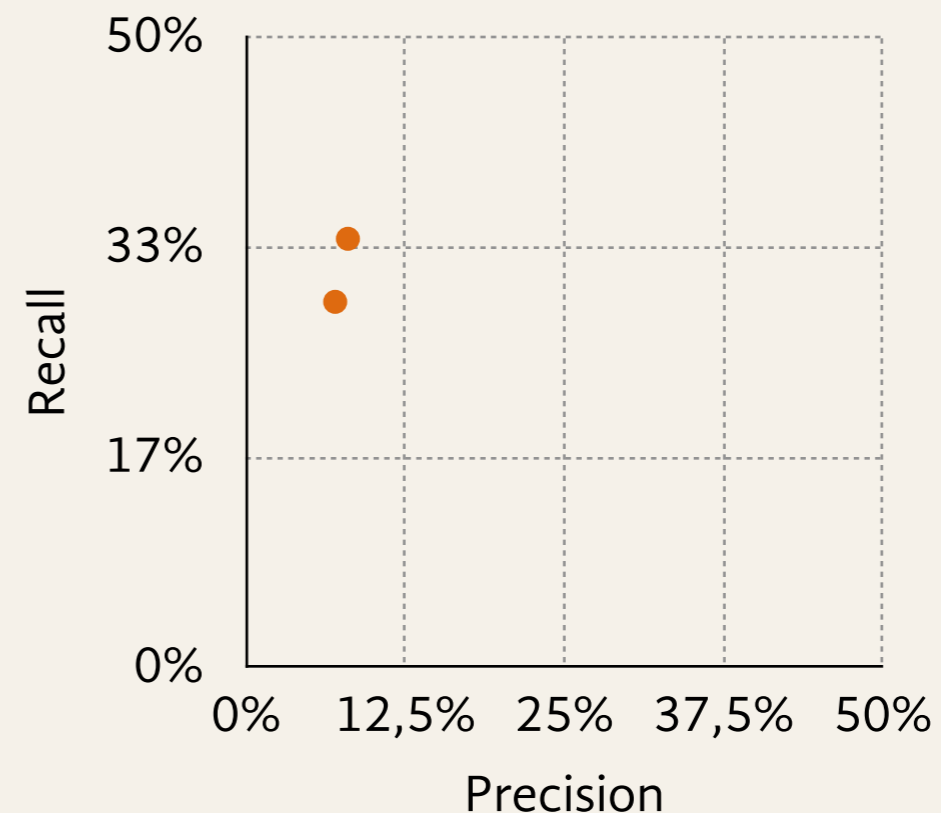
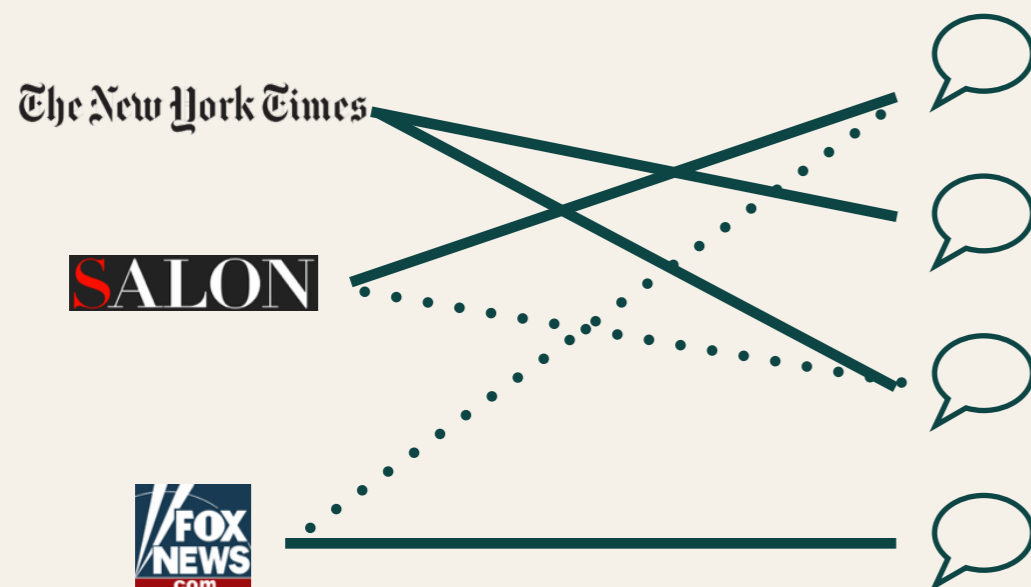
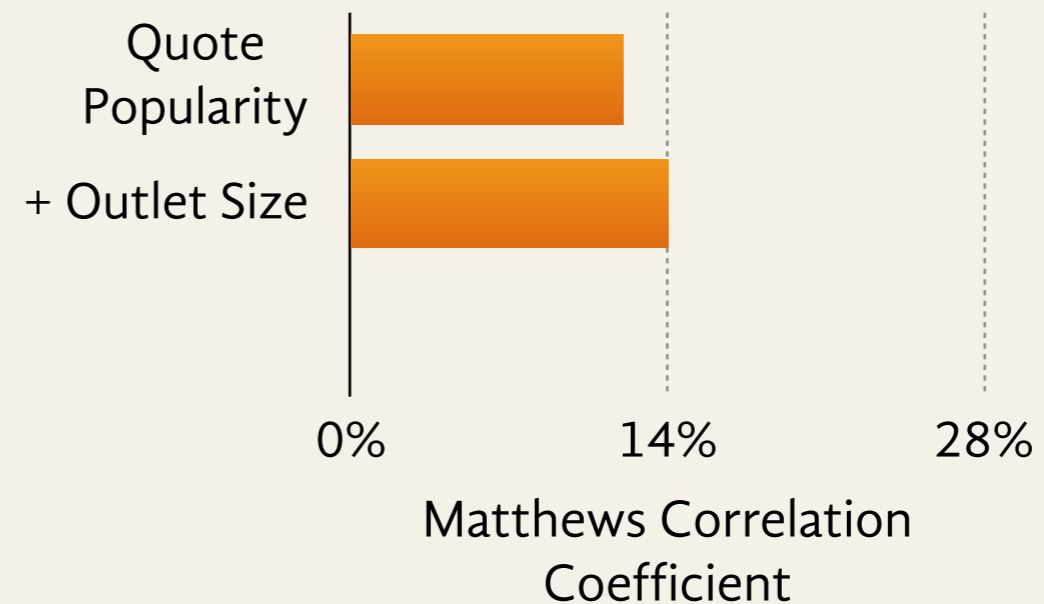
Are quoting patterns *systematic*?

Can we predict quoting?

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

In a “**fair world**” without bias:

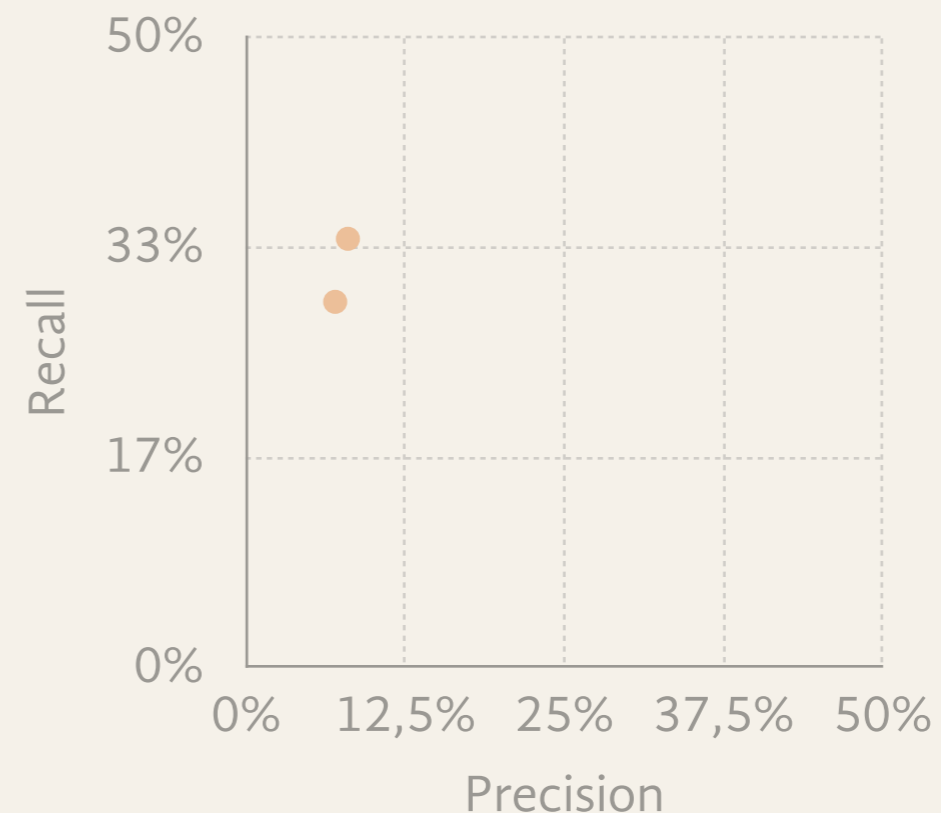
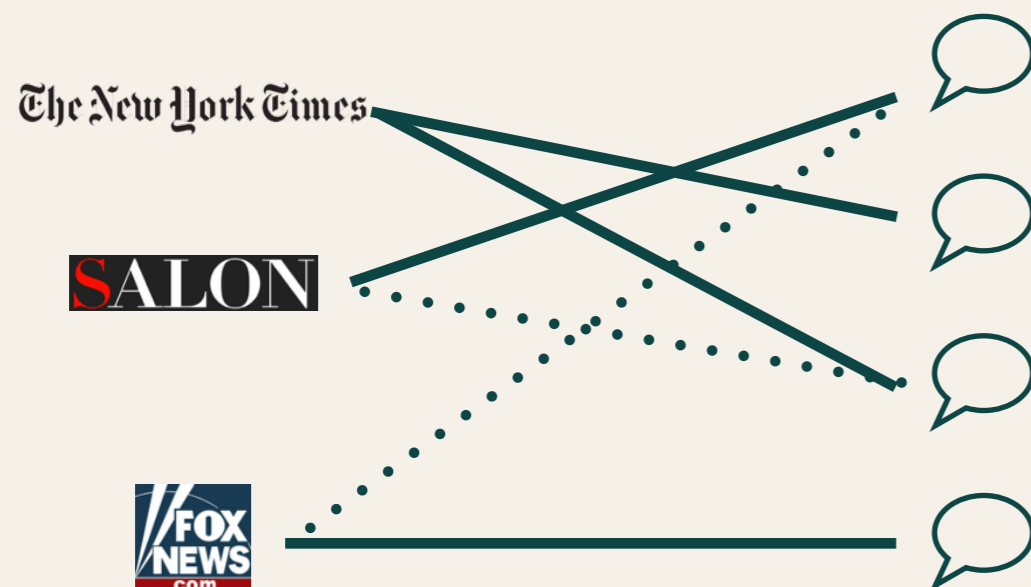
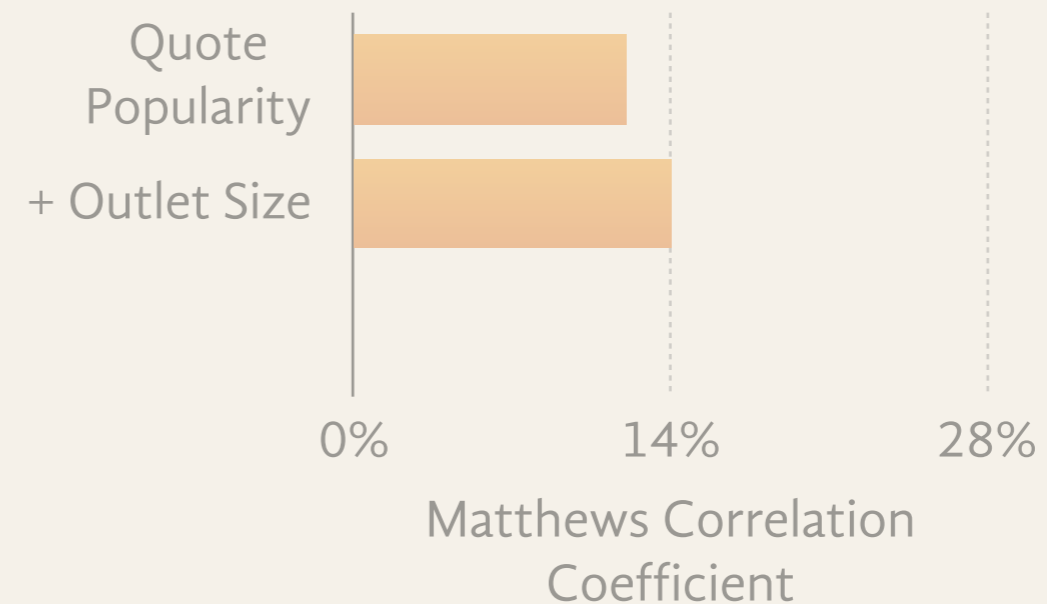
- popular quotes get cited more,
- and big outlets quote more.



Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

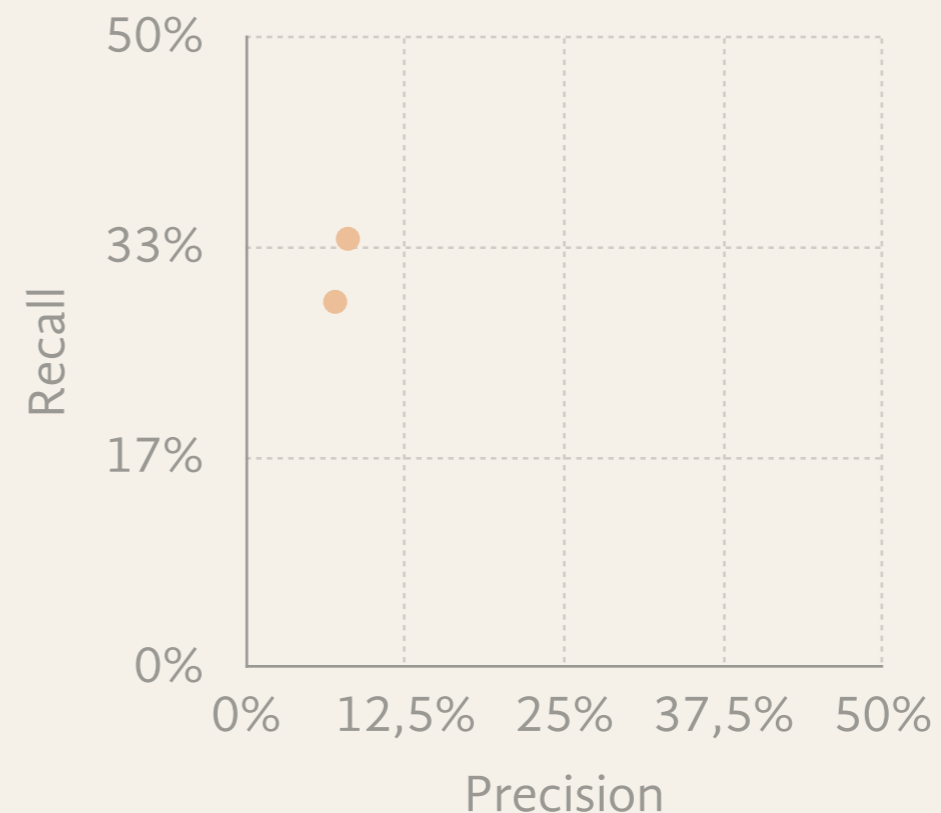
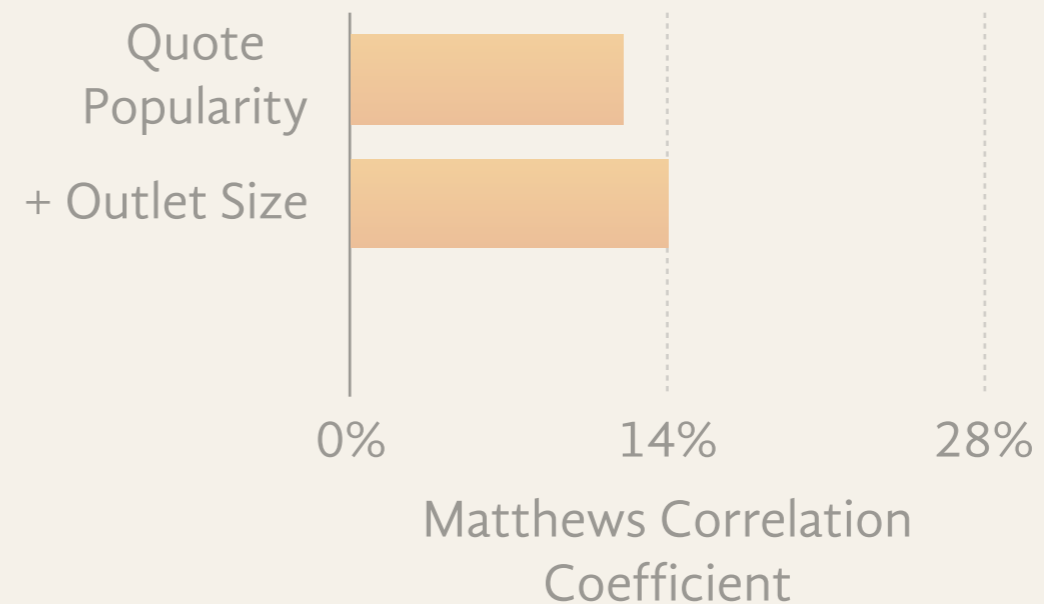
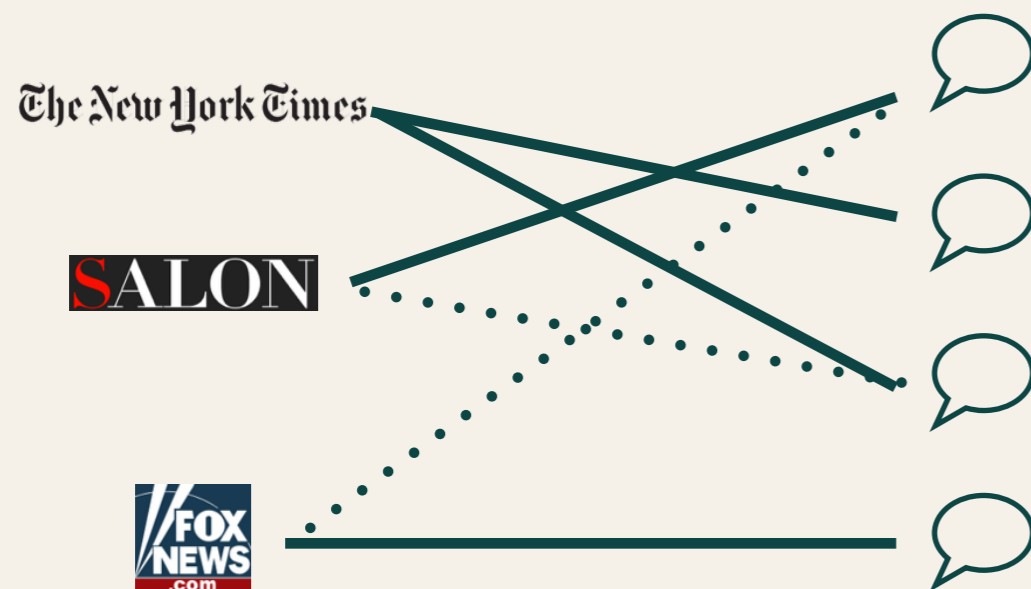


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

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

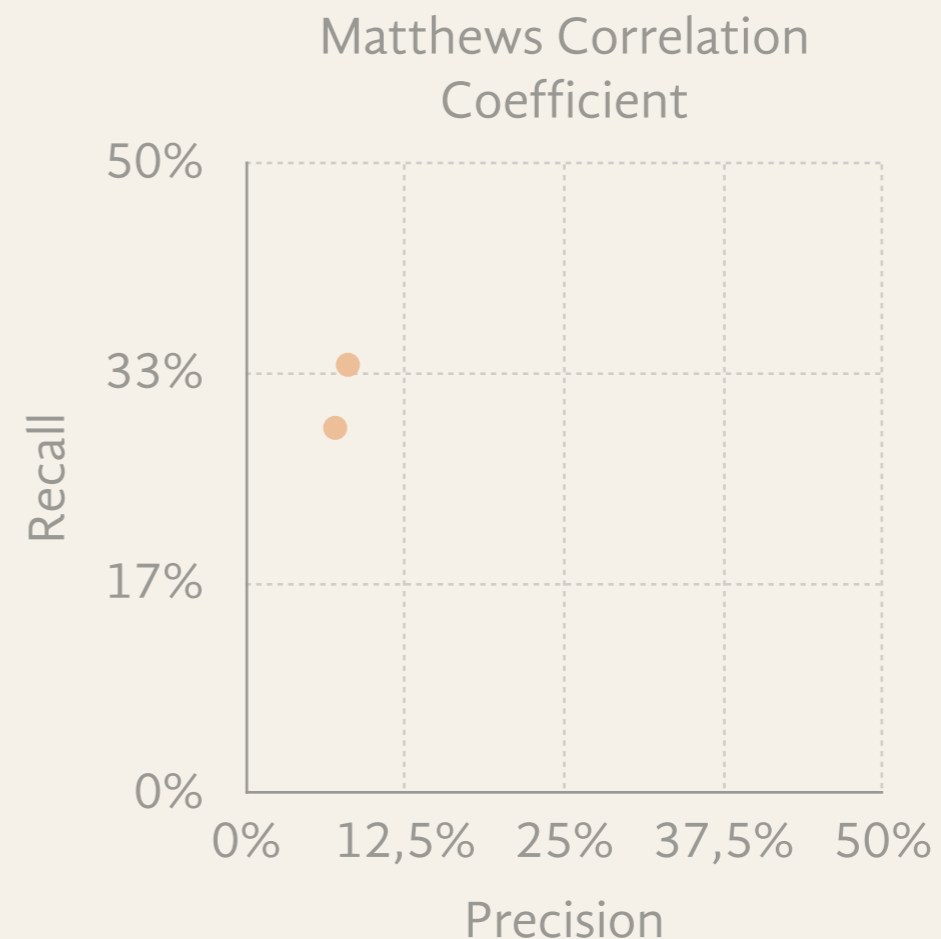
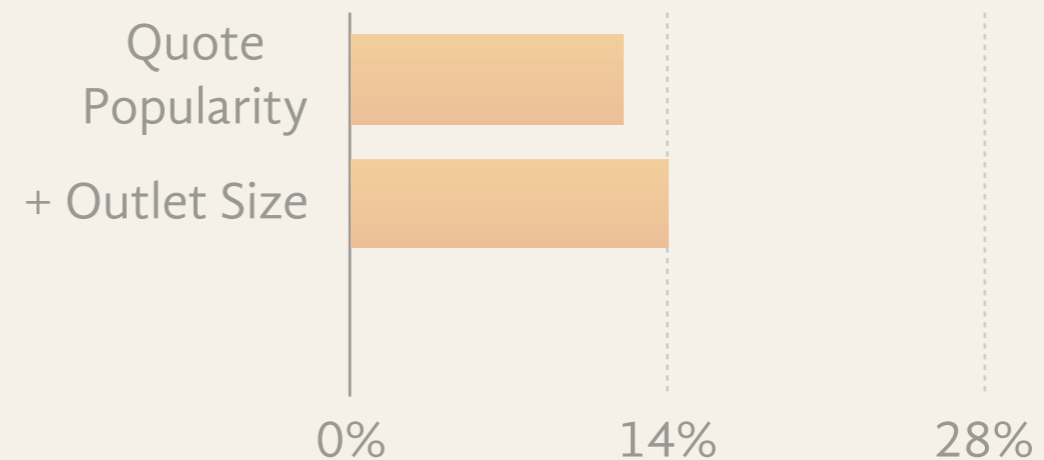
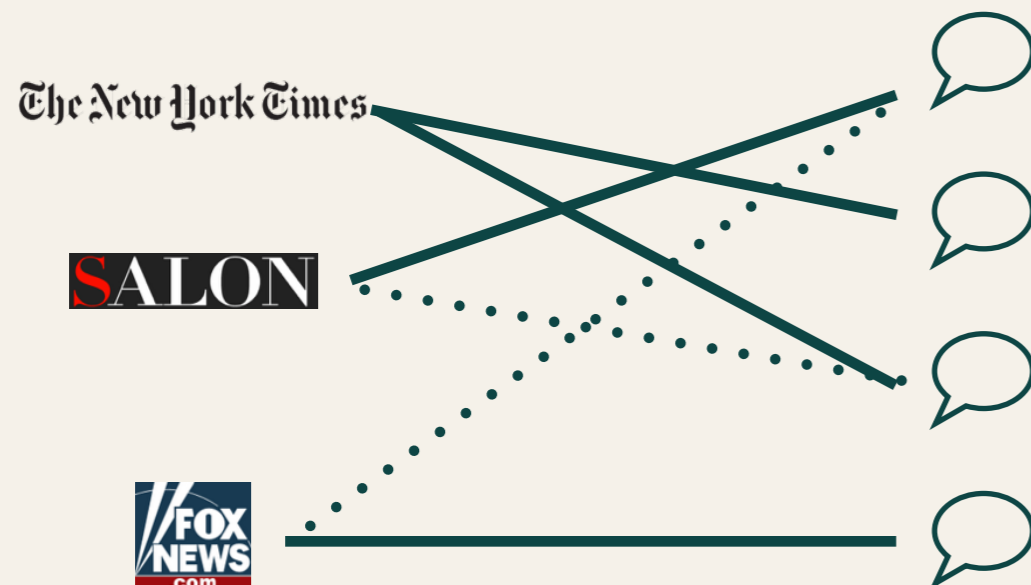
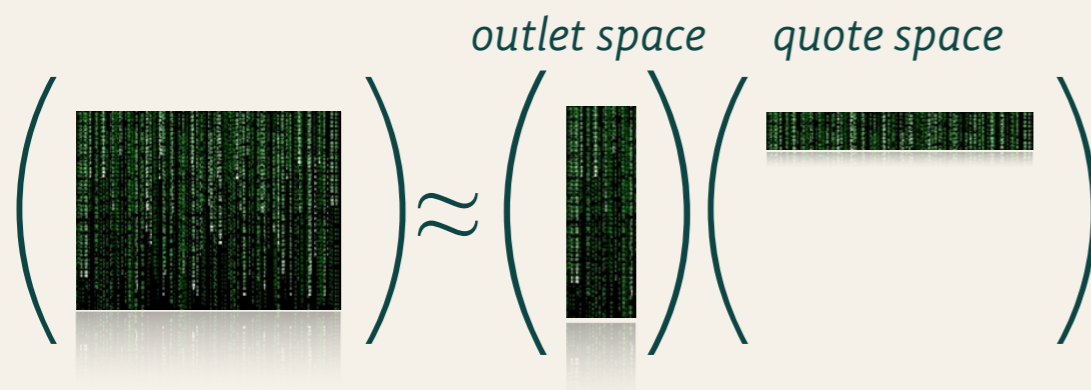


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

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

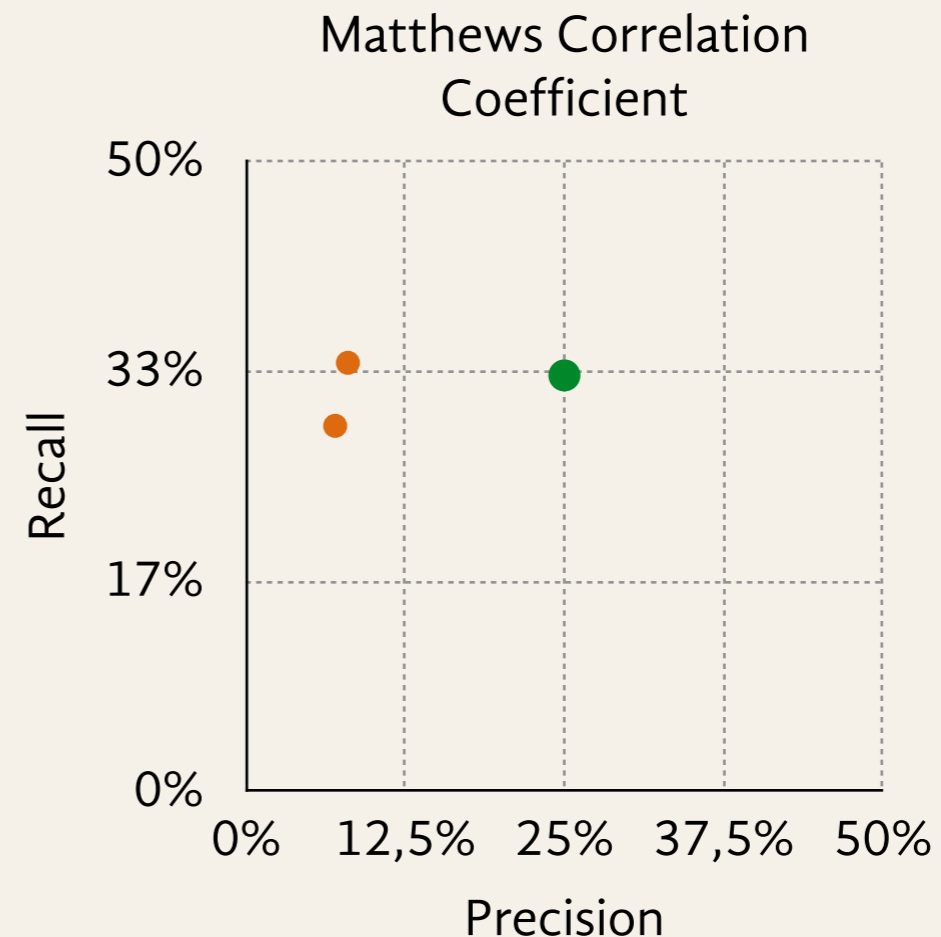
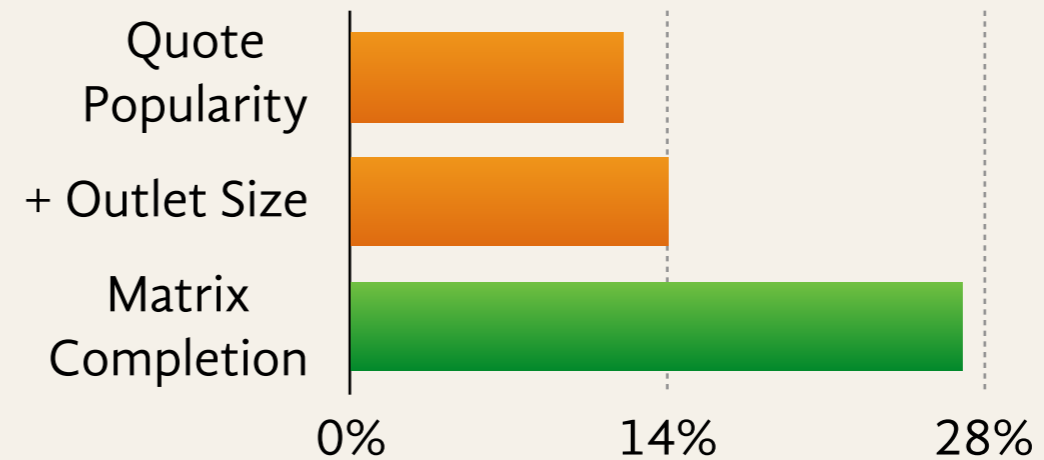
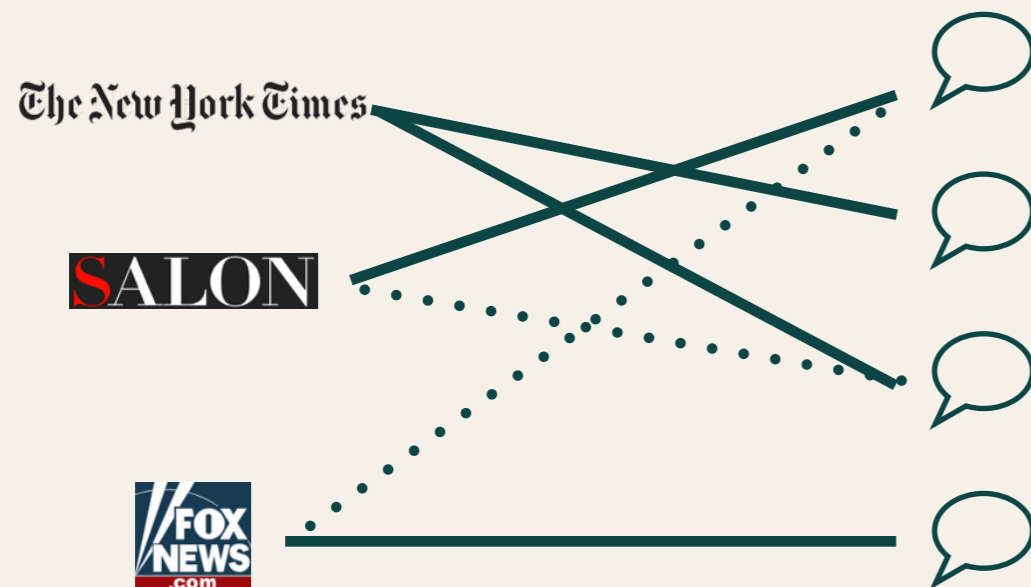
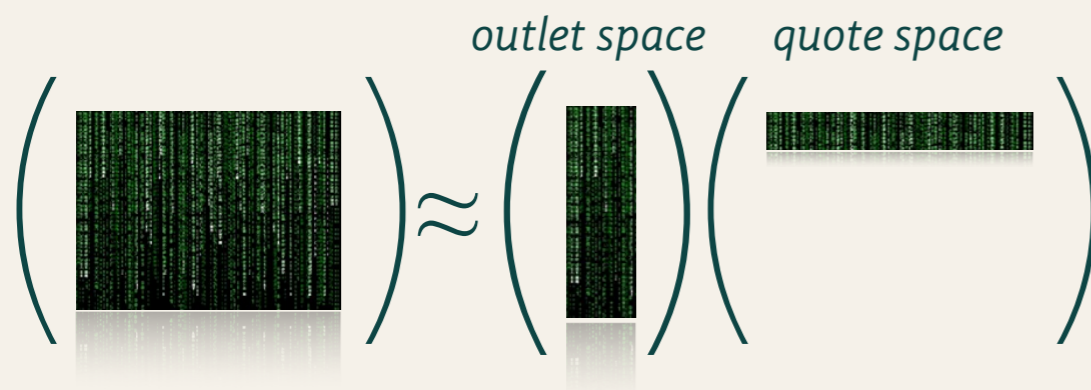


Are quoting patterns *systematic*?

Can we predict quoting?

Matrix completion by
low-rank factorization

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



Are quoting patterns *systematic*?

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Are quoting patterns *systematic*?

Do

on?

How do

language?

Yes!

**There is systematicity
beyond newsworthiness.**

Are quoting patterns *systematic*? Yes!

Do they correspond to *intuition*?

How do the patterns manifest in *language*?

Do they correspond to *intuition*?

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

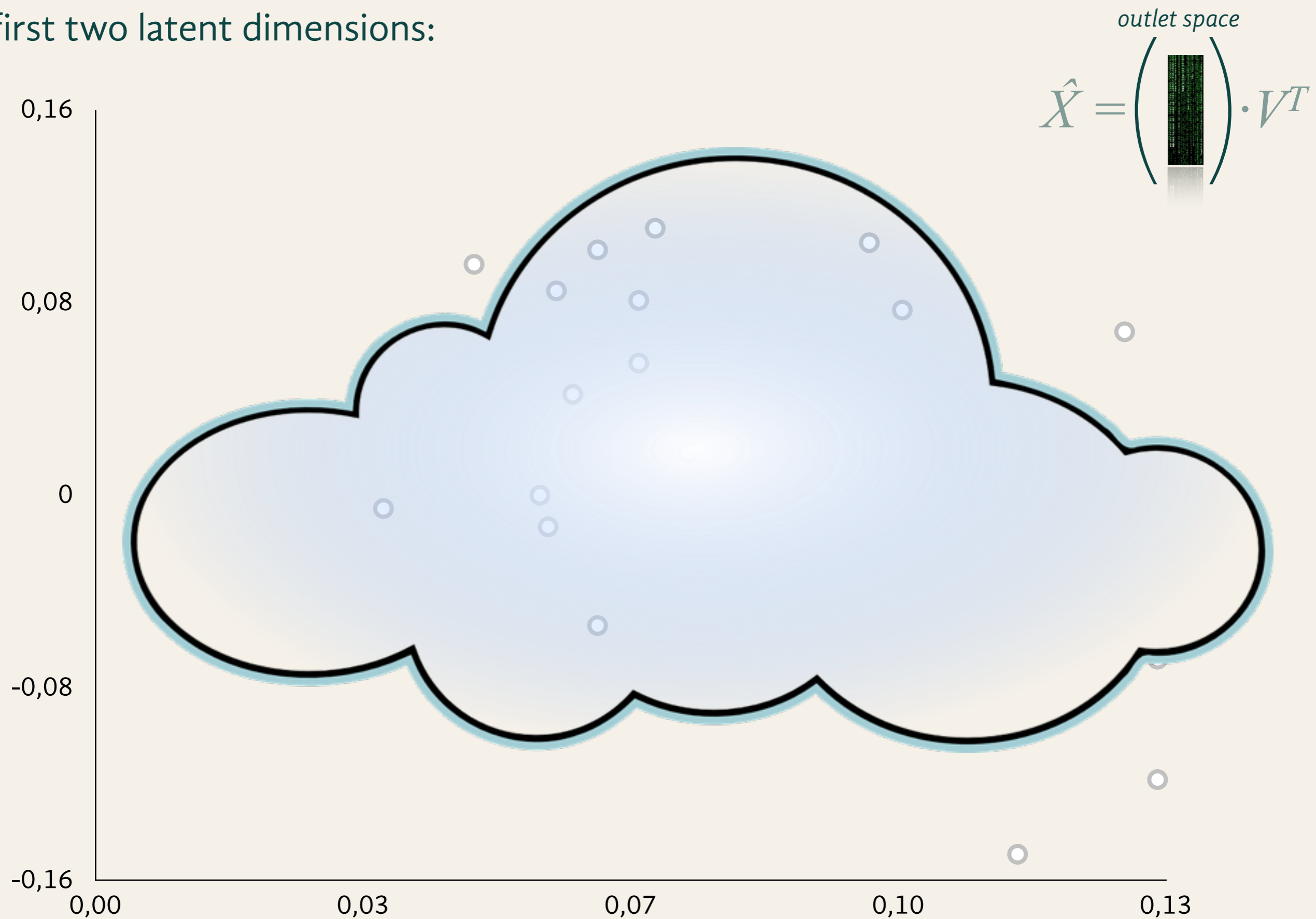
$$\left(\begin{array}{c} \text{matrix} \end{array} \right) \approx \left(\begin{array}{c} \text{outlet space} \\ \text{matrix} \end{array} \right) \left(\begin{array}{c} \text{quote space} \\ \text{matrix} \end{array} \right)$$

Do they correspond to *intuition*?

$$\begin{pmatrix} \text{inlet space} \end{pmatrix} \approx \begin{pmatrix} \text{outlet space} \end{pmatrix} \begin{pmatrix} \text{quote space} \end{pmatrix}$$

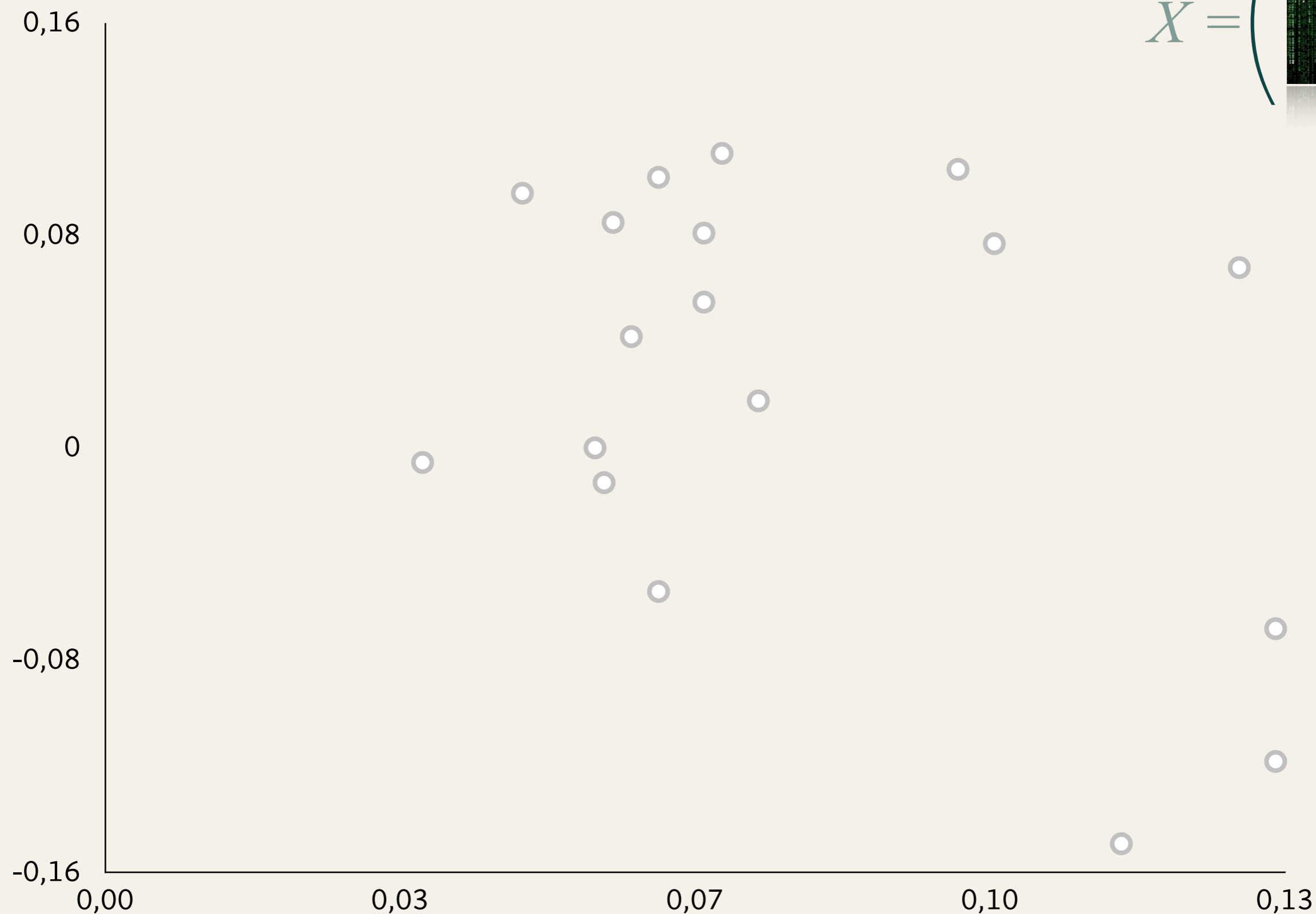
Do they correspond to *intuition*?

The first two latent dimensions:



Do they correspond to *intuition*?

The first two latent dimensions:

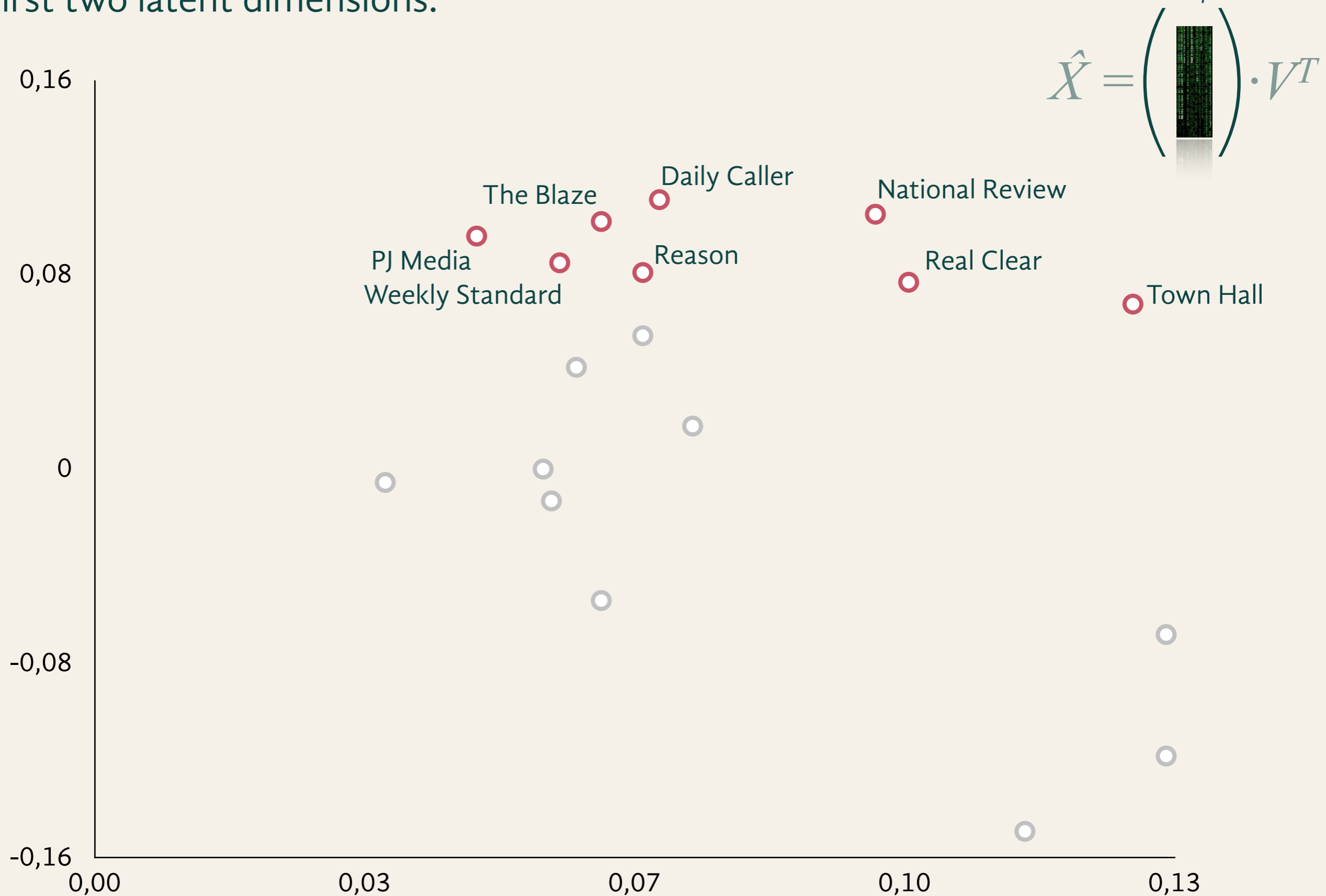


outlet space

$$\hat{X} = \begin{pmatrix} \text{outlet space} \end{pmatrix} \cdot V^T$$

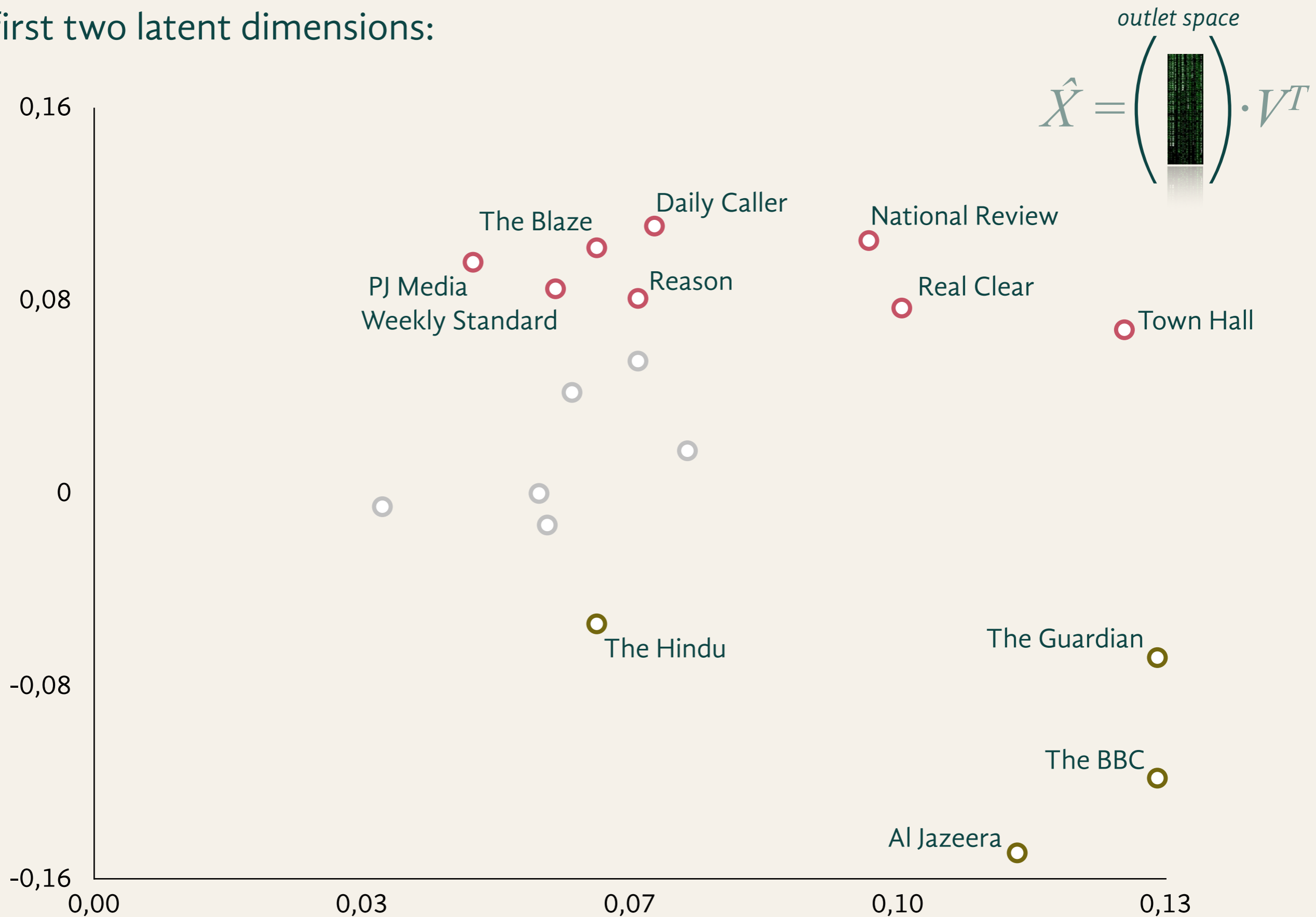
Do they correspond to *intuition*?

The first two latent dimensions:



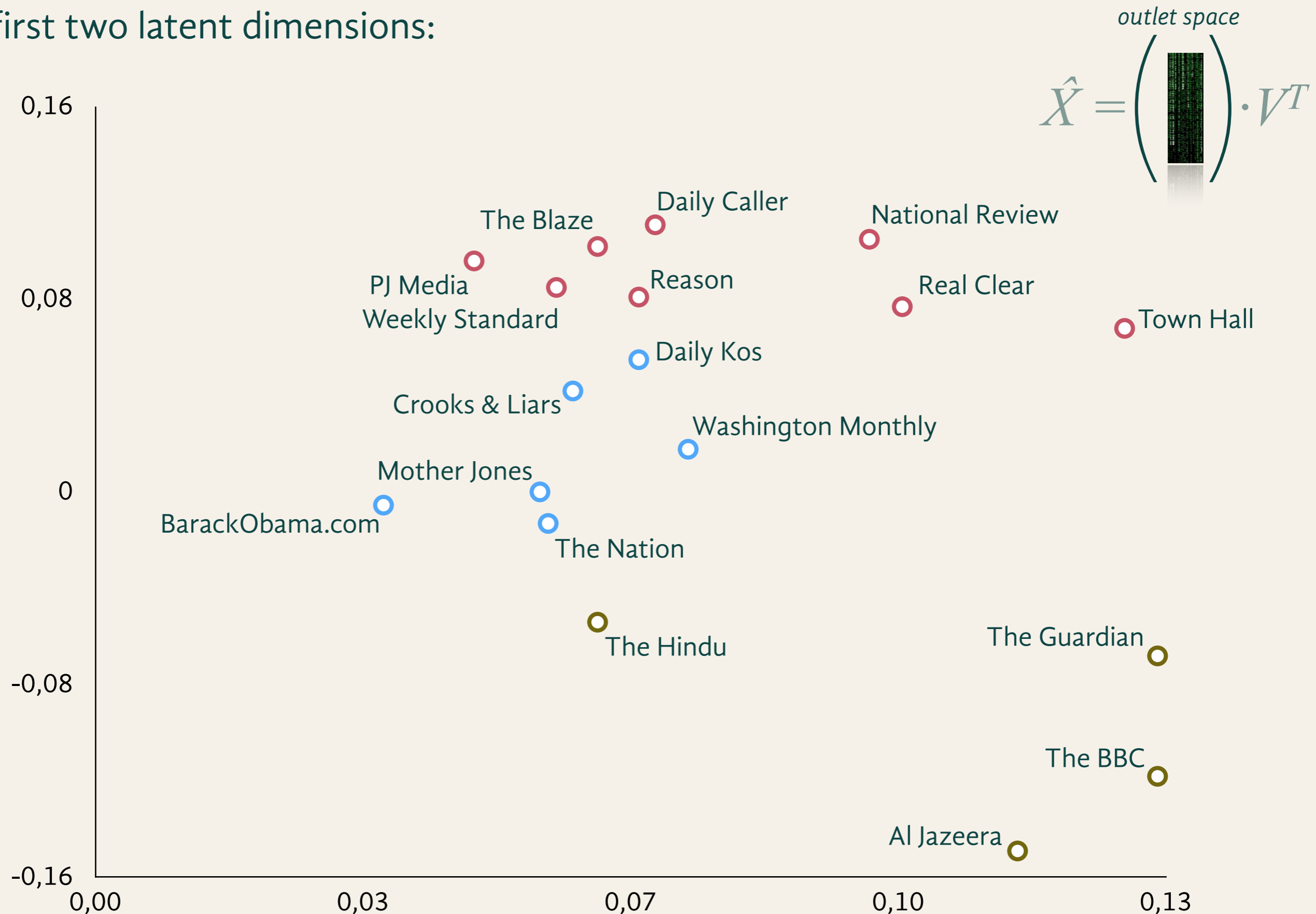
Do they correspond to *intuition*?

The first two latent dimensions:



Do they correspond to *intuition*?

The first two latent dimensions:



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:



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*?

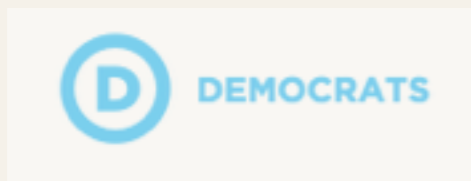
At the top of dimension 2:



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 *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*?

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

$$\left(\begin{array}{c} \text{matrix of green code} \end{array} \right) \approx \left(\begin{array}{c} \text{column of green code} \end{array} \right) \left(\begin{array}{c} \text{row of green code} \end{array} \right)$$

outlet space quote space

How do the patterns manifest in *language*?

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

outlet space

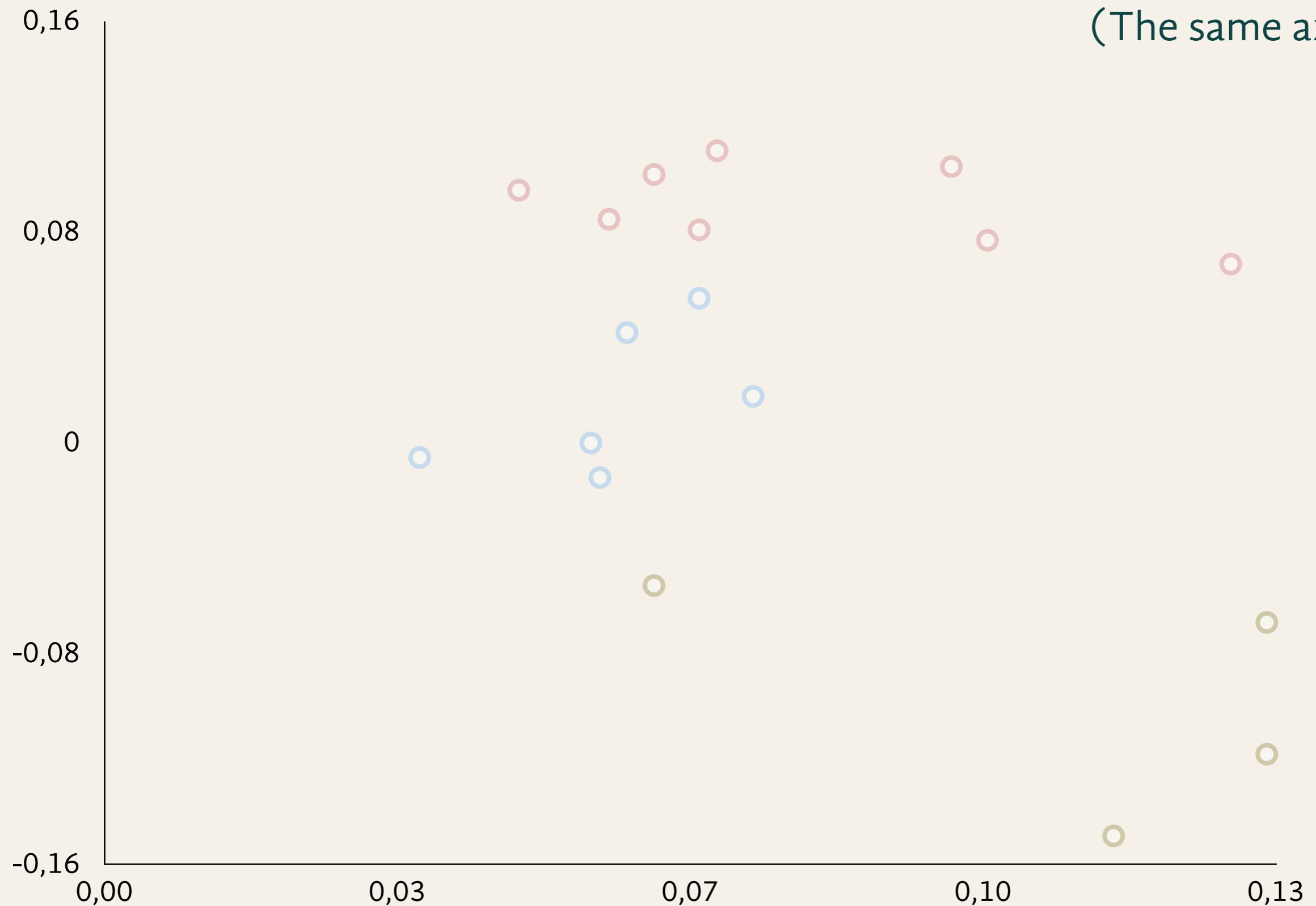
quote space

The diagram shows the equation $X \approx \hat{X} = U \cdot V^T$. Below the equation, there are three matrix representations. The first matrix, X , is a wide, short matrix of green code characters, representing the input space. The second matrix, \hat{X} , is a tall, narrow matrix of green code characters, representing the outlet space. The third matrix, V^T , is a wide, short matrix of green code characters, representing the quote space. The first two matrices are separated by an approximation symbol \approx . The third matrix is enclosed in a red box. The labels 'outlet space' and 'quote space' are positioned below the second and third matrices, respectively.

How do the patterns manifest in *language*?

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

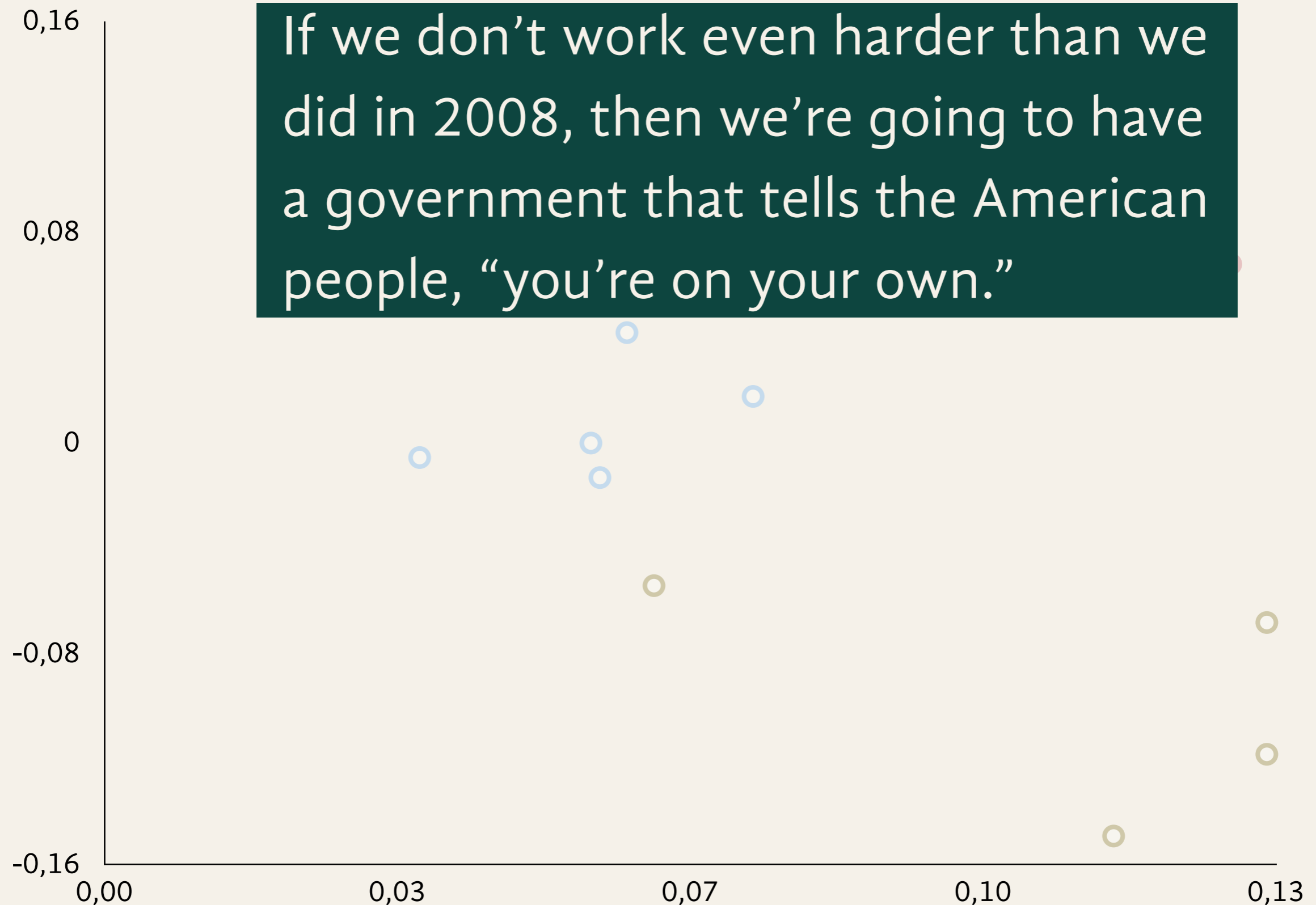
(The same axes!)



How do the patterns manifest in *language*?

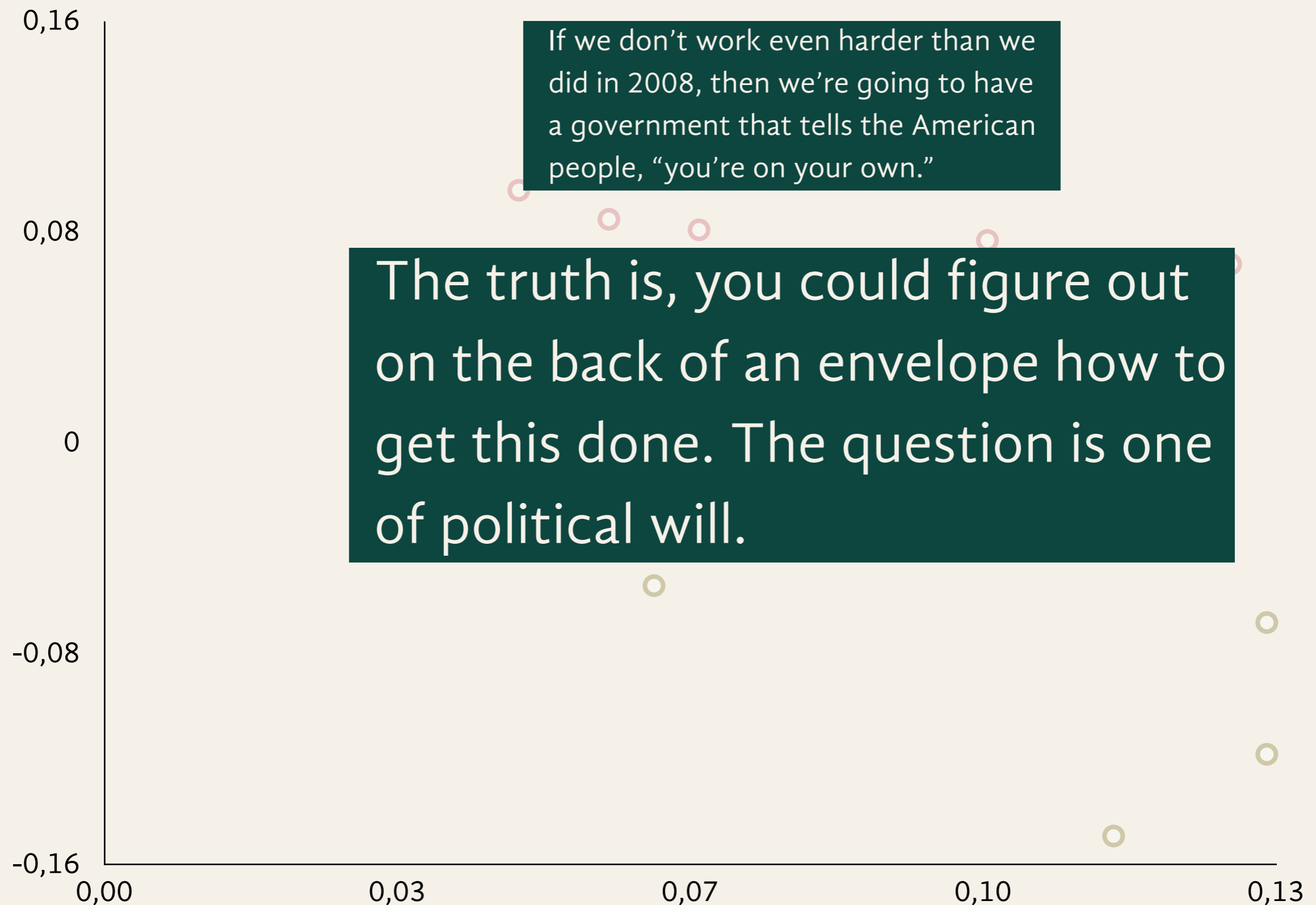
$$\hat{X} = U \cdot \left(\begin{array}{c} \text{quote space} \\ \text{[Matrix of green characters]} \end{array} \right)$$

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."



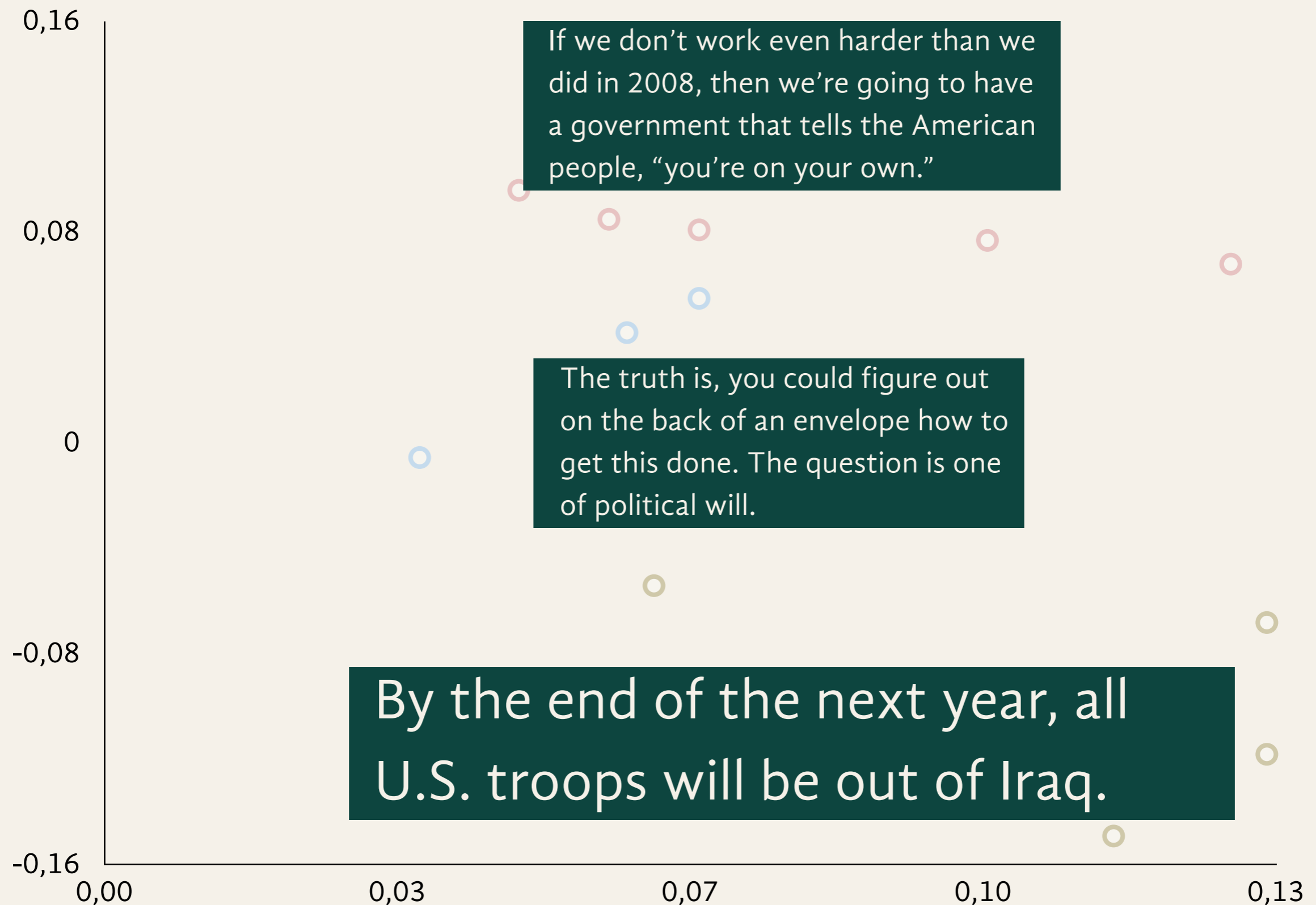
How do the patterns manifest in *language*?

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



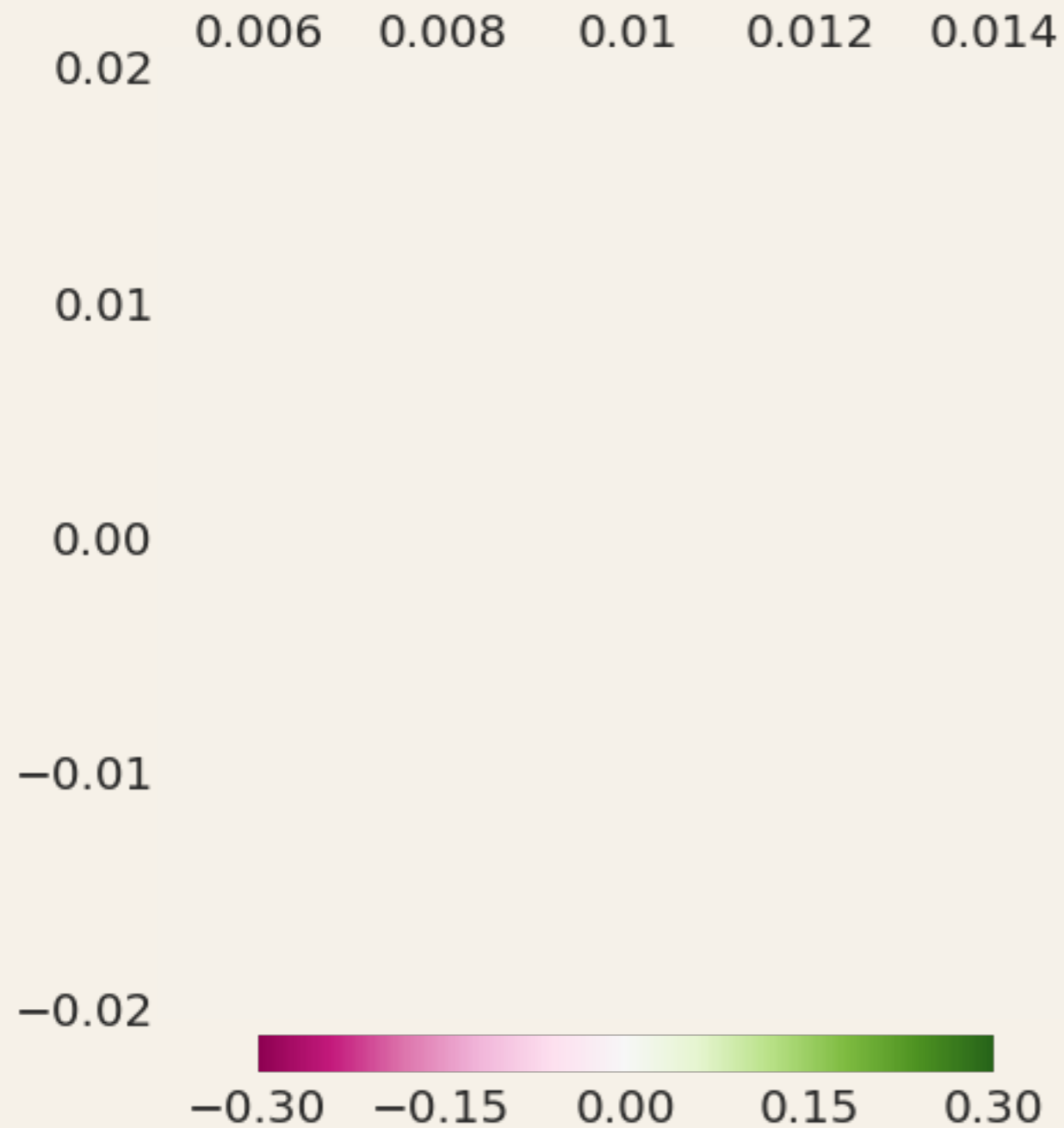
How do the patterns manifest in *language*?

$$\hat{X} = U \cdot \left(\begin{matrix} \text{quote space} \\ \text{[Matrix of green characters]} \end{matrix} \right)$$



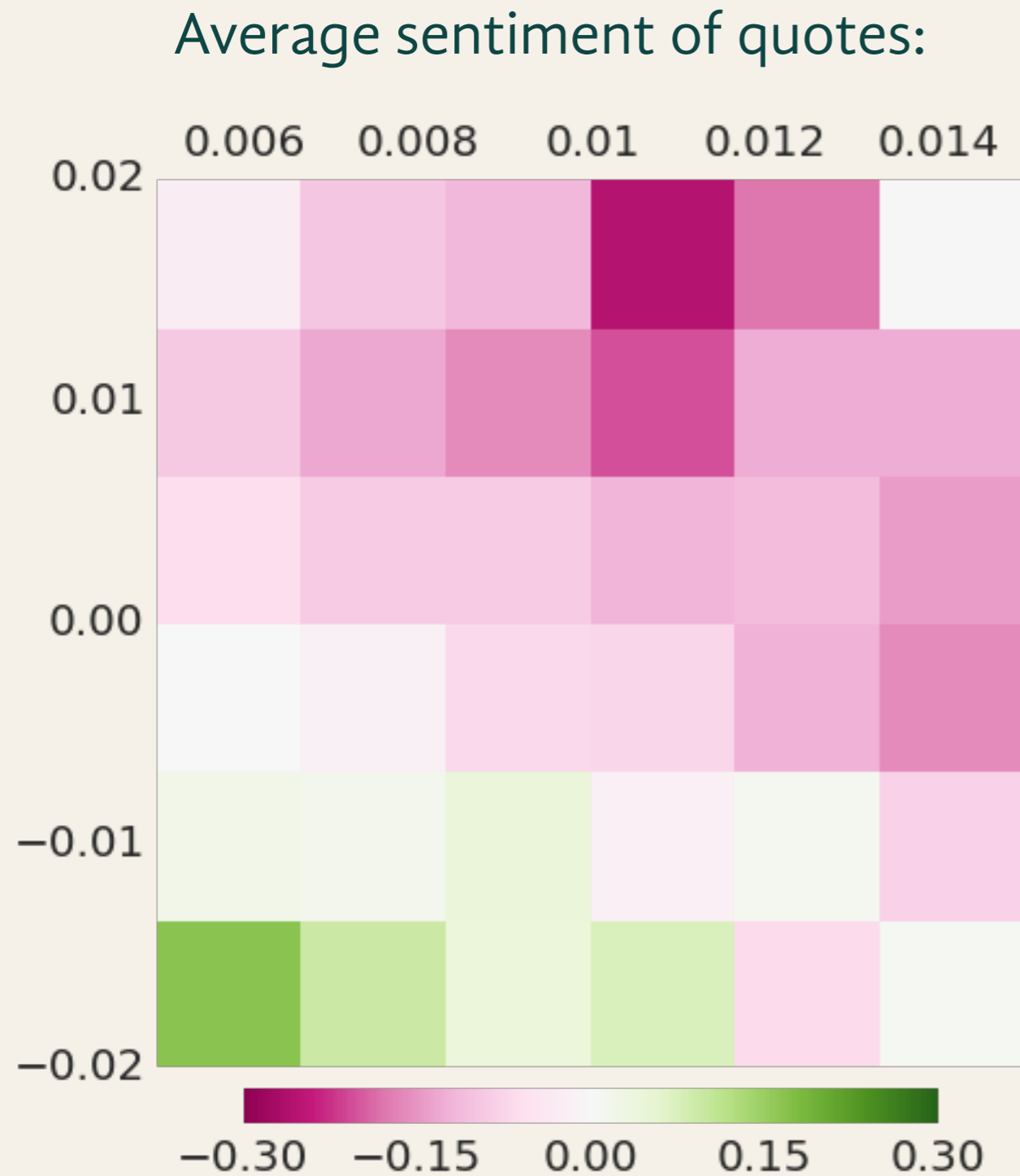
How do the patterns manifest in *language*?

Average sentiment of quotes:



(The same axes!)

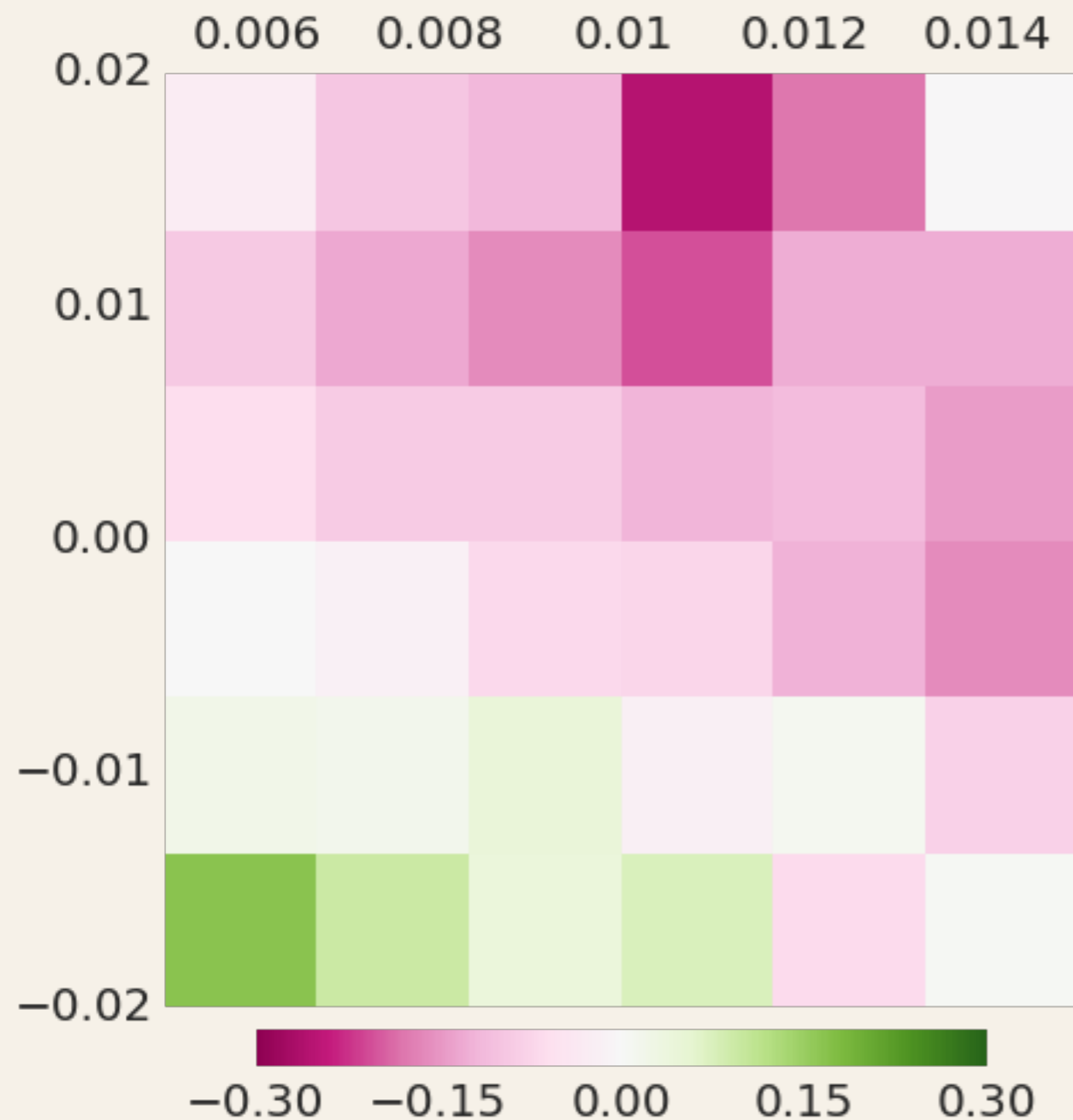
How do the patterns manifest in *language*?



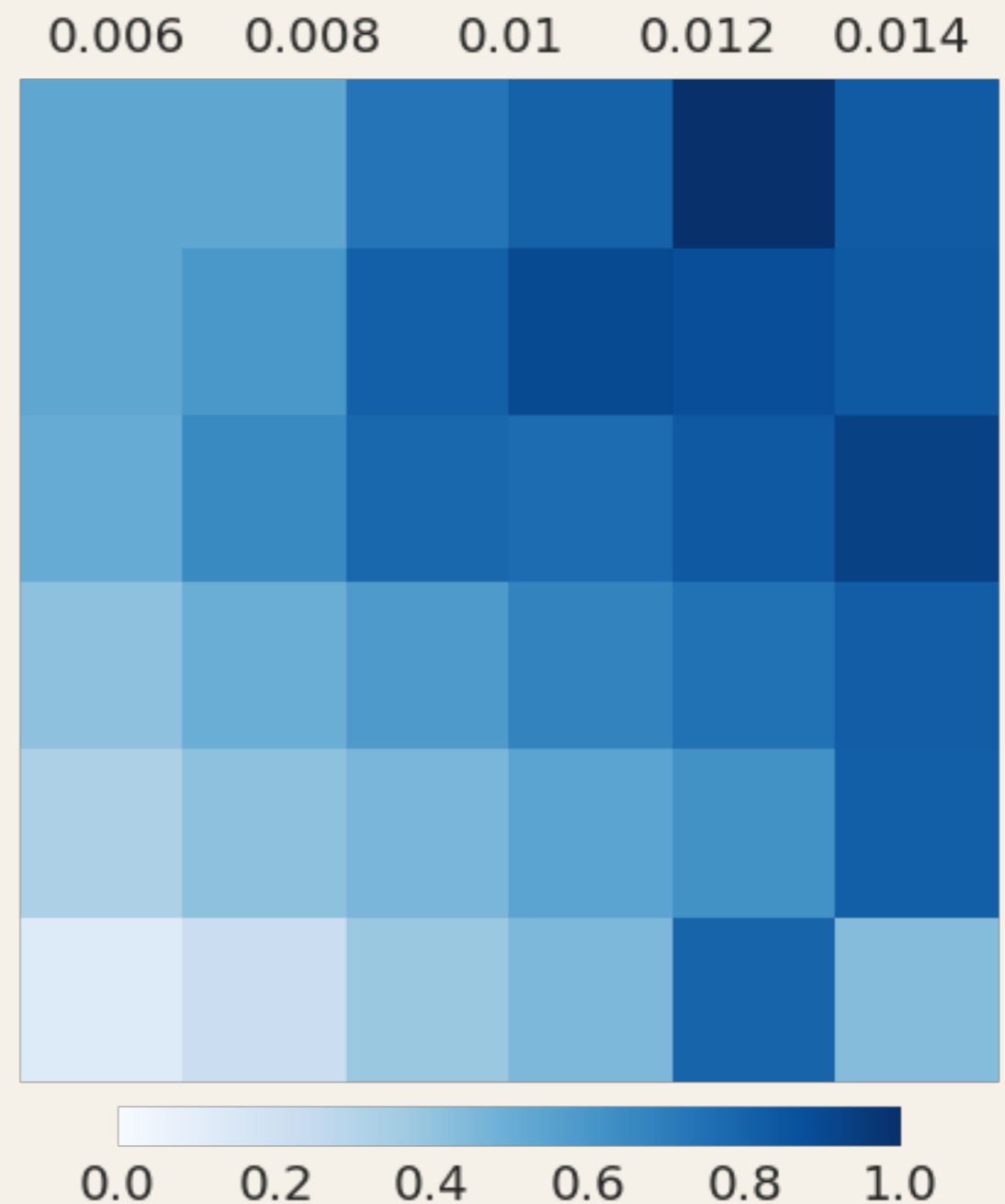
(The same axes!)

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*?

**Linguistic properties
map consistently
to the bias space.**

Are qu

? Yes!

Do they

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/>

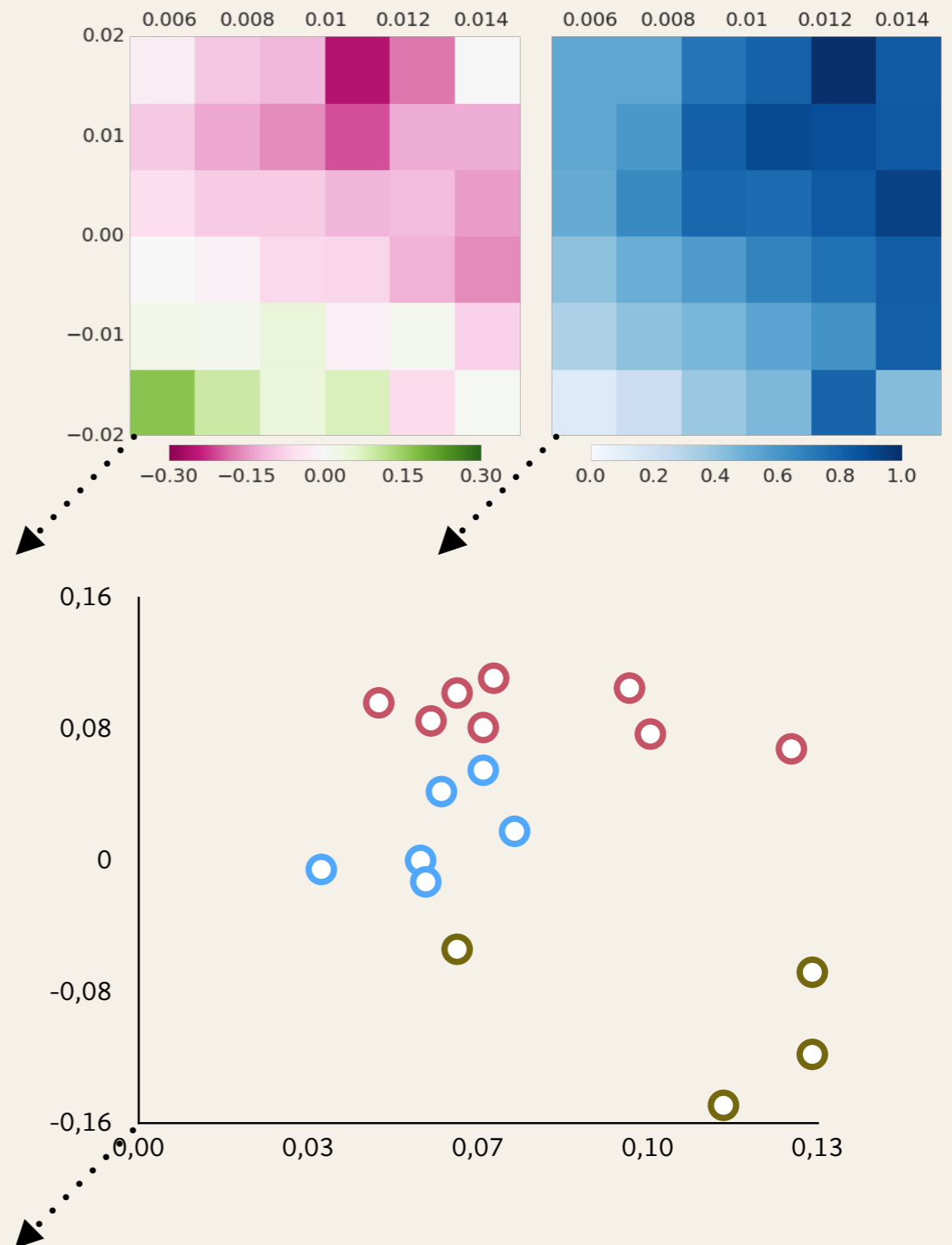
Conclusion

Quoting patterns are systematic:
We can predict them.

Selection bias is multidimensional.

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

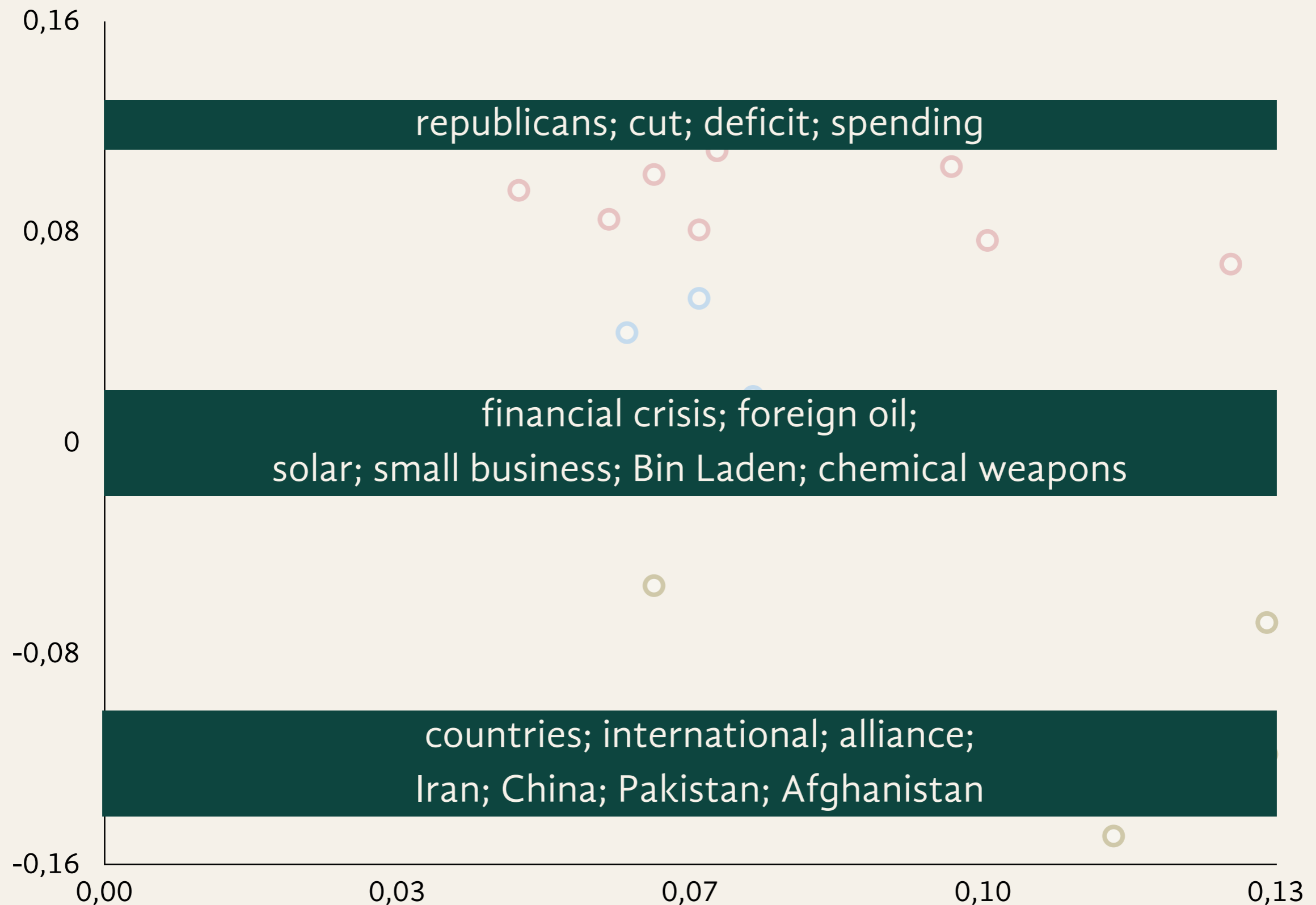
Online visualization & data:
<http://snap.stanford.edu/quotus/>



extra slides

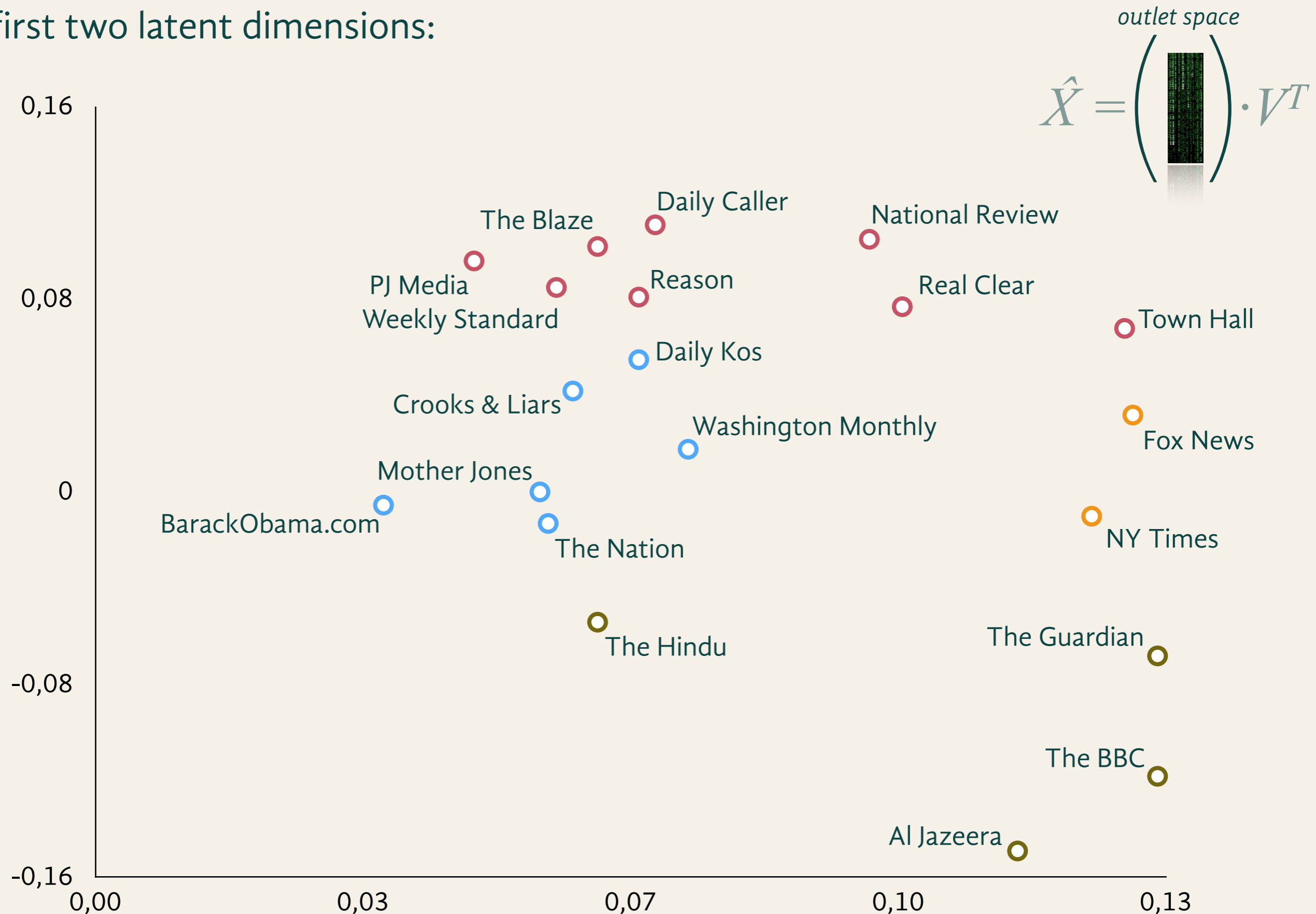
Can *language* characterize the bias?

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



Do they correspond to *intuition*?

The first two latent dimensions:



Can *language* characterize the bias?

$$\hat{X} = U \cdot \left(\begin{matrix} \text{quote space} \\ \text{matrix} \end{matrix} \right)$$

