

# Fanning the Flames of Hate:

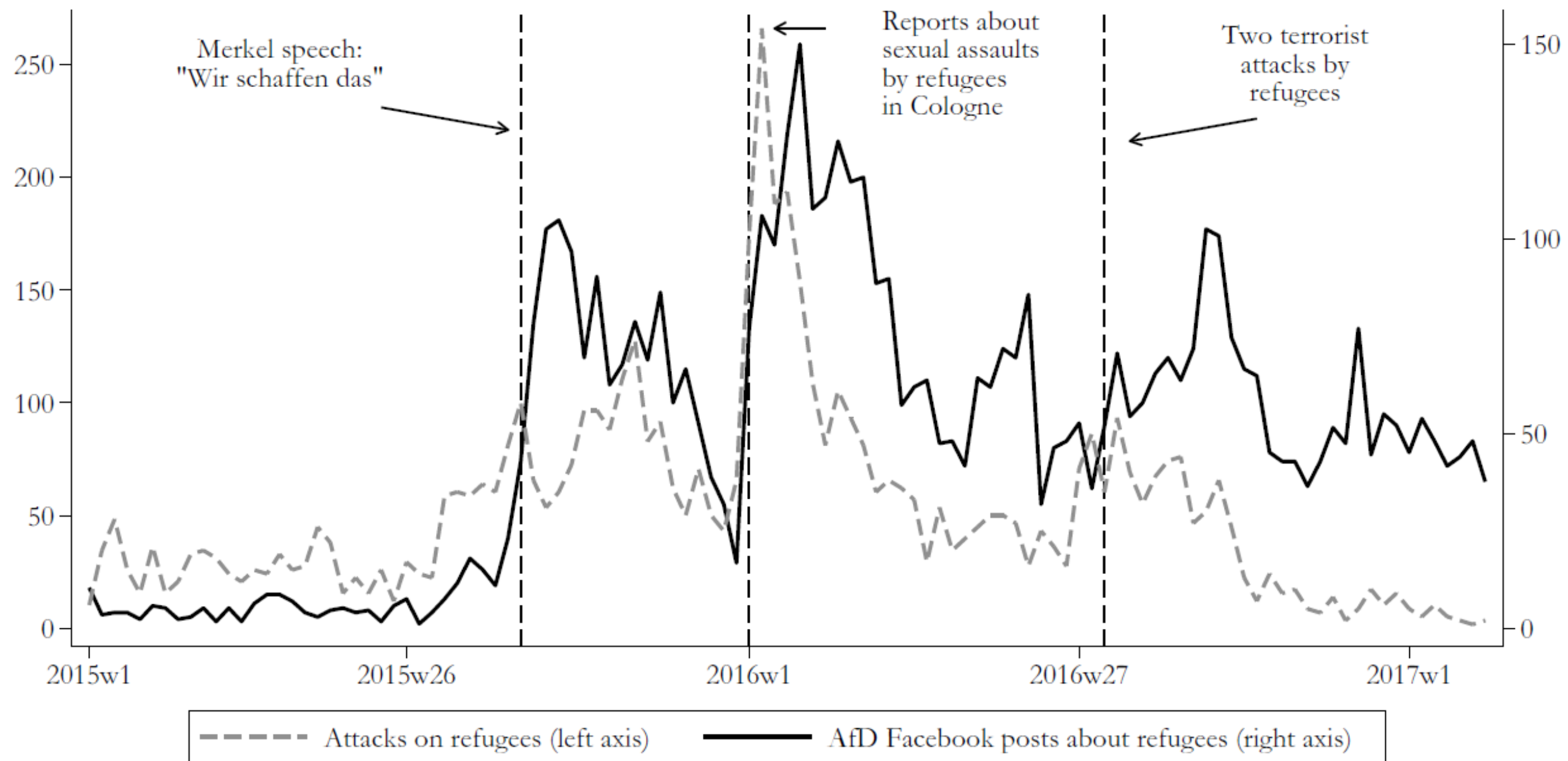
*Social Media and Hate Crime*

*Karsten Müller, Princeton University*

*Carlo Schwarz, Bocconi University*

# Motivation

- The paper investigates the connection between Facebook and anti-refugee incidents in Germany during the refugee crisis (2015-2017)
- Increasing worries about the way social media influences people's behaviour
- Led the German government to introduce the "Netzwerkdurchsetzungsgesetz" requiring social media companies to remove hateful content
- Strong correlation between anti-refugee sentiment on Facebook and anti-refugee incidents



# Motivation

- The paper investigates the connection between Facebook and anti-refugee incidents in Germany during the refugee crisis (2015-2017)
- Increasing worries about the way social media influences people's behaviour
- Led the German government to introduce the "Netzwerkdurchsetzungsgesetz" requiring social media companies to remove hateful content
- Strong correlation between anti-refugee sentiment on Facebook and anti-refugee incidents
- Paper aims to show that there is a causal relationship

## Introduction

# Contribution

### Social media effects:

- Hate speech: Oksanen, Holkeri, Näsi, Räsänen (2014); Siegel, Tucker (2017)
- Polarization: Sunstein (2017); Alcott, Gentzkow (2017); Boxell, Gentzkow, Shapiro (2018); Levy (2019)
- Protests: Enikolopov, Makarin, Petrova (2018) **Happiness:** Allcott, Braghieri, Eichmeyer, Gentzkow (2019)
- Political campaigns: Petrova, Sen, Yildirim (2017); Liberini, Redoano, Russo, A. Cuevas, R. Cuevas (2018);
- Hate crimes: Müller, Schwarz (2019); Bursztyn, Enikolopov, Makarin, Petrova (2019)

➡ **We show the importance of social media for the propagation of hate crime**

### Media and Violence:

- Anti-minority violence: Yanagizawa-Drott (2014); DellaVigna, Enikolopov, Mironova, Petrova, Zhuravskaya (2014); Adena, Enikolopov, Petrova, Santarosa, Zhuravskaya (2015)
- Other violence: Dahl, DellaVigna (2009); Card, Dahl (2011); Bhuller, Havnes, Leuven, Mogstad (2013);

➡ **Evidence for link between media and anti-minority violence in a democracy**

➡ **Evidence for the coordination of incidents as a plausible mechanism**

# Outline

## 1) Data

- Highly granular data on 3,335 anti-refugee incidents
- Measures of municipal Facebook usage and anti-refugee sentiment

## 2) Empirical Strategy and Main Results

- Exploit quasi-experimental variation in Facebook access

## 3) Difference between Social Media and Traditional Media

- Tone on social media consistently more negative

## 4) Mechanism

- Findings most consistent with a coordination of incidents

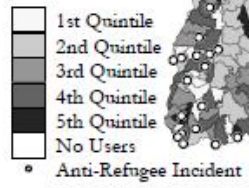
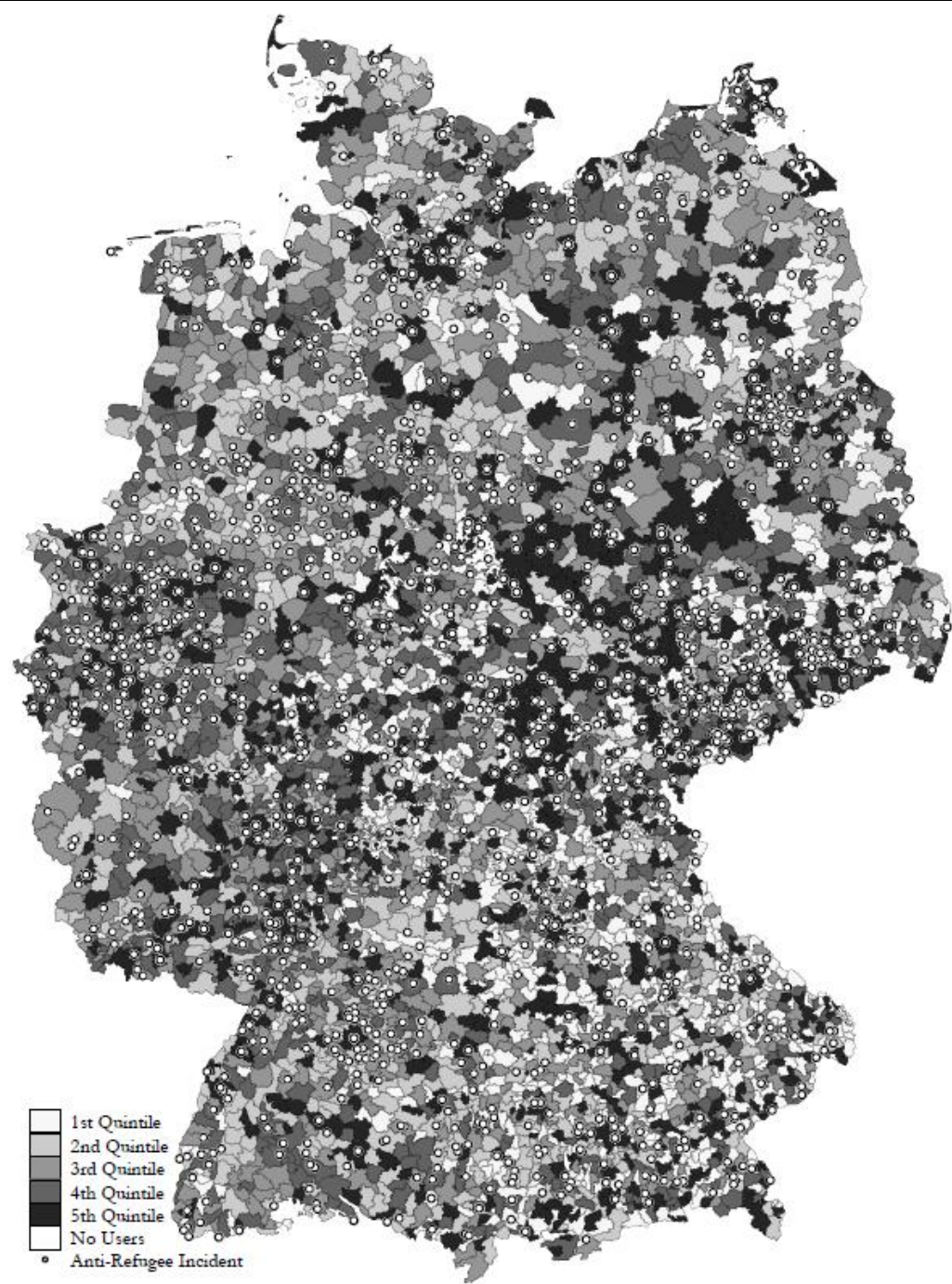
Data

## Data

# Data on anti-refugee incidents

- Data on anti-refugee incidents were collected by the Amadeu Antonio Foundation in collaboration with Pro Asyl (pro asylum NGO)
- Cover 3,335 incidents from January 2015 until February 2017
- All incidents are geo-coded with an exact location and feature a short description
- Data feature 2,226 incidents of property damage to refugee homes, 534 cases of assault, 339 incidents during anti-refugee protests and 225 cases of arson





## Data

# Facebook data

- Scraped information on all posts, comments and likes from the Facebook page of the AfD (far-right anti-refugee party)

Facebook Post with Text

The image shows a screenshot of a Facebook post from the AfD (Alternative für Deutschland) party. The post is dated August 16th and contains a satirical article about a nasal spray developed at the University of Bonn. The article's headline is "Noch mehr Willkommenskultur mit Nasenspray der Uni Bonn! Für die Nebenwirkungen als Kombi mit Pfefferspray?". The post has 1,818 likes and 805 shares. A comment from a user is visible at the bottom, stating "Wird bestimmt an Sylvester in allen größeren Städten an unsere Frauen verteilt werden .....". The post is annotated with red boxes and lines pointing to various elements: the main text, the article image, the number of likes and shares, a user comment, and a list of user-generated posts in the search results on the right.

**Facebook Post with Text**

Alternative für Deutschland

AfD @alternativ fuerde

Startseite

Beiträge

Info

Trau Dich Deutschland

Mitglied werden!

Themen

Fotos

Videos

Livestream

Community

Veranstaltungen

Eine Seite erstellen

Gefällt mir Abonnieren Teilen ...

Registrieren

flüchtling

Besucherbeiträge

Heute um 10:13

<http://www.epochtimes.de/.../bundestag-ge...>

Gefällt mir · Kommentieren

Heute um 10:12

<https://www.welt.de/.../Experten-beobacht...>

Gefällt mir · Kommentieren

Heute um 10:06

"Deutsches Volk", das ist völkisch, rassistisch, geht gar nicht. Naz... Mehr anzeigen

Gefällt 1 Mal

Gefällt mir · Kommentieren

Deutsch · English (UK) · Español · Português (Brasil) · Français (France) +

Datenschutz · Nutzungsbedingungen · Werbung · Datenschutzinfo · Cookies · Mehr +

LEIDER KEINE SATIRE

Noch mehr Willkommenskultur mit Nasenspray der Uni Bonn!

Für die Nebenwirkungen als Kombi mit Pfefferspray?

Alternative für Deutschland

TRAU DICH DEUTSCHLAND!

Gefällt mir Kommentieren Teilen

1.818

805 Mal geteilt

Chronologisch

Vorherige Kommentare anzeigen 52 von 424

Wird bestimmt an Sylvester in allen größeren Städten an unsere Frauen verteilt werden .....

FÜR EINE STARKE Alternative für Deutschland IM BUNDESTAG

Gefällt mir · Antworten · 2 · 17. August um 01:42

2 Antworten

**Nr. of Likes and Shares**

**Facebook Post by Users**

**Comment**

Nr. of Likes and Shares

Facebook Post by Users

Comment

## Data

# Facebook data

- Scraped information on all posts, comments and likes from the Facebook page of the AfD (far-right anti-refugee party)
- Data contain over 176k posts, 290k posts, 510k likes from 93k individuals
- Many posts exhibit clearly negative language towards refugees

"Maybe there is a plot for German genes to be exterminated by the large streams of refugees. But what could be the reason, [revenge for] WW2?"

## Data

# Facebook data

- Scraped information on all posts, comments and likes from the Facebook page of the AfD (far-right anti-refugee party)
- Data contain over 176k posts, 290k posts, 510k likes from 93k individuals
- Many posts exhibit clearly negative language towards refugees
- We use the number of posts mentioning refugees as a proxy for anti-refugee sentiment
- We use the number of AfD users in a municipality as the a proxy for the exposure to anti-refugee content

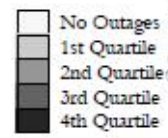
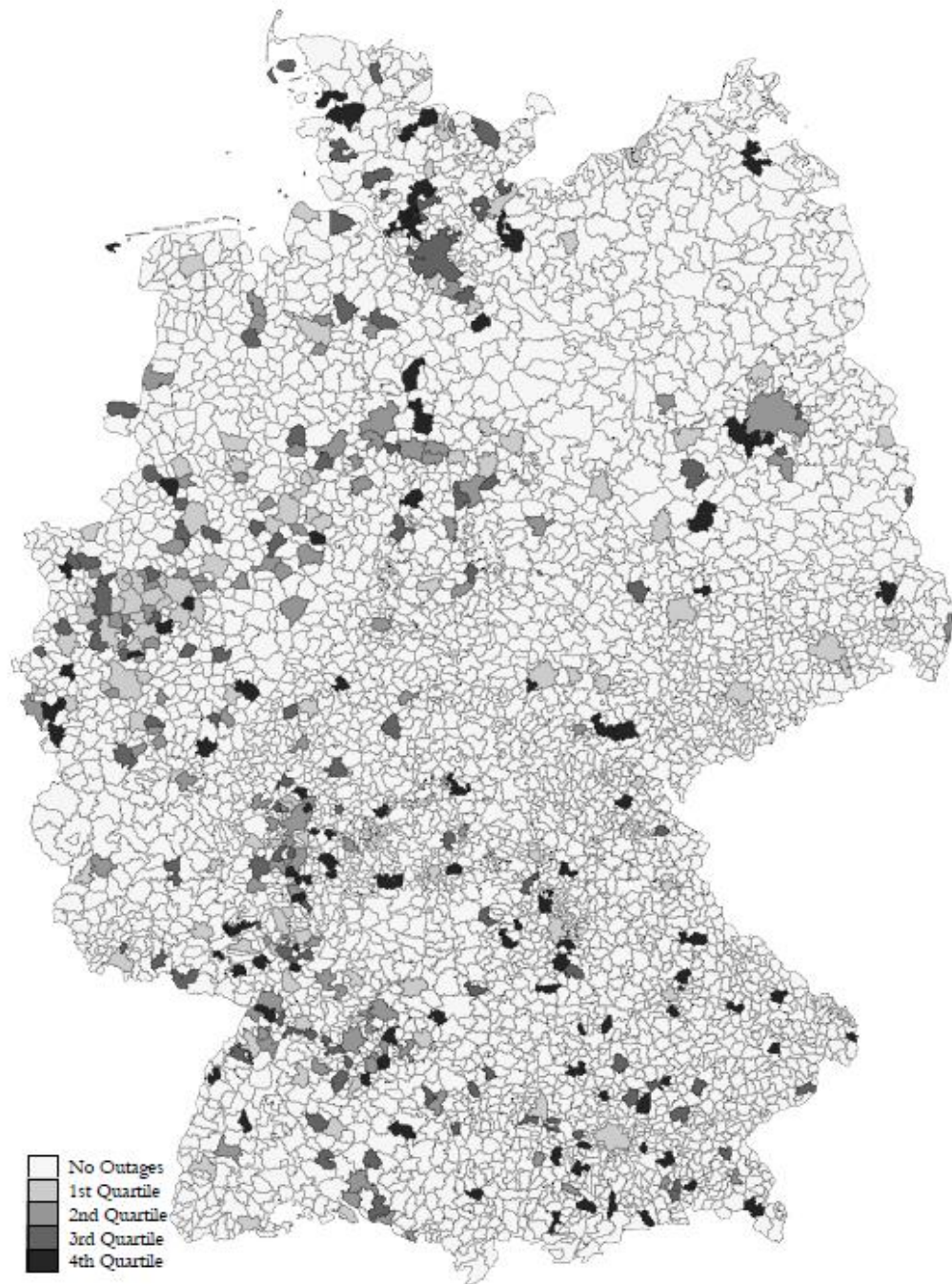
## Data

# Outage Data

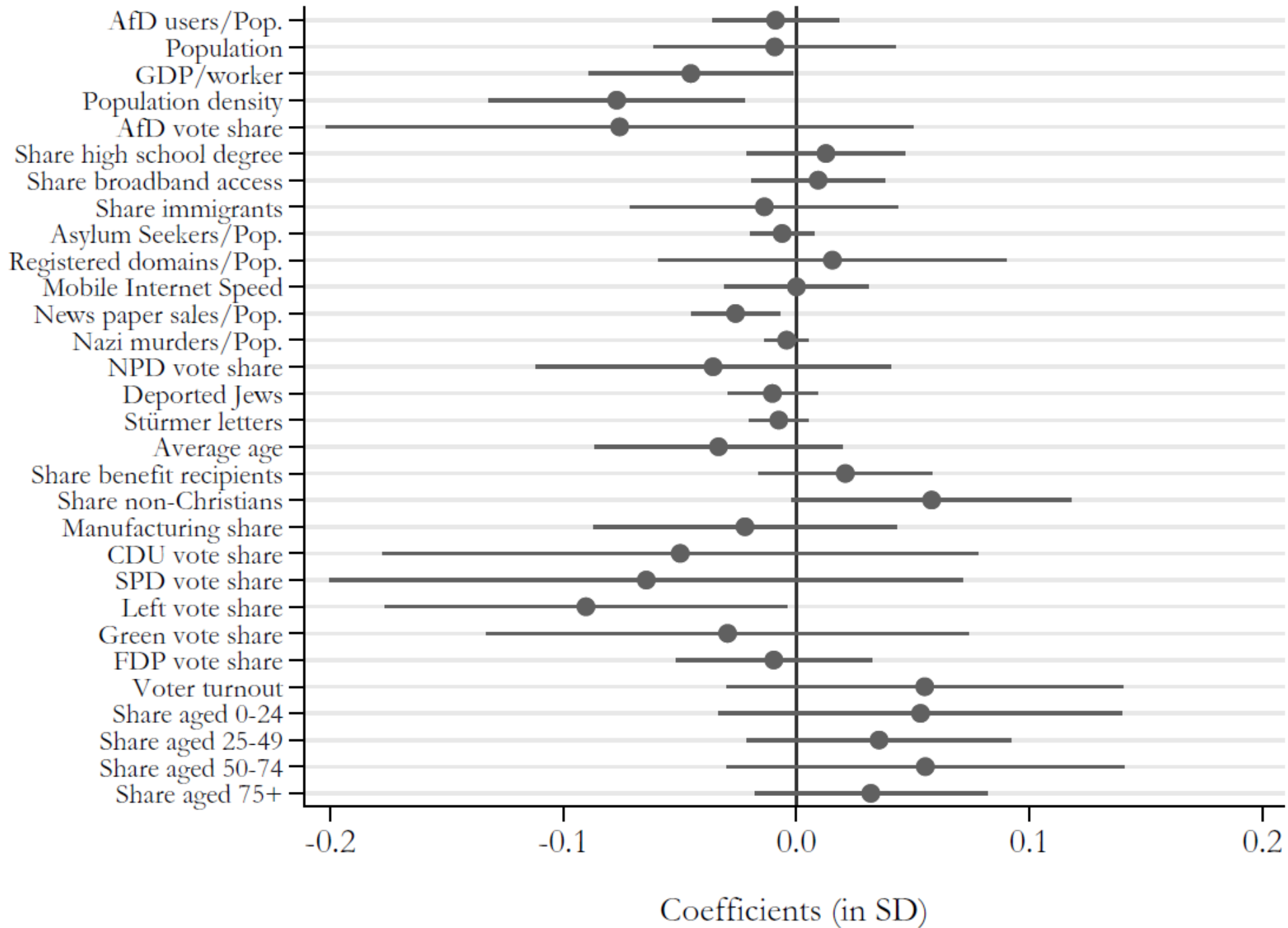
## 1) Internet Outages

- Collected user reports on internet outages from *Heise.de*
- Outages reports are geographically dispersed and do not correlate with observables
- We restrict our analysis to outages that last longer than 24h and define a dummy for the upper quartile of the outages per capita distribution
- This tags 313 municipality-week pairs as suffering an severe internet outage









# Outage Data

## 1) Internet Outages

- Collected user reports on internet outages from *Heise.de*
- Outages reports are geographically dispersed and do not correlate with observables
- We restrict our analysis to outages that last longer than 24h and define a dummy for the upper quartile of the outages per capita distribution
- This tags 313 municipality-week pairs as suffering an severe internet outage

## 2) Facebook Outages

- As additional variation, we use 8 weeks with major Facebook outages
- These outages are uncorrelated with the frequency of internet outages

# Empirical Strategy and Main Results

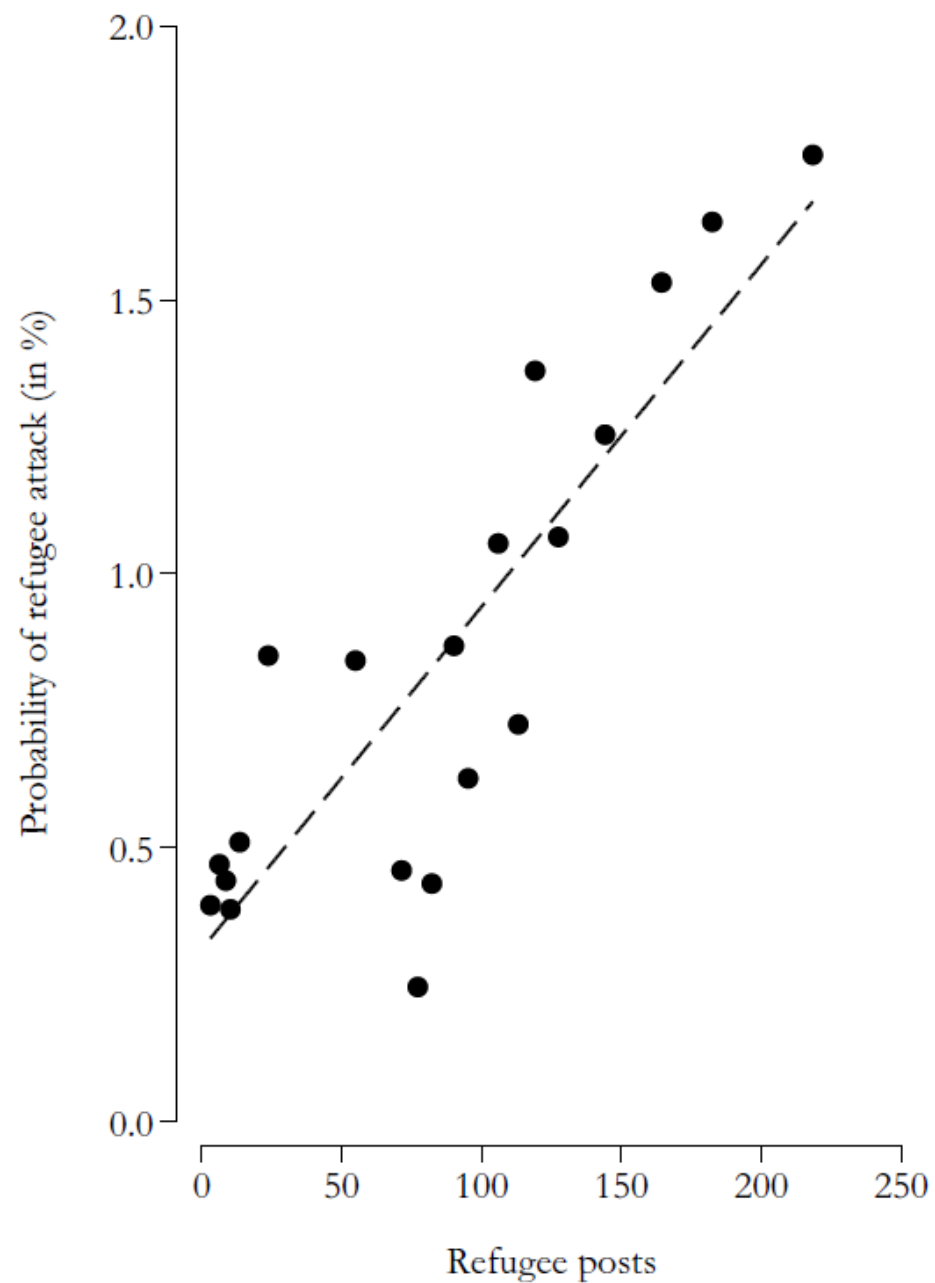
## Empirical strategy

# Empirical strategy

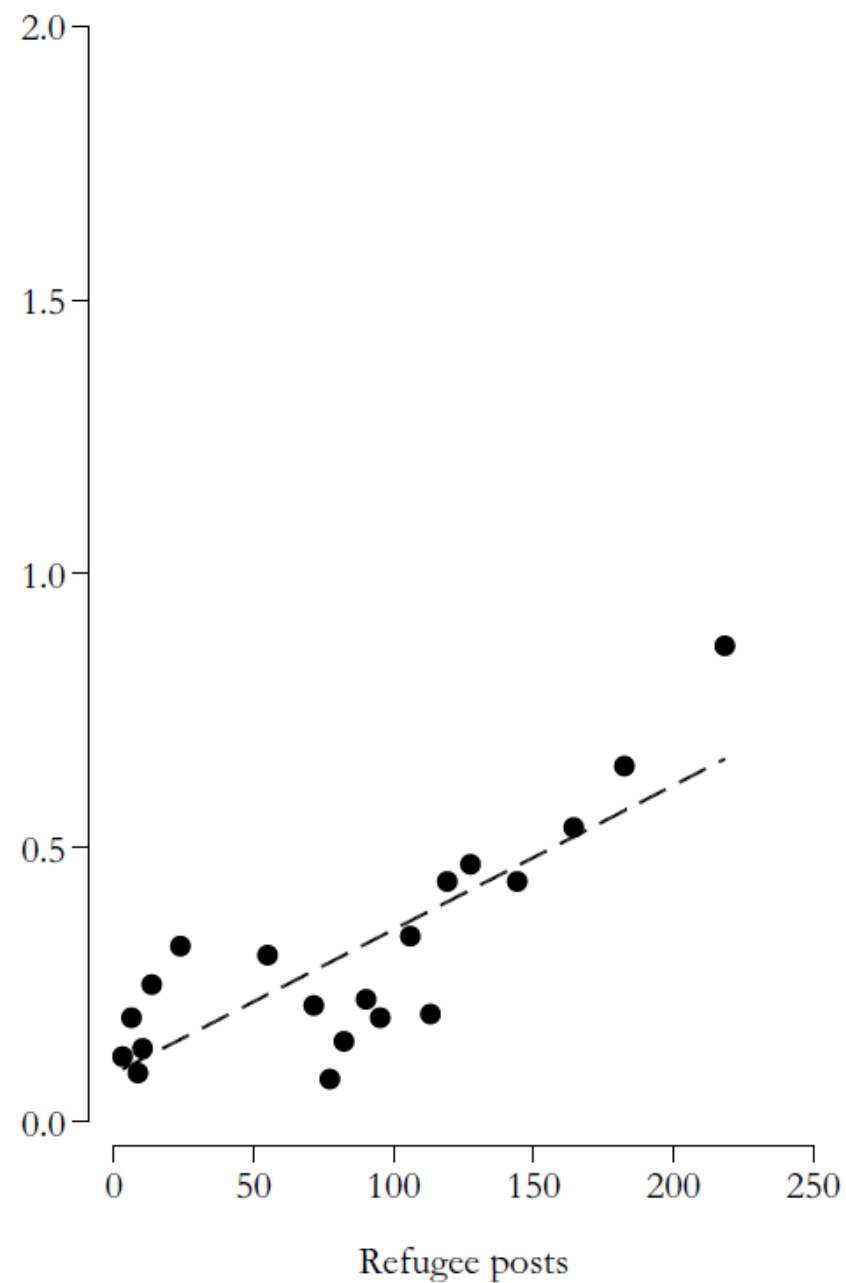
$$\begin{aligned} Attacks_{mw} = & \beta \cdot Refugee Posts_w \times AfD User/Pop_m \\ & + Refugee Post_w \times Controls_m \\ & + Week FE_w + Municipality FE_m + \epsilon_{mw} \end{aligned}$$

- **Attacks** is an indicator for at least one anti-refugee incident
- **Refugee Posts** is the number of posts about refugees on the AfD's Facebook page in week  $w$
- **AfD User** is the number of users on the AfD's Facebook page in municipality  $m$
- **Standard errors** are clustered at the municipality level (robust to alternatives)

a) AfD users/Pop.  $\geq$  Median



b) AfD users/Pop.  $<$  Median



## Main results

# Baseline correlations

	Additional interacted controls						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline controls	Right Wing controls	Media controls	Socio-economic controls	2017 vote controls	Age structure controls	All controls
AfD users/Pop. $\times$ Refugee posts	0.024*** (0.009)	0.020** (0.008)	0.023** (0.009)	0.024** (0.009)	0.021** (0.009)	0.023** (0.009)	0.016** (0.008)
Observations	479,964	479,964	479,964	474,303	479,964	476,856	474,303
R-squared	0.082	0.083	0.082	0.083	0.083	0.083	0.084
Municipalities	4324	4324	4324	4273	4324	4296	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] $\times$ Posts	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Right-wing controls [4] $\times$ Posts		Yes					Yes
Media controls [4] $\times$ Posts			Yes				Yes
Socio-econ. controls [4] $\times$ Posts				Yes			Yes
Election controls [7] $\times$ Posts					Yes		Yes
Age controls [4] $\times$ Posts						Yes	Yes

## Empirical strategy

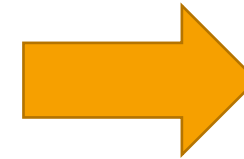
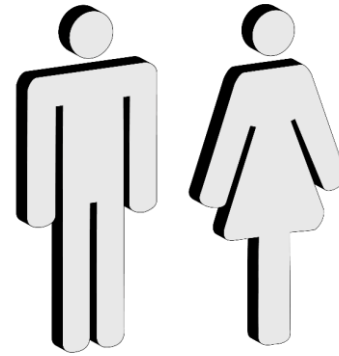
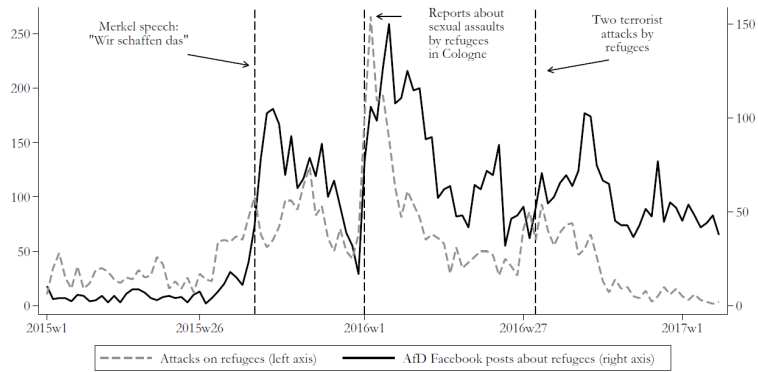
# Interpretation

- The results indicate that anti-refugee incidents disproportionately increase in municipalities with many AfD Facebook users at times high anti-refugee sentiment
- This correlation is not easily explained by observable characteristics of the municipalities
- There remain obvious concerns about unobservable characteristics
- **Solution:** exploit quasi-experiment variation to Facebook exposure created by internet and Facebook outages

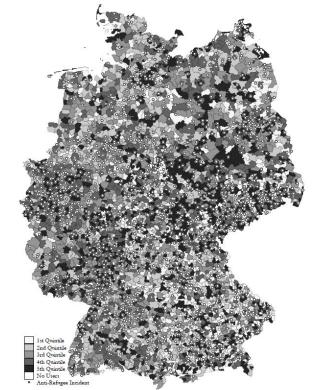
## Empirical strategy

# Social media and hate crime

### Anti-refugee sentiment on social media



### Anti-refugee incidents



Outage



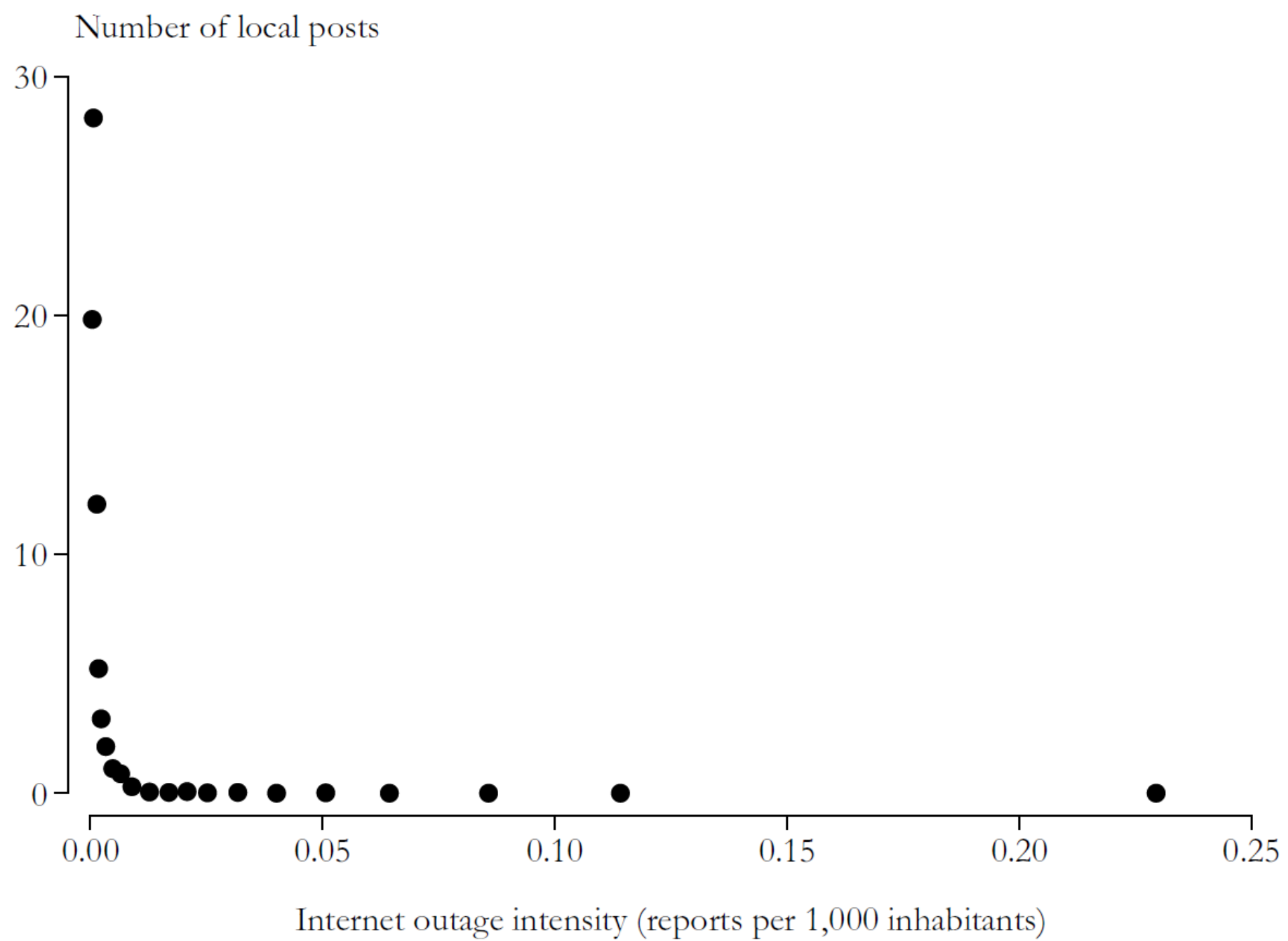


## Main results

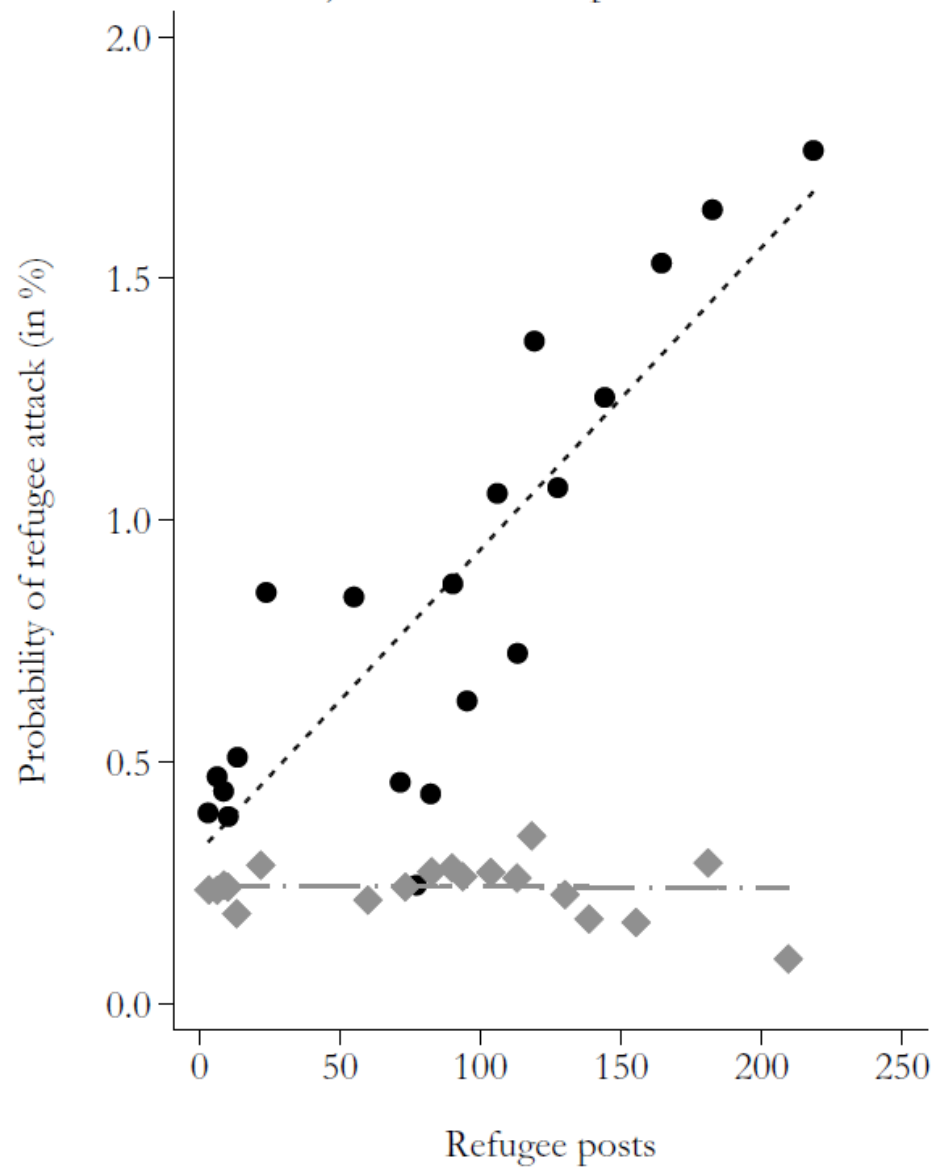
# Internet outages

$$\begin{aligned} Attacks_{mw} = & \beta \cdot Refugee Posts_w \times AfD User/Pop_m \\ & + \lambda \cdot Refugee Posts_w \times AfD User/Pop_m \times Outage_{mw} \\ & + [\dots] \\ & + Refugee Post_w \times Controls_m \\ & + Week FE_w + Municipality FE_m + \epsilon_{mw} \end{aligned}$$

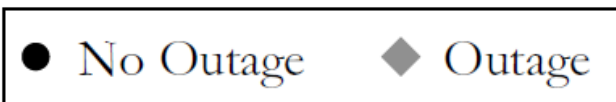
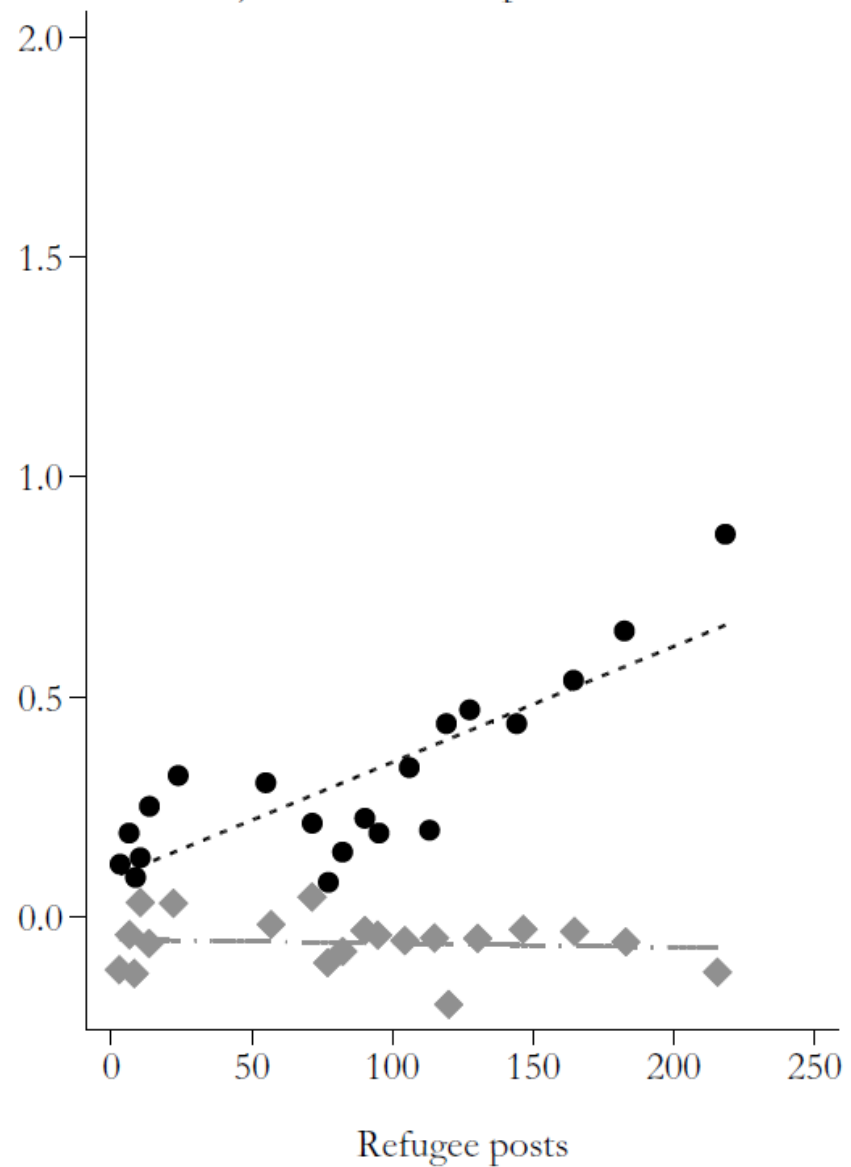
- **Outage** is an indicator variable for severe internet outages at the municipality-week level
- **Standard errors** are clustered at the municipality level (robust to alternatives)



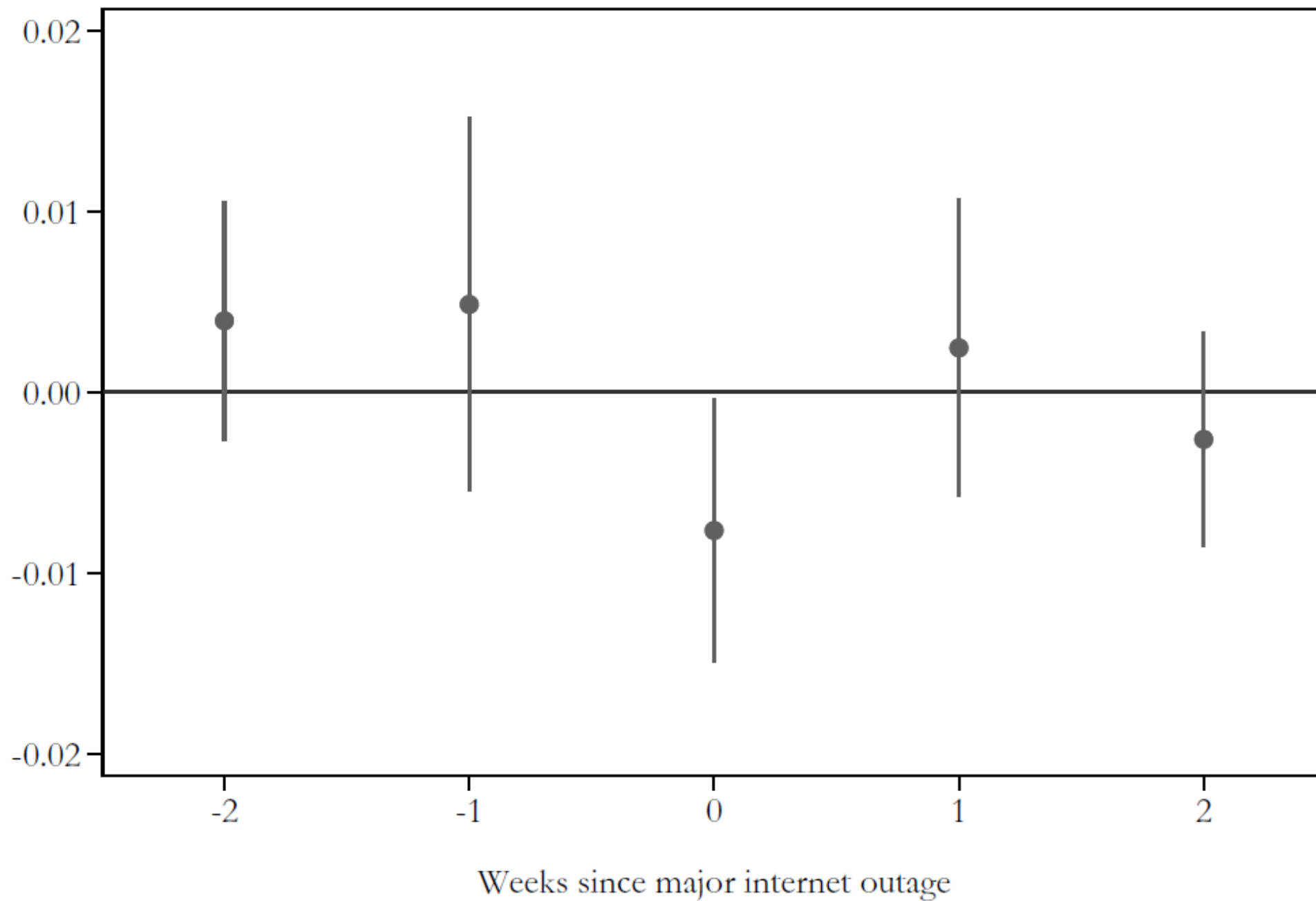
a) AfD users/Pop.  $\geq$  Median



b) AfD users/Pop.  $<$  Median



Change in probability of anti-refugee incident



## Main results

# Internet outages

	(1)	(2)	(3)	(4)	(5)	(6)
<b>Baseline Interaction</b>						
AfD users/Pop. × Refugee posts				0.024*** (0.009)	0.016** (0.008)	0.016** (0.008)
AfD users/Pop. × Posts × Outage				-0.181*** (0.058)	-0.184*** (0.058)	-0.172*** (0.057)
<b>Outage Interaction</b>						
Outage	-0.003*** (0.001)	-0.000 (0.001)	-0.003** (0.001)	-0.001 (0.002)	-0.002 (0.002)	-0.007 (0.008)
Refugee posts × Outage		-0.005*** (0.001)		-0.000 (0.002)	0.001 (0.002)	0.000 (0.002)
AfD users/Pop. × Outage			-2.685 (3.464)	4.441 (4.384)	4.455 (4.054)	4.391 (4.058)
<b>Internet Usage Interaction</b>						
Share broadband access × Outage						-0.000 (0.000)
Internet domains/Pop. × Outage						0.021* (0.012)
Mobile Broadband Speed × Outage						0.000 (0.000)
Observations	479,964	479,964	479,964	479,964	474,303	474,303
R-squared	0.082	0.082	0.082	0.082	0.084	0.084
Municipalities	4324	4324	4324	4324	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes
All other controls [22] × Posts					Yes	Yes

## Main results

# Internet outages: Robustness

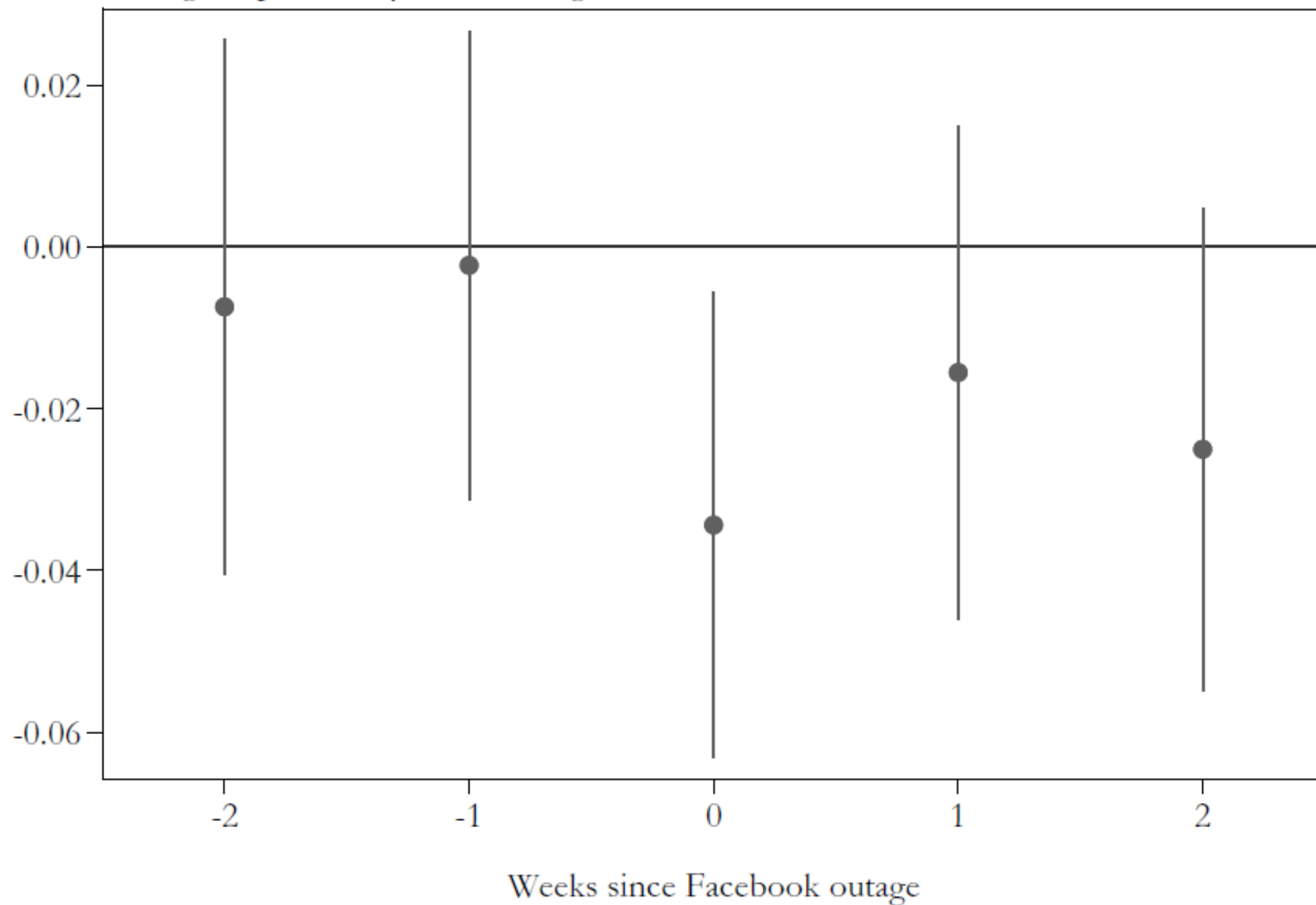
	(1)	(2)	(3)	(4)
	Official	Leave	Lagged	Google
	Reports	One Out	Posts	Sentiment
		Estimator		Measure
AfD users/Pop. $\times$ Refugee posts	0.009* (0.005)	0.057*** (0.021)	0.011 (0.008)	0.103*** (0.032)
AfD users/Pop. $\times$ Posts $\times$ Outage	-0.137*** (0.042)	-0.372*** (0.115)	-0.164*** (0.062)	-0.571*** (0.219)
Observations	474,303	474,303	470,030	474,303
R-squared	0.045	0.084	0.084	0.084
Municipalities	4273	4273	4273	4273
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
All controls [30] $\times$ Posts	Yes	Yes	Yes	Yes

## Main results

# Facebook outages

	(1)	(2)	(3)	(4)	(5)	(6)
<b>Baseline Interaction</b>						
AfD users/Pop. × Refugee posts			0.027***	0.027***	0.021**	0.021**
			(0.010)	(0.010)	(0.009)	(0.009)
AfD users/Pop. × Posts × Outage			-0.040*	-0.040*	-0.046**	-0.046**
			(0.021)	(0.021)	(0.022)	(0.022)
<b>Additional Outage Coefficients</b>						
Outage	-0.001***					
	(0.000)					
AfD users/Pop. × Outage		-2.222*	1.164	1.164	1.367	3.230
		(1.273)	(1.833)	(1.833)	(1.862)	(1.969)
Observations	479,964	479,964	479,964	479,964	474,303	474,303
R-squared	0.079	0.082	0.082	0.082	0.084	0.084
Municipalities	4324	4324	4324	4324	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE		Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes
All other controls [22] × Posts					Yes	Yes
All controls [30] × Outages						Yes

Change in probability of anti-refugee incident





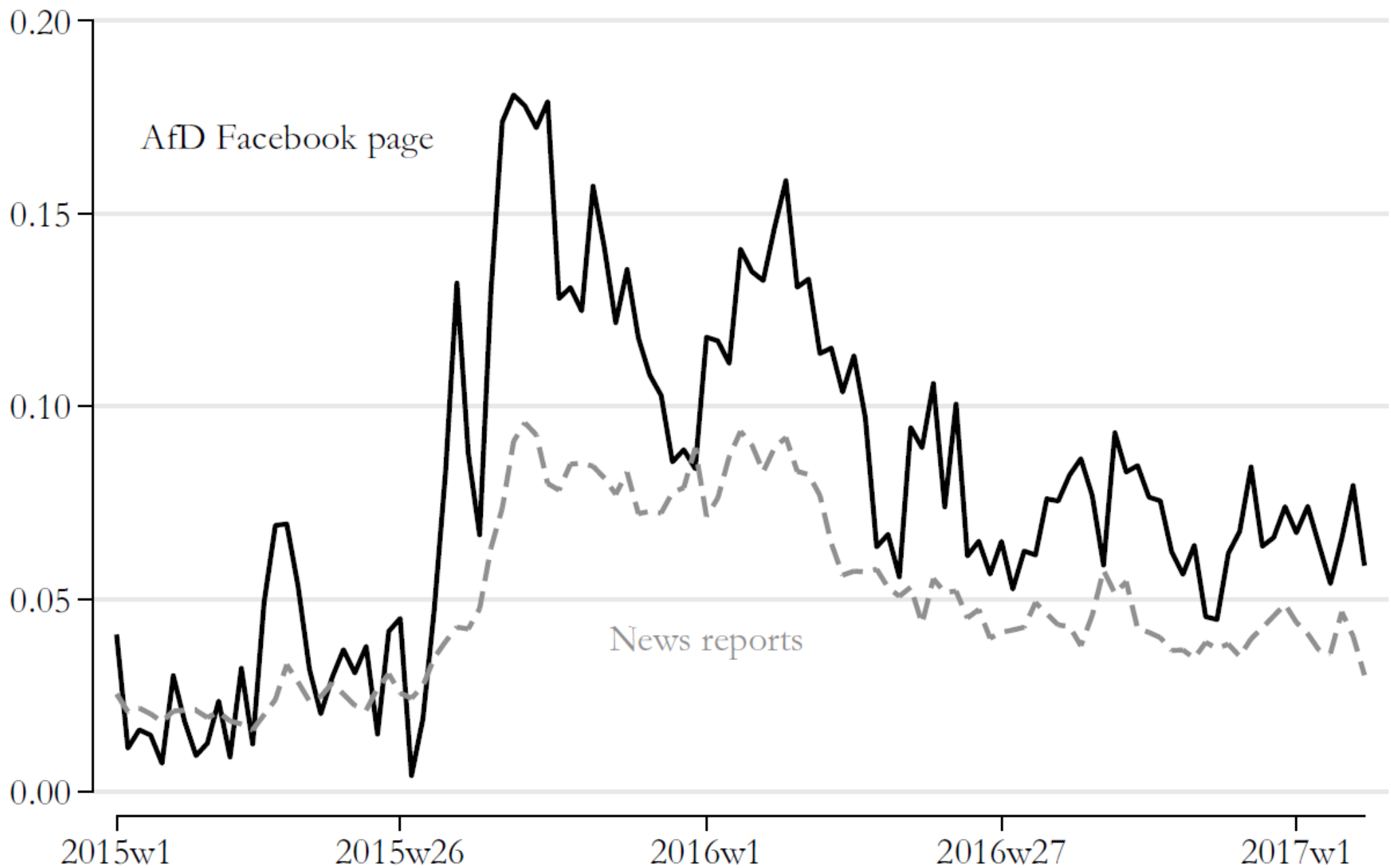
# Social vs. Traditional Media

# Major Differences

## 1) Share of refugee content

- Share of messaging on refugees is consistently higher on the AfD's Facebook page when compared to traditional media outlets
- Particularly when tensions are high, the share is nearly 100% higher

Share of Posts/News About Refugees



# Major Differences

## 1) Share of refugee content

- Share of messaging on refugees is consistently higher on the AfD's Facebook page when compared to traditional media outlets
- Particularly when tensions are high, the share is nearly 100% higher

## 2) Language more negative

- The language on the AfD's Facebook page is far more likely to mention crimes committed by refugees, corrupt elite, express distrust in the press
- Similar results using a machine learning approach

Rank	Word	Translation	Relativ prob.
<i>Panel A: Flücht (refugee)</i>			
1	Flüchtlingsenklaven	refugee enclave	780
2	Flüchtlingslüge	refugee lie	693
3	Flüchtlingsirrsinn	refugee insanity	650
4	Flüchtlingsmafia	refugee mafia	520
5	Flüchtlingsbefürworter	refugee supporter	520
<i>Panel B: Krimi (crime)</i>			
1	Regierungskriminalität	government crime	1300
2	Diskriminierungsgesetze	anti-discrimination laws	520
3	Schwerstkriminellen	dangerous criminals	260
4	Fluechtlingskriminalität	refugee crimes	260
5	Kriminalittssteigerung	increase in crime	260
<i>Panel C: Presse (media)</i>			
1	Freie Presse	free press	390
2	Propagandapresse	propaganda press	260
3	Presseempfang	press meeting	260
4	Meinungspresse	opinionated media	260
5	Nazipresse	nazi media	260
<i>Panel D: Volk (people)</i>			
1	Volksbetrug	betrayal of the people	1715
2	volksfeindlich	hostile to the people	780
3	volksverdummenden	brainwashing the people	520
4	Volksverhetzungsparagrafen	law against incitement	520
5	Volksprotesten	protest by the people	260
<i>Panel E: Verrat (betrayal)</i>			
1	Volksverrats	betrayal of the people	130
2	Vaterlandsverrat	betrayal of the fatherland	43
3	Volksverrat	betrayal of the people	43
4	Hochverrat	high treason	36
5	verratenen	betrayed	32



# Major Differences

## 1) Share of refugee content

- Share of messaging on refugees is consistently higher on the AfD's Facebook page when compared to traditional media outlets
- Particularly when tensions are high, the share is nearly 100% higher

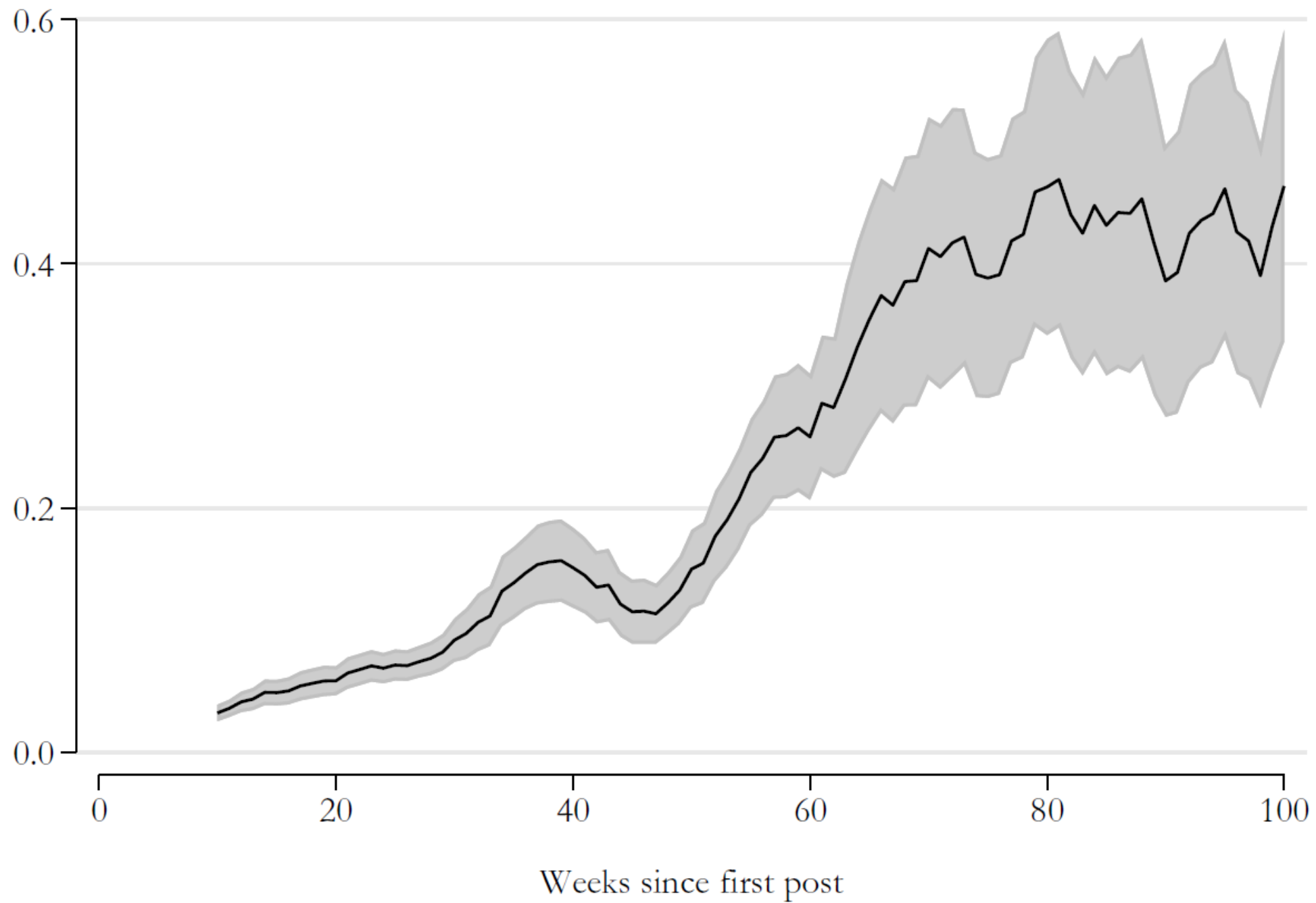
## 2) Language more negative

- The language on the AfD's Facebook page is far more likely to mention crimes committed by refugees, corrupt elite, express distrust in the press
- Similar results using a machine learning approach

## 3) Behavioural changes over time

- The longer people are active on the AfD's Facebook page the more frequently they post about refugees

Number of refugee posts





# Major Differences

## 1) Share of refugee content

- Share of messaging on refugees is consistently higher on the AfD's Facebook page when compared to traditional media outlets
- Particularly when tensions are high the share is nearly 100% higher

## 2) Language more negative

- The language on the AfD's Facebook page is far more likely to mention crimes committed by refugees, corrupt elite, express distrust in the press
- Similar results using a machine learning approach

## 3) Behavioural changes over time

- The longer people are active on the AfD's Facebook page the more frequently they post about refugees

Mechanism

## Mechanism

# Potential Mechanisms

## 1) Information Exchange

- People learn about locations of refugees on social media

## 2) Persuasion

- People's opinion about refugees is changed by social media

## 3) Coordination

- Attacks against refugees are coordinated on social media

## 4) Local Spillovers

- Social media inspire copy-cat attacks

## Mechanism

# Mechanism I: Information Persuasion

## Information Exchange

- Search for location information in posts (e.g. zip codes, street and town names)
- Location information only used in references to politicians and crimes *by*, not against refugees

## Persuasion

- High-frequency nature of variation makes persuasion an unlikely channel
- People's attitudes do not change on a week-to-week basis
- In particular, people do not become less xenophobic in a week that an internet outage cuts them off from Facebook

## Mechanism

# Mechanism II: Coordination

- The social media effect is stronger for incidents with multiple perpetrators

	(1) Known perp. sample	(2) 1 perp.	(3) <4 perp.	(4) ≥4 perp.
AfD users/Pop. × Refugee posts	0.010** (0.005)	0.003 (0.002)	0.004 (0.003)	0.007** (0.003)
Observations	479,964	479,964	479,964	479,964
R-squared	0.081	0.037	0.046	0.055
Municipalities	4,324	4,324	4,324	4,324
Share of attacks	1	0.245	0.494	0.534
Mean of DV	0.002	0.000	0.001	0.001
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes

## Mechanism

# Mechanism II: Local Spillover

- Attacks in neighboring municipalities increase likelihood of incidents

	(1)	(2)	(3)
AfD users/Pop. × Refugee posts	0.024*** (0.009)	0.022*** (0.008)	0.016* (0.008)
Attack in neighboring municipality	0.004*** (0.001)	-0.000 (0.002)	0.004** (0.002)
Attack in neighboring municipality × Posts		0.000 (0.001)	-0.004** (0.002)
Attack in neighboring municipality × AfD users/Pop.		13.765*** (4.782)	1.610 (4.914)
Attack in neighboring municipality × AfD users/Pop. × Posts			0.121** (0.052)
Observations	479,964	479,964	479,964
R-squared	0.082	0.082	0.083
Municipalities	4324	4324	4324
Municipality FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes

## Mechanism

# Potential Mechanism



### 1) Information Exchange

- People learn about locations of refugees on social media



### 2) Persuasion

- People's opinion about refugees is changed by social media



### 3) Coordination

- Attacks against refugees are coordinated on social media



### 4) Local Spillovers

- Social media inspire copy-cat attacks

Conclusion



## Conclusion

# Our findings

1. Empirical evidence for relationship between social media and hate crimes
2. Novel Facebook data allow construction of anti-refugee sentiment and social media measures
3. Facebook and internet outages provide exogenous variation in Facebook exposure and we show that the relationship between social media usage and hate crimes disappears during outages
4. Plausible mechanism driven by coordination of attacks
5. Limitation: Empirical strategy focuses only on short term shifts
6. In follow up research, "From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment", we provide evidence for mid- to long-term effects