

文字探勘與訊號理論：
以微信中美貿易戰資料為例

陳至潔



國立中山大學政治學研究所

Institute of Political Science | National Sun Yat-sen University

簡介

- 微信資料特徵
- 訊號理論 (Signaling theory)
- 威權中國的外交訊號
- 案例比較：中美貿易戰 (05/2019)

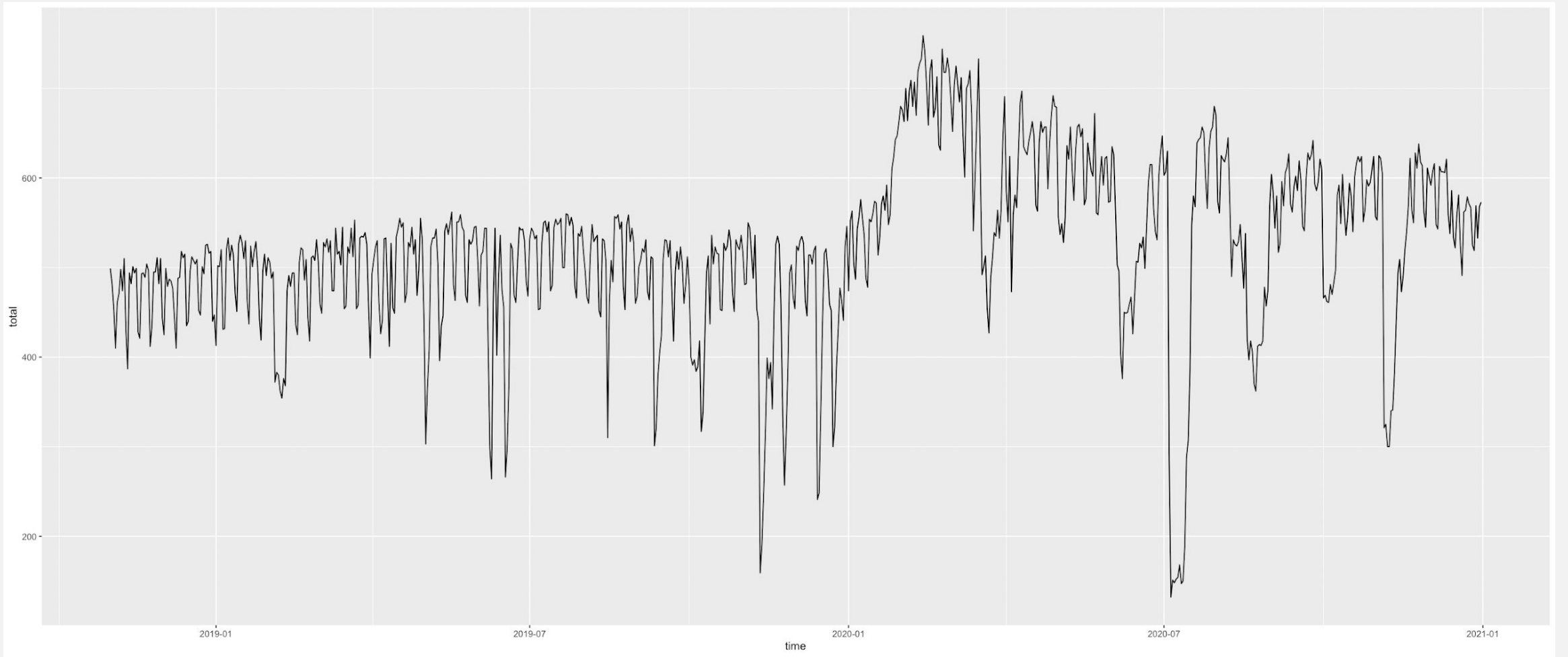
微信資料特徵

微信數據主要特徵

- 收錄時間：Nov. 1, 2018 ~ Dec. 31, 2020
- 收錄對象：
 - 根據『新榜』2018年最有影響力的100個微信公眾號
 - 去除內容農場帳號與非時政類帳號
 - 剩下47個公眾號
- 收錄文章量：412,339篇
 - 官媒文章數：306,234 (73.4%)
- 文章平均字元數：1,536
- 遭刪除文章數：10,137 (2.5%)
 - 非官媒遭刪除文章數：9,914 (97.8%)

time	account	author	title	content	official	license	forprofit	media	ensor	url
2018-11-01	讲武堂	空军世界加特林	美国飞天神器竟有一颗俄罗斯的心！想戒掉却是“忘记你我...	声明：除《踢馆》外，讲武堂所刊登文章均为授权转载，目...	0	0	1	5	0	http://
2018-11-01	铁血军事	金灿荣	从当前国际形势和贸易形势，看懂中国周边外交关系	本文经授权转载自微信公众号：政委灿荣（ID：jincanrong...	0	0	1	16	0	http://
2018-11-01	人民网	原创	毕业生获北京户口一年后离职，法院判其赔单位31万	非京籍毕业生梁某，通过应聘的学校获得了北京市户口，但...	1	1	1	2	0	http://
2018-11-01	界面新闻	原创	涪陵榨菜不愁卖，7个单品提价约10%	图片来源：视觉中国因为卖得好而被视作“消费降级”信号的...	1	1	1	2	0	http://
2018-11-01	瞭望智库	陆克文	如果认定中国是对手，美国该考虑这十个问题	今年11月，我们将迎来第一次世界大战结束100周年的纪念...	1	0	0	4	0	http://
2018-11-01	澎湃新闻	原创	彩票系统反腐：福彩中心已有5名重磅官员落马	澎湃新闻记者 蒋子文 岳怀让10月31日，中央纪委国家监委...	1	1	0	5	0	http://
2018-11-01	环球网	原创	“日本人在世界杯时的素质都哪去了”？	10月31日，万圣节前夕，日本东京涩谷每年都会举行盛大...	1	1	1	12	0	http://
2018-11-01	环球网	原创	“防弹少年团”反日？日韩又吵起来了！	10月24日下午，在首尔奥林匹克公园奥林匹克大厅，防弹...	1	1	1	7	0	http://
2018-11-01	界面新闻	李媚玲	FF美国正与不同投资者接洽，能否找到下一个白衣骑士？	“每天（高层）都在跟不同背景的投资人开会。”美国当地时...	1	1	1	2	0	http://
2018-11-01	人民日报	薛笙	【夜读】天赋，就是你看不见的努力	1 我的书法老师小蔡，是个90后。第一次在周末的成人...	1	1	0	4	0	http://
2018-11-01	观察者网	原创	学校这是什么令人窒息的操作。。。	NA	0	1	1	34	0	http://
2018-11-01	中国搜索	原创	一家5口先后3人患癌，儿子已不幸离世！癌症难道传染吗？	编者按：陕西一家五口人，四年之内先后有三人患癌，一人...	1	1	0	5	0	http://
2018-11-01	中国青年报	原创	热点 西昌发生5.1级地震，主持人淡定安抚同事：不要慌...	昨天一则地震消息引起很多人关注，据中国地震台网正式测...	1	1	0	23	0	http://
2018-11-01	澎湃新闻	原创	让大学生最慌的事发生了！学校给家长寄成绩单还包邮	澎湃特约评论员 胡乐乐深圳大学上大学“最害怕的事”发生了...	1	1	0	5	0	http://
2018-11-01	中国经济网	央视财经	整容大国如今竟流行“反整容”？这个视频火了...	韩国的美容整形产业世界闻名，但据俄罗斯第一频道报道，...	1	1	0	10	0	http://

微信數據欄位



訊號理論

訊號理論

- 一件公共政策的訊息 (signal) 由一政治行為者 (傳訊者) 傳送並由另一方 (收訊者) 所觀察而得之
- **間接**顯示傳訊方對施政的決心 (resolve)
- 決心是對方無法得知的資訊 (private information)
- 而且觀察者無法得知對手的訊號是否屬於欺敵 (deception) 行為
- 時常發生在外交緊張/軍事危機時刻, 具時間壓力

- 觀察者尋求 **可信** (credible) — 能顯示 **政府意圖** — 的外交訊號

訊號理論

- 昂貴訊息論 (costly signal) - 高昂成本的訊號較為可信
 - 將訊號的可信程度與訊號的成本掛鉤
- 沈沒成本 (sunk cost) 邏輯 - 訊號發送方所承擔、不可逆轉的施政成本
 - 對盟國安全承諾、海外駐軍、經濟援助
 - 直接訊號

訊號理論

- **自縛手腳 (tying hands) 邏輯** – 執政者所面對的**國內受眾成本 (domestic audience costs)**
 - 政府在對外談判過程中從強硬立場退縮而失去國內輿論支持或影響正當性
 - 政府對外政策前後不一致[先強後弱]而遭受的政治風險
 - 政府藉由**容忍國內輿論**來表明其承受國內壓力，從而展示了政府表現強硬外交立場的動機
 - 對國內輿論壓力的敏感程度成為一個政府對外談判的有力工具
 - 間接訊號
 - 政府向其國際談判對手傳遞了一個**代價高昂、因此是可信 (costly, hence credible)**的政治訊號，表明它在對外談判中**堅持立場的決心**。

威權中國的外交訊號

威權中國的外交訊號

- 不透明、無外部競爭的政治
- 受控管、被操弄的媒體
- J. C. Weiss 將中國民間的反外民族主義示威視為可信的外交訊號
- 據此推斷出威權政府在國內的脆弱程度和其外交強硬程度
- 藉由在外交衝突中容忍**反外民族主義**示威，威權政府領導者發出了一個昂貴的外交訊號，表明其對國內輿論的敏感程度，從而可信地展示其在對外談判中保持強硬立場的決心。
- 基於『自縛手腳』邏輯的昂貴訊號

威權中國的外交訊號

- 設計一套包含sunk-cost & tying-hands 邏輯的訊號判斷標準，能鑑別中共外交訊號的信息操縱模式和其內容特徵
- 可以從中國外交訊號的模式中推斷出中共政權領導層**威脅感知的來源和強度**
- 呈現威權外交訊號的文本：官方高層講話/聲明、官方媒體報導、社群媒體（微信/微博）發文文本與發文模式
- 選擇社群媒體（微信）發文，因其能同時呈現『沈沒成本邏輯』與『自縛手腳邏輯』的訊號成本

威權中國的外交訊號：沈沒成本邏輯

- 按照沉沒成本邏輯：透過測量微信帳號的政治發文的內容相似程度而得知在微信平台上發布外交訊號的成本。
- 微信政治發文的內容相似程度越高，就越需要政府在微信媒體採編新聞的過程中進行強有力的干預。
 - 微信發文內容的相似程度是政府將大量政治與行政資源用於策略性傳播的結果
- 政府在社群媒體平台上強調(priming)一個外交訊號的成本，可以透過檢驗微信媒體協同發佈內容高度相似的發文的時間點(timing)而得知。
- 相似性(similarity)和同步性(synchronicity)是一個昂貴的政治訊號的關鍵指標，這個訊號傳達了北京在外交緊張時期的決心

威權中國的外交訊號：自縛手腳邏輯

- 按照自縛手腳邏輯：在微信上傳散外交訊號的成本可以理解為國內受眾成本。
- 在中國，國內受眾成本可以通過官方立場與非官方意見在外交爭端上的情感差異來衡量。
- 國內受眾成本高，意謂著非官方微信媒體的發文在討論外國對手時的語氣比官方媒體更為負面或更加批判。
- 在外交爭端期間，若是非官方微信媒體傳播反外的民族主義言論，會有利於中共政權發出昂貴的對外訊號，向外部觀察者表明其對國內輿論壓力的敏感性(即脆弱性)，從而可信地表明在國際談判中立場堅定的決心。

威權中國的外交訊號類型

信息協同成本		高協同成本	低協同成本
國內受眾成本	高受眾成本	第一類訊號: 嚇阻 (Deterrence)	第三類訊號: 虛張聲勢 (Bluffing)
國內受眾成本	高受眾成本	<ul style="list-style-type: none"> • 高程度外部與國內威脅感知 • 展現高度不讓步的決心 	<ul style="list-style-type: none"> • 高程度國內威脅感知 • 展現無實質作為的決心
國內受眾成本	低受眾成本	<ul style="list-style-type: none"> • 高程度外部威脅感知 • 展現對外交協議的意願 	<ul style="list-style-type: none"> • 低程度外部與國內威脅感知 • 展現對外交談判的低度意願

案例比較：中美貿易戰

中美貿易戰:2019年5月的激烈化

- May 6: 川普宣布對價值2000億美元的中國輸美產品加徵10% ~ 25%的關稅, 作為對中方反悔談判結果的報復
- May 16: 川普宣布制裁華為與70家華為協力廠商, 將它們列入禁止交易名單
- 研究問題: 中方對兩個事件的反應是否相同? 中方對兩個事件是否發展出相似的威脅感知?

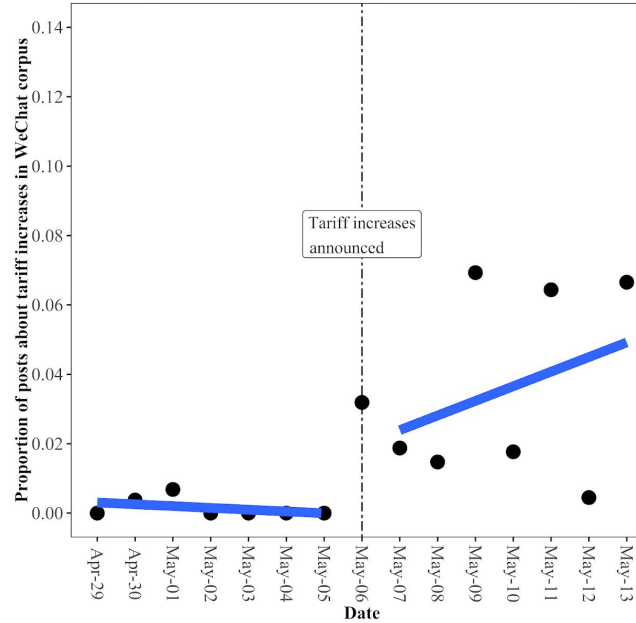
數據收集

- 關稅事件：收集 04/29 ~ 5/13 期間主題為關稅的微信發文
 - 關鍵詞：关税, 加征, 磋商, 贸易摩擦, 贸易战, 中美贸易战
 - 7060篇微信發文中, 148篇以關稅事件為主題(2.1%)
- 華為事件：收集 05/09 ~ 05/23期間主題為華為的微信發文
 - 關鍵詞：华为 任正非, 禁令, 芯片, 断供, 海思, 操作系统, 半导体.
 - 7670篇微信發文中, 348篇以華為為主題(4.5%)

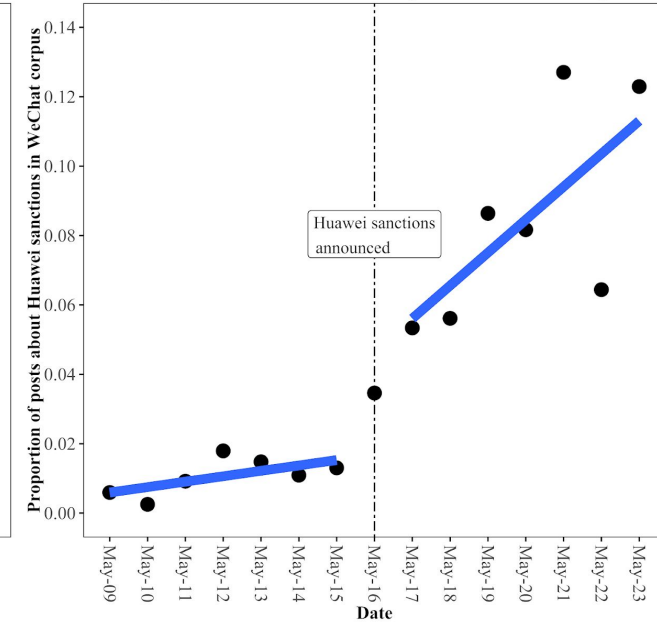
關稅 VS. 華為

- 關稅報導斜率:
 - -0.0005 -> 0.004
- 華為報導斜率:
 - 0.0017 -> 0.0095
- 關稅事後報導佔比: 3.7%
- 華為事後報導佔比: 8.1%

6.1a Proportion of tariff-related posts in WeChat corpus



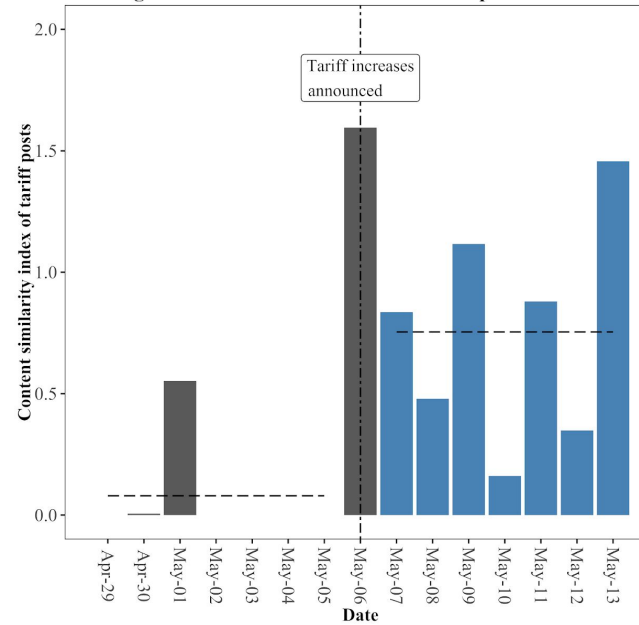
6.1b Proportion of Huawei-related posts in WeChat corpus



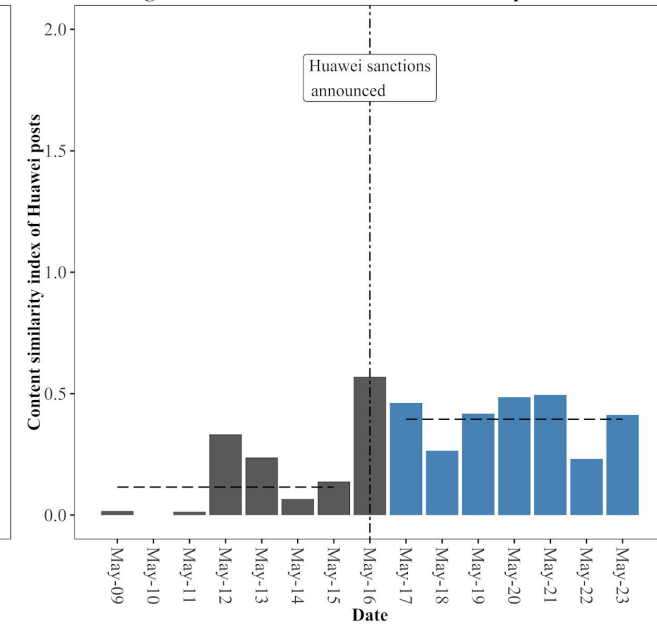
關稅VS.華為

- 關稅: 內容相似程度高
- (0.734)
- 華為: 內容相似程度低
- (0.41)

6.2a Degree of content coordination of tariff posts

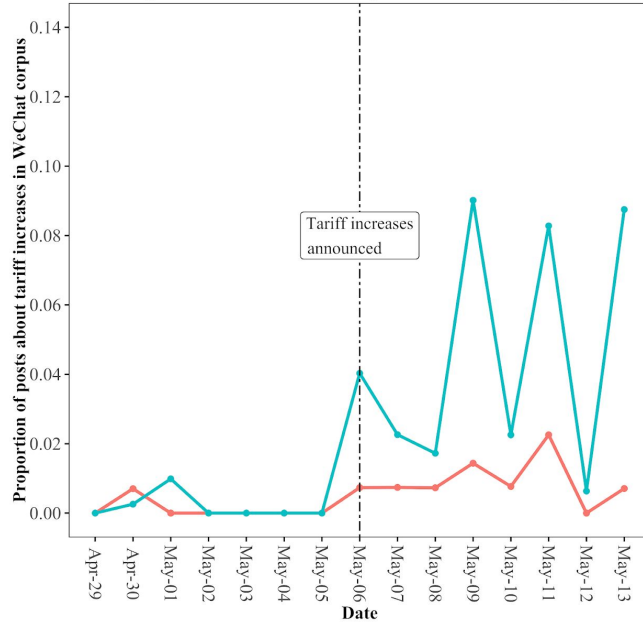


6.2b Degree of content coordination of Huawei posts

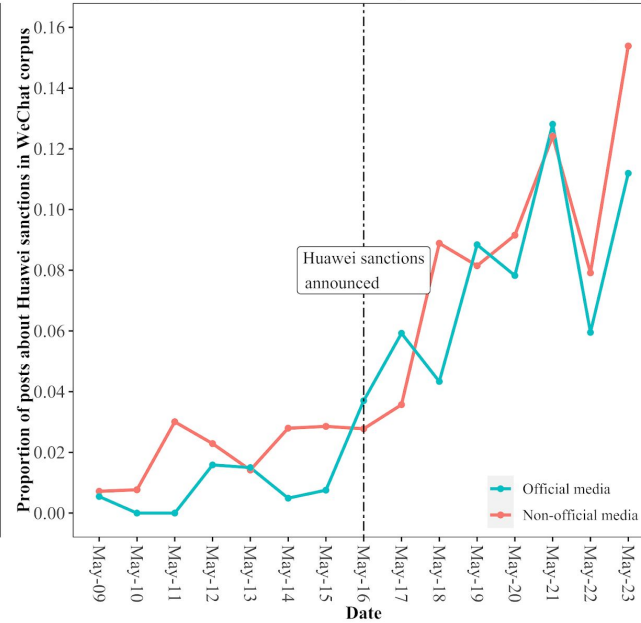


關稅VS.華為

6.3a Proportion of tariff-related posts in WeChat corpus

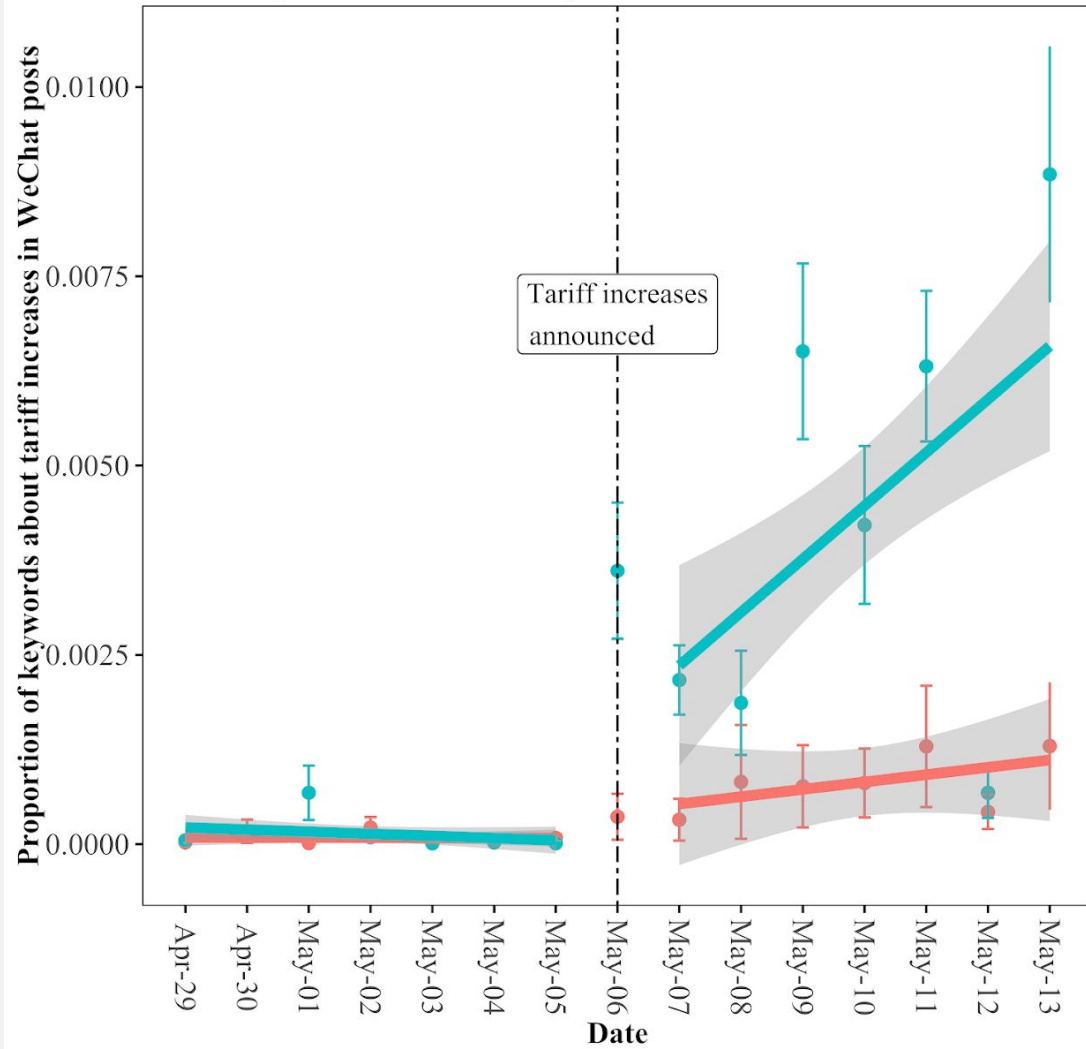


6.3b Proportion of Huawei-related posts in WeChat corpus

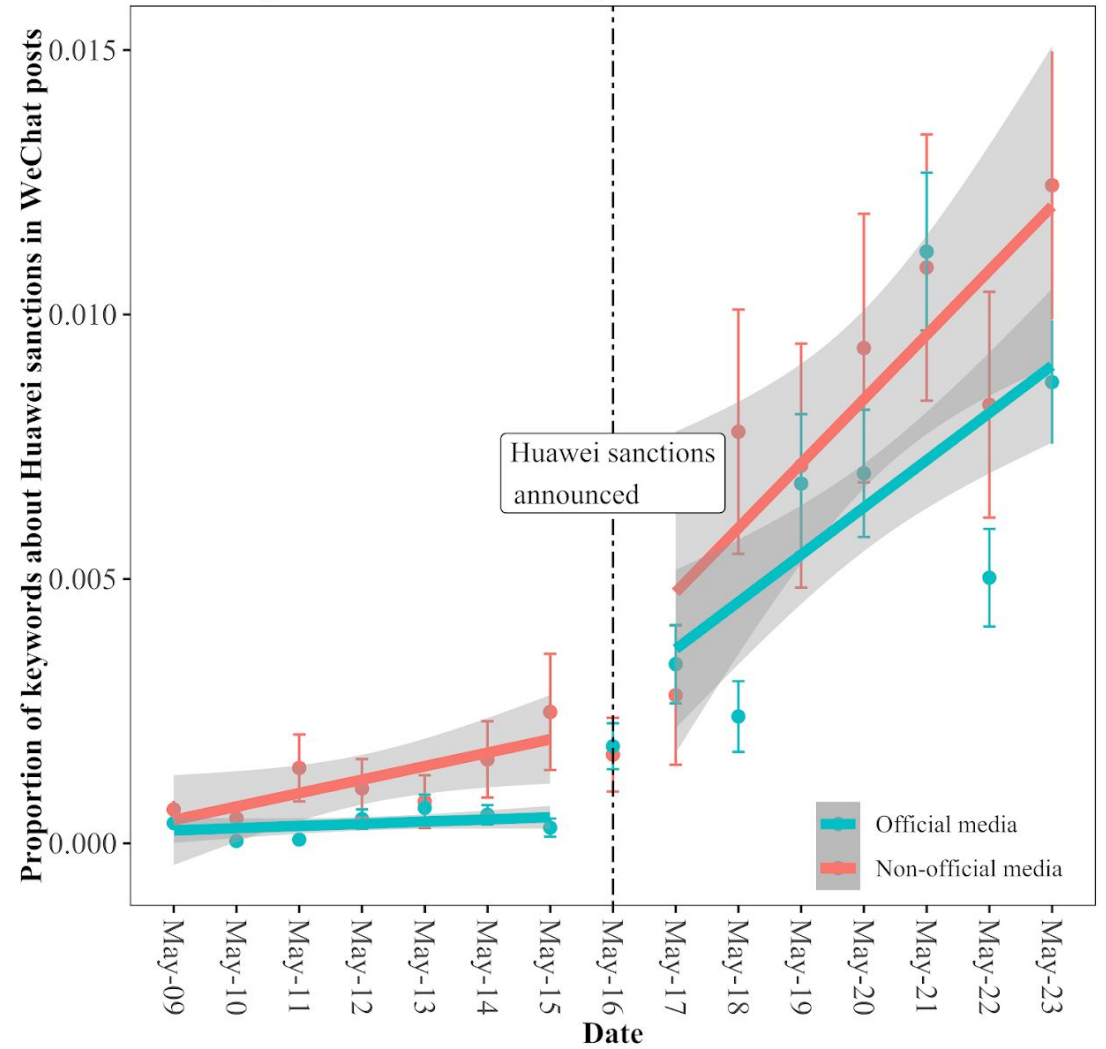


- 關稅: 官方媒體報導比例多於非官方媒體
- 華為: 官方與非官方媒體報導比例幾乎一樣

6.4a Proportion of tariff keywords in WeChat content



6.4b Proportion of Huawei keywords in WeChat content



情緒分析的 統計結果

ZERO-INFLATED NEGATIVE
BINOMIAL MODEL

	Response variables:					
	Tariffs Positive tone	Tariffs Negative tone	Tariffs Critical tone	Huawei Positive tone	Huawei Negative tone	Huawei Critical tone
Official	0.06 (0.11)	0.19* (0.10)	0.16* (0.10)	0.02 (0.07)	-0.07 (0.06)	-0.06 (0.07)
License	-0.01 (0.06)	-0.02 (0.06)	-0.04 (0.06)	0.01 (0.05)	0.03 (0.04)	0.00 (0.05)
For_profit	0.02 (0.06)	0.08 (0.07)	0.09 (0.06)	0.03 (0.06)	-0.02 (0.04)	-0.01 (0.05)
USA	-0.02 (0.03)	-0.03* (0.02)	-0.02 (0.02)	-0.09*** (0.01)	0.00 (0.01)	-0.02 (0.01)
PRC	0.01 (0.03)	-0.01 (0.02)	-0.02 (0.02)	-0.01 (0.02)	0.04* (0.02)	0.01 (0.03)
Bilateral	-0.00 (0.04)	0.04 (0.03)	0.07** (0.03)	-0.08 (0.06)	-0.14 (0.12)	0.00 (0.12)
log(Word_num)	-0.01 (0.03)	-0.04* (0.03)	-0.05* (0.03)	0.01 (0.01)	0.00 (0.01)	-0.04** (0.01)
Date	0.00 (0.01)	0.03*** (0.01)	0.03*** (0.01)	-0.02*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Variance: account	0.00	0.01	0.00	0.01	0.00	0.00
Dispersion:	5.63	51.41	151.41	4.32	5.75	5.81
Dispersion: SD	0.81	25.42	212.61	0.17	0.34	0.42
Zero inflation:	0.68	0.69	0.73	0.58	0.75	0.81
Zero inflation: SD	0.01	0.01	0.01	0.00	0.00	0.00
AIC	6224.52	6187.56	5587.90	43541.40	29279.00	23296.40
BIC	6293.62	6256.66	5657.00	43630.55	29368.15	23385.55
Log Likelihood	-3100.26	-3081.78	-2781.95	-21758.70	-14627.50	-11636.20
Num. obs.	2341	2341	2341	12448	12448	12448
Num. groups:	43	43	43	45	45	45

Standard errors in parentheses.

***p < 0.01; **p < 0.05; *p < 0.1

綜合討論

- 中國政府投注於關稅報導的沈沒成本 > 投注於華為報導的沈沒成本
 - 網管的成本, 網宣的成本
 - 對關稅報導的限制 > 對華為報導的限制
 - 官媒對關稅報導的比例 > 官媒對華為報導的比例
- 中國政府對關稅報導的國內受眾成本 < 對華為報導的國內受眾成本
 - 在關稅報導中, 官媒比非官媒更加負面與批判
 - 在華為報導中, 官媒與非官媒的負面批判情緒差異不顯著

綜合討論

- 中國政府在關稅案例傳送了第二類外交訊號(再保證)
 - 對外部威脅的高度感知
 - 對內部民族主義聲量的壓制
 - 對外交談判與達成協議的興趣
- 中國政府在華為案例傳送了第四類外交訊號(迴避)
 - 對外部與國內的威脅感知都不高
 - 並未壓制國內民族主義的聲量
 - 對外交談判與達成協議較無興趣

感謝聆聽

Q & A