Replication Package for: Platform Attention and the Twitter Takeover

Course: Platform Economics
Date: April 2025
Author: Elya Madambekova Uysal
Institution: Jönköping University, Master of Applied Economics and Data Analysis

1. Overview

This replication package examines search interest for six major social media platforms before and after the 2022 Twitter acquisition. The analysis is conducted using a Difference-in-Differences (DiD) framework with standardized Google Trends data from 2019 to 2024.

2. Data Collection and Sources

Source: Google Trends (https://trends.google.com/)
Platforms: Twitter, Facebook, TikTok, Instagram, YouTube, LinkedIn
Timeframe: 2019-01-01 to 2024-12-31
Frequency: Weekly search interest (0–100 scale)

Why `pytrends` and VPN?

Initially, data was downloaded manually using Google Trends' export feature. However, there are:
- Truncated timeframes,
- Inconsistent row counts across platforms,
- Partial periods with missing or zero values.

Therefore, programmatic scraping was used via the `pytrends` library and used a VPN to bypass regional rate limits and ensure full weekly data access across all years.

3. Methodology

Preprocessing:
- Data collected year-by-year per platform with retry logic
- Cleaned and merged into long format (columns: `date`, `Platform`, `SearchVolume`)
- Standardized using z-scores within each platform

Model:
- Difference-in-Differences (DiD) regression with controls:
  - `Treatment`, `Post`, `Treatment × Post`
  - Time trends and interaction terms
  - Month and platform fixed effects

- Additional event study estimates dynamic treatment effects ±12 weeks

4. Summary of Results

Event Study:
- Twitter shows a short-term dip after April 14, 2022, of approximately. **–0.26 SD**, which is **not statistically significant** (p > 0.05)
- The effect disappears within 4–5 weeks
- Other platforms show stable or rising trends, especially TikTok

Barplot:
- TikTok and Instagram experience growth post-event
- Twitter shows minimal change in average interest
- LinkedIn and Facebook remain largely stable

5. File Structure

File/Folder	Description
`scraper.py`	Downloads weekly search volume for all platforms using `pytrends`
`prepare_data.py`	Normalizes and reshapes raw data for DiD modeling.
`main.py`	Runs the DiD model and generates output plots and a regression table.
`data/`	Folder containing raw and cleaned `.csv` files
`output/plots/`	Event study and barplot figures
`output/tables/`	Regression output file
`requirements.txt`	Python library dependencies

6. Software and Environment

Package	Version
Python	3.9
pandas	1.5.x
matplotlib	3.6.x
seaborn	0.12.x
statsmodels	0.13.x
pytrends	4.9.x

Tested on: **Windows 11** using **Spyder IDE**

7. Integrity Statement

All data was obtained from publicly accessible sources. No simulated or fabricated data was used. VPN usage was only intended to ensure full-year completeness and did not involve unauthorized access or scraping beyond Google's public interface.

8. References

1. Google Trends: https://trends.google.com
2. `pytrends` – Unofficial Google Trends API for Python: https://github.com/GeneralMills/pytrends
3. Wooldridge, J. M. (2010). *Econometric Analysis of Cross-Section and Panel Data*. MIT Press.
4. Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics*. Princeton University Press.
5. Statsmodels documentation: https://www.statsmodels.org/
6. Jonathan Roth April 2023: "What's trending in Difference-in-Differences" https://www.jonathandroth.com/assets/files/DiD_Review_Paper.pdf

9. License

10. Contact

email: mael24mm@student.ju.se
For academic use only.