Fighting Organized Crime by Targeting their Revenue: Screening, Mafias and Public Funds

Data generation process

Main database: "data/data_funds.dta"

1. Our data on European Subsidies comes from the public database **Opencoesione (https://opencoesione.gov.it/en/opendata/#!progetti_section),** which records information on all awarded subsidies co-financed by the European Union.
    a. The database contains information at the subsidy level and includes data on the identity of the awarded company, characteristics of the project for which the award was assigned (time to complete it, type of project, amount funded), as well as the co-financing institutions (Ministry, Region, Municipality, ...).
    b. We use data from 2007 to 2015 (last year available at the time of writing the paper) including subsidies between 50,000 and 250,000 euros (the range with the highest density of subsidies around the 150,000 euros threshold). We also show robustness to extending the window of analyses below this group.
2. We use companies name and fiscal code to match this data with a proprietary database of Italian firm information, **AIDA (https://aida.bvdinfo.com/version-20211230/Home.serv?product=AidaNeo).**
    a. We are able to match 45% of our sample to this database. Note that, while we use the matched database for our main analyses, we also repeat the analyses and robustness tests on the full sample from Opencoesione.
    b. Aida allows us to obtain information on the firm (sector, time of operation, location) and on its owners (sex, age, place of birth).
3. We use region and city names to additionally merge information on **mafia presence**. In particular, we merge the following databases:
    a. Real estates and firms seized to mafias from the ANBSC - Italian Ministry of Interior (https://benisequestraticonfiscati.it/).

b. City councils dissolved due to mafia infiltration from the Ministry of Interior (https://www.interno.gov.it/it/notizie/enti-sciolti-mafia-presentata-relazione-ministro-dellinterno-sulle-attivita-commissioni).
c. Location of mafia victims collected from the NGO "Vittimemafia" (https://vittimemafia.it/).
    i. We define a city as having mafia presence if any of the indicators in points a,b,c scores one or above.
d. Provinces infiltrated by 'Ndrangheta as in Calderoni (2011).

4. We use region and city names to merge data on **corruption** and **government inefficiency** (speed of payment) at the city level in the years 2003-2013 from the Ministry of Interior.
    a. We define corruption as a dummy equal to one if any of the following crimes were committed in the municipality in the period 2003-2012, when the data is available: corruption, concussion, abuse of power, embezzlement.
    b. We define a measure of government inefficiency as a dummy equal to one if the speed of payments in a municipality is lower than median.
5. Finally, we merge data on **spatial coordinates** to map the data. We obtain shapefiles on spatial coordinates from the Italian census 2011 and transform them into .dta files.

Firm owners database: "data/data_owners.dta"

- We gather information on business owners from AIDA and match them to information on subsidies. In this version of the database, we use data at the business owner level rather than at the subsidy level.
- *Use*: analysis on figureheads (Table 3)

Open coesione database: "data/data_opencoesione.dta"

- Database from Opencoesione including all subsidies awarded, independently on whether they are or not matched to AIDA. The data is also matched with information on mafia presence.
- *Use*: robustness to the main analysis (Appendix B)

Database below 50,000 Euro: "data/data_below50000.dta"

- Database from Opencoesione including all subsidies awarded, including those below 50,000 Euros.
- *Use*: descriptive histogram (Figure 6, Appendix) and robustness to the main analysis (Figure 14, right panel).

Placebo database: "data/data_placebo.dta"

- Data on cities dissolved due to mafia infiltration in the past 5 years matched with data on subsidies that were previously merged to AIDA.
- *Use*: Placebo test in Appendix E, Figure 15, left panel

Placebo for agricultural funds: "data/data_agea.dta"

- Data on subsidies to agriculture, obtained by AGEA (*Agenzia per le Erogazioni in Agricoltura*).
- *Use*: Placebo test in Appendix E, Figure 15, right panel

Placebo for Opencoesione database: "data/data_placebo_opencoesione.dta"

- Data on cities dissolved due to mafia infiltration in the past 5 years matched with full data on subsidies (whether or not they were merged to AIDA).
- *Use*: Placebo test in Apppendix B, Figure 4, right panel.