

Exploratory analysis of accesses to support centers for gender-based violence in Apulia

Data

The dataset employed regards the counts of accesses to gender-based violence support centers in the Apulia region by residence municipality of the women victims of violence in 2021-2024. R codes to generate the dataset are in the R script posted here which this report is based on.

Here, we only take into account the violence reports which support centers actually take charge of, at the risk of underestimating the counts of gender-based violence cases. This choice is driven by the need of avoiding duplicated records, since e.g. it may happen that a support center redirects a victim to another support center.

In order to avoid singletons, i.e. municipalities with no neighbours, in the spatial structure of the dataset, the Tremiti Islands need to be removed from the list of municipalities included (0 accesses recorded so far).

Therefore, the municipality-level dataset in scope consists of 256 areas.

We can only take into account the accesses to support centers for which the origin municipality of victims is reported; therefore the total count of accesses in scope is 1477, 1516, 1822 and 1778 for the reference years respectively:

Here, we plot the log-access rate per residence municipality, i.e. the logarithm of the ratio between access counts and female population. Blank areas correspond to municipalities from which zero women accessed support centers (82 municipalities).

Covariates

Our target is explaining the number of accesses to support centers, y , defined at the municipality level, on the basis of a set of candidate known variables. y is modelled with simple Poisson regression.

We have at disposal a number of candidate explanatory variables, which include the distance of a municipality from the closest support center and a set of variables measuring social vulnerability under different dimensions; these latter covariates are provided by the ISTAT and are among the components of the Municipality Frailty Index (MFI). A more detailed description of MFI components is in this excel metadata file.

All covariates are quantitative variables and to ease model interpretation are scaled to have null mean and unit variance.

- TEP_th, i.e. the distance of each municipality from the closest municipality hosting a support center. Distance is measured by road travel time in minutes (acronym TEP stays for Tempo Effettivo di Percorrenza, i.e. Actual Travel Time). Since to the best of our knowledge the list of active support centers changed between 2022 and 2023, we employ the list of centers active until 2022 for 2021-2022 data, and the list of centers active in 2023 for 2023 data.
- AES, the distance from the closest infrastructural pole, always measured in travel time. Infrastructural poles are defined as municipalities or clusters of neighbouring municipalities provided with a certain endowment in health, education and transport infrastructure (more details can be found, e.g., here).
- MFI, i.e. the decile of municipality frailty index.

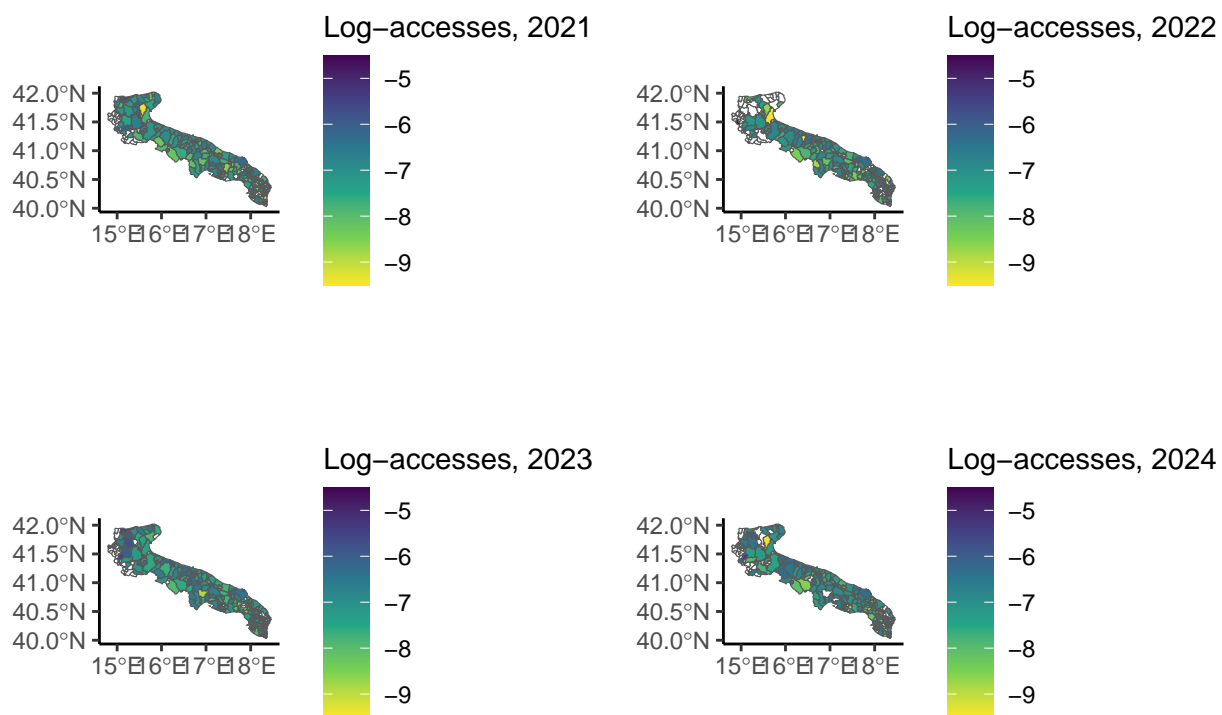


Figure 1: Log-access rate

- PDI, i.e. the dependency index, i.e. population either ≤ 20 or ≥ 65 years over population in $[20 - 64]$ years.
- ELL, i.e. the proportion of people aged $[25 - 54]$ with low education.
- ERR, i.e. employment rate among people aged $[20 - 64]$.
- PGR, i.e. population growth rate with respect to 2011.
- UIS, i.e. the ventile of the density of local units of industry and services (where density is defined as the ratio between the counts of industrial units and population).
- ELI, i.e. the ventile of employees in low productivity local units by sector for industry and services.

First, we visualise the correlations among these explanatory variables:

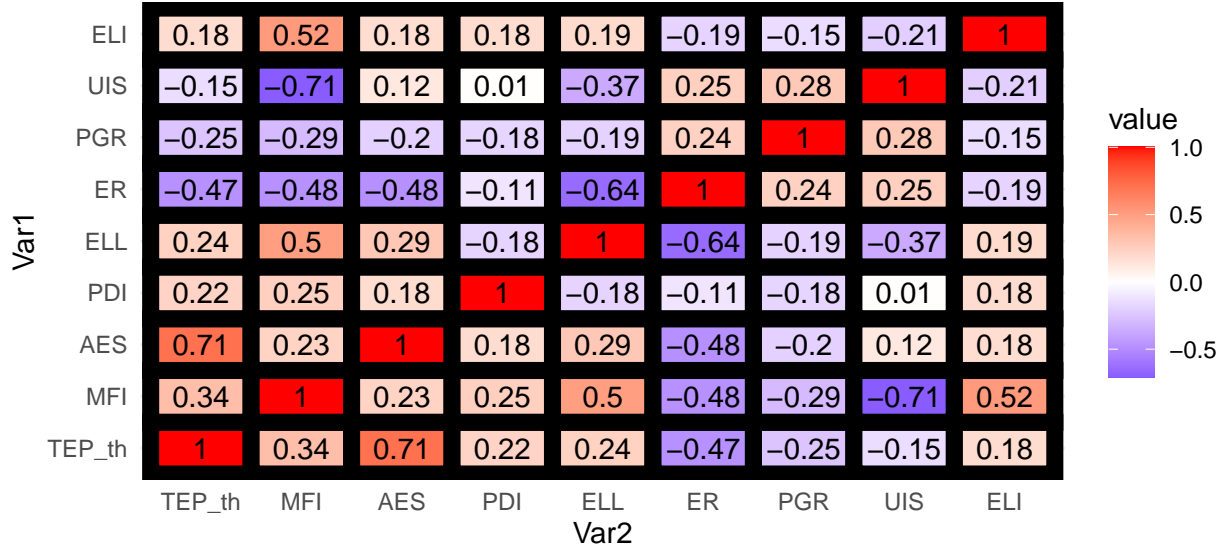


Figure 2: Correlations in explanatory variables

We see the correlation between the two distances is very high (0.72), and so is the correlation between the fragility index decile and the density of productive units.

In the first case, we drop the distance from the nearest infrastructural pole. In the latter we drop MFI, which is a combination of all covariates except for TEP_th, and is a weakly informative choice.

Nonspatial regression

We regress the counts of accesses y to support centers on the aforementioned explanatory variables. To interpret regression coefficients in an easier way, all covariates are scaled to zero mean and unit variance.

$$y_{it} \mid \eta_{it} \sim \text{Poisson}(E_{it} e^{\eta_{it}}) \quad \text{where} \quad \eta_{it} = X_{it}^{\top} \alpha \quad (1)$$

Where X are the covariates defined earlier, α are covariate effects, and E_{it} is the female population aged ≥ 15 in municipality i and year t .

To gain more insight on the role of all explanatory variables we show the posterior summaries of the regression model; the posterior distribution is approximated with the INLA (INLA?).

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Thu Aug 7 19:34:03 2025

Effect	Mean	Sd	0.025quant	0.975quant
Int_2021	-7.333	0.030	-7.392	-7.274
Int_2022	-7.305	0.030	-7.364	-7.247
Int_2023	-7.128	0.028	-7.183	-7.073
Int_2024	-7.152	0.028	-7.208	-7.097
TEP_th	-0.227	0.017	-0.260	-0.195
ELI	-0.088	0.015	-0.117	-0.059
PGR	0.017	0.020	-0.021	0.056
UIS	0.040	0.016	0.009	0.071
ELL	-0.110	0.021	-0.150	-0.070
PDI	-0.065	0.021	-0.106	-0.025
ER	-0.212	0.022	-0.256	-0.169

- **TEP_th_22:** The distance from the closest support center appears to play an important role. The easiest interpretation is that the physical distance represents a barrier to violence reporting. This is quite intuitive if we think of the material dynamics of reporting gender-based violence: one could reasonably expect violent men to prevent their partners to take a long rout and come out to report the violence suffered.
- **ELI:** The (ventile of the distribution of the) share of employees in low productivity economic units is a clear indicator of (relative) economic underdevelopment. The most naive interpretation would be that in underdeveloped areas reporting gender violence is somewhat harder than in developed ones.
- **PGR:** The association with population growth rate does not appear to be significantly different from zero.
- **UIS:** The (ventile of the distribution of the) density of production units does not appear to have a significantly $\neq 0$ association with VAW reporting.
- **ELL:** The association with the proportion of people with low educational level has negative sign and is high in absolute value. The interpretation seems quite easy: cultural development, in general, would encourage reporting violence.
- **PDI:** The association with population dependency index is negative, and its interpretation is ambiguous as well since the proportion of old or very young people can be read as a limiting factor both for VAW incidence and reporting.
- **ER:** The association with employment rate is very strong and bears negative sign for 2021 and 2023 data.

Spatial regression

Exploratory analysis of residuals We plot the log-residuals ε of the nonspatial regression models, defined as $\varepsilon := \ln y_{it} - \ln \hat{y}_{it}$ being \hat{y}_{it} the fitted value.

Residuals may exhibit spatial structure. To assess it, we employ the Moran and Geary tests. Since

Please notice that log-residuals only take finite values across the municipalities whose female citizens have reported at least one case of violence.

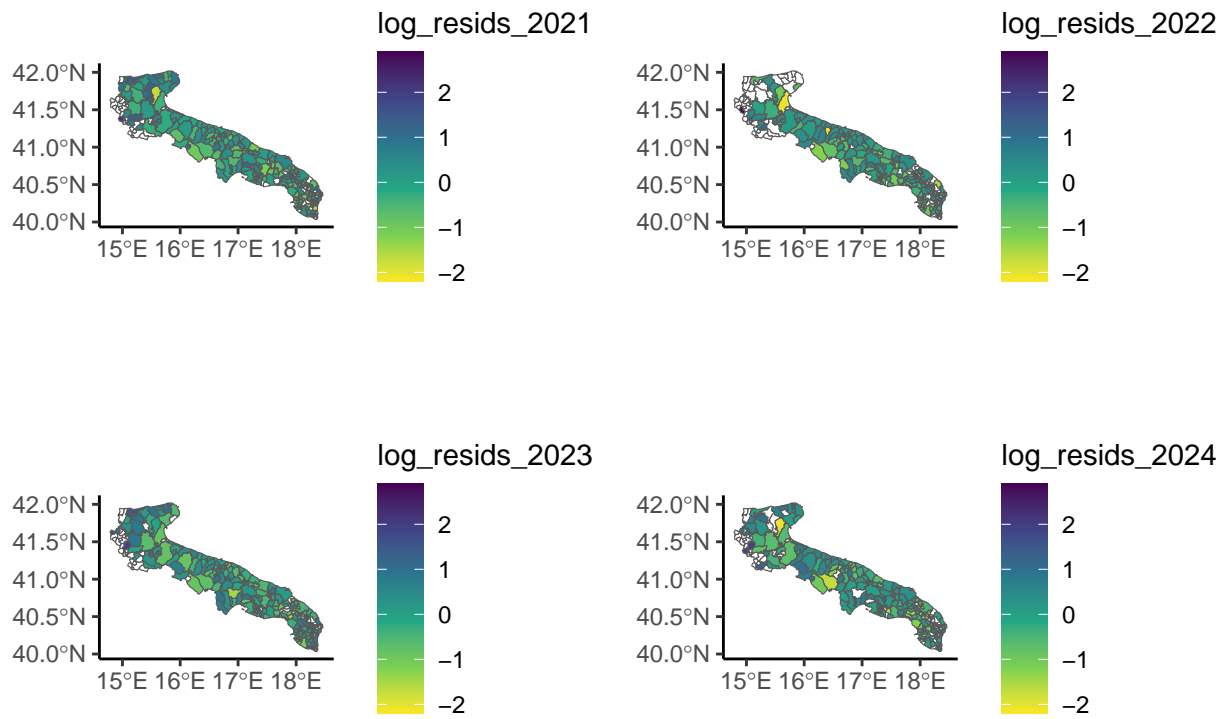


Figure 3: Log-residuals in GLM regression