SENTENCED WITHOUT EVIDENCE? AN HONEST DIFF-IN-DIFF EVALUATION OF ITALY'S ANTI-DEPOPULATION NATIONAL STRATEGY FOR INNER AREAS

Gabriele Dore

Abstract. This study investigates the medium-term impact of the National Strategy for Inner Areas (SNAI) on internal emigration dynamics, aiming to assess the policy effectiveness in retaining residents in targeted inner areas. Leveraging Italian Census data (2012-2023 period) and OpenCoesione's 2025 release of SNAI's geocoded project records, a Doubly Robust Difference-in-Differences design, complemented with partial identification, was used to provide robust causal estimates under incremental violations of the identifying assumption. Results suggest that the policy did not produce significant effects on municipalities treated between 2017 and 2019. However, even under modest violations of the parallel trend assumption, causal estimates become fragile and inconclusive. This fragility persists even after accounting for the confounding effect of the COVID-19 pandemic, signalling a structural difficulty in achieving identification via Difference-in-Differences with available data. This calls for caution in attributing conclusive demographic effects to the policy in both the short and medium term.

1. Introduction

In July 2025, the Italian Government formally acknowledged that a significant proportion of Italian inner areas is in a condition of irreversible demographic decline (PSNAI, 2025), foreseeing for these territories institutional support on a "*socially dignified pathway of chronic shrinkage and ageing*" as a Cohesion Policy objective (PSNAI, *p. 45*). Such acknowledgement reflects the inherent difficulty of reversing decline once it becomes entrenched (Copus *et al*., 2021). In fact, the phenomenon can not be understood solely as a contraction of the resident population, but rather as a self-perpetuating syndrome of decline where negative natural and migration balance intersects with processes of economic restructuring, geographic disadvantages, and austerity-driven disinvestments (Elshof *et al*., 2014; Copus *et al*., 2021; Newsham and Rowe, 2025). This syndrome erodes local tax bases, weakens essential public services, and undermines economic vitality, hampering territories' endogenous capacity to react (Copus *et al*., 2021).

The unprecedented recalibration of Italy's policy effort to counter depopulation follows nearly a decade of targeted action addressing territorial disparities. Through the National Strategy for Inner Areas (SNAI), Italy established a place-based policy to reverse negative demographic trends by improving service provision (education, healthcare, transport) and promoting local development (Barca *et al*., 2014; Lucatelli *et al*., 2022). Because migration is the most responsive component of population change in the short term (Newsham and Rowe, 2025), SNAI's effects are expected to emerge first through adjustments in internal mobility. Evidence shows that out-migration reacts strongly to both service availability and local economic opportunities: improved service access reduces pressures to leave, while enhanced employment prospects strengthen retention and attraction and can support fertility (Elshof *et al*., 2014; Sonzogno *et al*., 2022; Benassi *et al*., 2023). By acting simultaneously on these determinants of mobility, SNAI potentially creates a coherent pathway through which interventions can curb out-migration and increase the attractiveness of inner areas, laying the groundwork for longer-term demographic stabilisation.

Nevertheless, it remains a matter of debate whether the SNAI is producing the desired demographic outcomes. Empirical evidence on its impact on population dynamics is scarce. The only contribution in this respect is by Monturano *et al*. (2025), who found no evidence of an effect on net migration rates, but positive impacts on new business creation in the short term through a Difference-in-Difference (DiD) design. However, their analysis is based on a short-term evaluation and a sample of treated municipalities as of 2020, which limits the possibility of providing the evidence needed to inform potential policy reforms.

To contribute to the ongoing debate, this study investigates the medium-term effects of the SNAI on internal emigration from Italy's inner areas. The empirical strategy combines the doubly robust properties of Callaway and Sant'Anna (2021)'s estimator with Rambachan and Roth (2023)'s partial identification approach, allowing for causal inference under incrementally relaxed identifying conditions. Compared to the existing literature, this study leverages newly available SNAI-related intervention data extending to 2025, allowing for a larger treated sample and a longer observation window; it adopts an empirical approach suited for medium-term evaluation that accounts for confounding factors such as the COVID-19 pandemic; and it moves beyond aggregated migration indicators by focusing on demographically granular internal emigration flows. By isolating out-migration, this study avoids conflating retention and attraction effects, which may follow different dynamics and temporal patterns after policy interventions. This focus also allows us to provide precise evidence on the capacity of the strategy to retain the resident population.

2. Policy Background and Treatment Assignment Mechanism

The SNAI is a place-based policy launched in 2012 under the EU Cohesion Policy 2014–2020. It is implemented through integrated area-based projects financed by national resources and EU structural funds (ERDF, ESF, EAFRD) for a total of €1.15 billion (Barca *et al*., 2014). By the end of the 2014–2020 programming period, 72 project areas had been selected, comprising approximately 1,060 municipalities and around 2 million residents. As of April 2025, only 20% of projects have been completed, and 54% are still ongoing, with a monitored public cost of €538 million, indicating that implementation remains at an early stage operationally.

Delving into the treatment assignment mechanism, unlike traditional open-call procedures, SNAI adopted a multi-level selection and concentration approach. The national territory was first classified by accessibility to essential services: municipalities with upper secondary schools, a DEA Level 1 hospital, and a certified railway station were designated as “centres,” while all others were grouped into belts (up to 20 min), intermediate (20–40 min), peripheral (40–75 min), and ultra-peripheral areas (over 75 min), with the latter three defining inner areas (Lucatelli *et al.*, 2022). Regional governments then proposed aggregates of IA municipalities as candidate project areas. These proposals were assessed by the national *Technical Committee for Inner Areas* using a standardised indicator grid covering demographic, socio-economic, and infrastructural conditions, followed by fieldwork and consultations with local stakeholders. This phase allowed adjustments reflecting territorial cohesion, shared development visions, and local leadership (Lucatelli *et al*., 2022). Although the process followed a coordinated national framework, the available documentation provides limited detail on how quantitative indicators were used. Similarly, the interplay between quantitative criteria and qualitative judgment introduced discretionary elements. For example, some municipalities in critical demographic conditions were initially excluded from the project areas but later included upon their explicit request; others were selected due to their importance in ensuring the development of the entire project area. These features suggest caution in assuming quasi-randomness or strong comparability between treated and untreated areas.

Importantly, the specificities of selection mechanisms may have direct implications for the plausibility of the parallel trends assumption (PTA) (see Ghanem *et al*., 2022). In the case of SNAI, these implications are multifold. Initial classifications based on largely stable territorial features could support similar pre-treatment trajectories across units. However, discretionary adjustments may introduce time-varying and unobserved factors, breaking parallel trends. Combined with the intrinsic heterogeneity of Italy's inner areas, this suggests that assuming unconditional parallel trends a priori is unwarranted. While conditioning on key pre-treatment covariates reduces observable differences between treated and eligible units and thereby makes parallel trends more plausible, unobserved or dynamic factors may still generate deviations, so conditional parallel trends may not be guaranteed. This motivates the use of covariate-adjusted, doubly robust estimation and complementary sensitivity analyses to enhance comparability and assess the robustness of the estimated effects to violations of the identifying assumption.

3. Data and Empirical Strategy

3.1. Data and Sample Construction

This empirical study draws on a balanced panel dataset of Italian municipalities observed annually from 2012 to 2023. The units of analysis are municipalities defined according to 2020 ISTAT administrative boundaries and SNAI 2014-2020's territorial map (2020 last update), ensuring territorial and temporal coherence among data sources.

The outcome variable is the *annual internal emigration rate*, defined as the ratio between the number of individuals relocating from one municipality to another within Italy and the average resident population during the same period, expressed per 1,000 inhabitants. This metric is derived from ISTAT's permanent and reconstructed intercensal population census for the 2012-2023 period.

The treatment variable identifies whether an eligible municipality was exposed to SNAI's intervention during the 2014–2020 programming period. A municipality is considered treated starting from the year in which the first SNAI project officially began its implementation, based on the "*Actual Start Date of Implementation*" recorded in OpenCoesione data (as of February 2025). Treatment status is irreversible: once a municipality is treated, it remains in the treated group for the rest of the observation period.

Additional covariates were selected for their relevance in predicting untreated potential outcomes and the probability of treatment. These include: average resident population, percentage of residents aged +75, percentage of foreigners, and average per capita income (sourced from municipal IRPEF data provided by the Ministry of Economy and Finance).

The initial raw sample included 7,904 municipalities. After data cleaning procedures to remove cases affected by boundary changes and inconsistencies in treatment assignment (i.e., municipalities formally included in a SNAI project but non selected in the SNAI framework), the final sample comprises 7,637 municipalities.

3.2. Empirical Strategy

The study's aim is to estimate the medium-term effects of SNAI-related interventions on municipalities treated within the 2014-2020 SNAI Framework. The target parameter is the *Average Treatment Effect on the Treated* (ATT), which measures the difference, on average, between what happened to units that received the intervention and what would have happened to them if they had not received it. To estimate the ATT, a *Difference-in-Differences* design was employed, given the non-random nature of the treatment. A key complication in this context arises from the staggered design of the policy: SNAI's interventions began in a staggered fashion from 2015 to 2023, with considerable variation in timing and treated cohorts' sample size. In such a context, standard Two-Way Fixed Effects estimators are prone to produce biased estimates due to inappropriate comparisons between earlier- and later-treated units (Callaway and Sant'Anna, 2021).

To address these issues, the study leveraged Callaway and Sant'Anna (2021)'s Doubly-Robust DiD estimator (DR-DiD). DR-DiD computes *group-time average treatment effects* (ATT_{gt}) (Eq. 1), which under no anticipation, common support, and parallel trends, remain consistent as long as either the outcome model or the generalised propensity score model is correctly specified.

$$ATT(g,t) = \mathrm{E}[Y_t(g) - Y_t(0) | G = g] \tag{1}$$

In this setting, anticipating behaviours are unlikely due to both the treatment assignment mechanism nature and the lack of information dissemination from central authorities to local beneficiaries (Monturano *et al*., 2025). On the other hand, violations of the common support and parallel trends assumptions may be plausible. For the former, due to the relatively small sample sizes of 2015 and 2016 treated cohorts (16 and 7 units, respectively), and for the latter, due to both the COVID-19 pandemic shock that impacted treated cohorts in between their pre- and post-treatment periods, as well as a hybrid selection mechanism that may not accommodate the PTA. Thus, units treated in 2015, 2016, and from 2020 to 2023 were excluded from the sample. These exclusions, while necessary to preserve internal validity, limit the generalisability of the findings. The omitted cohorts may differ systematically in their structural characteristics or exposure to external shocks, meaning that estimated effects should be interpreted as reflective of the 2017–2019 treated municipalities rather than the full set of SNAI beneficiaries.

Additionally, the full time series including COVID-19 years was considered to enable a medium-term evaluation of the strategy. Although the presence of an exogenous shock with heterogeneous impact on inner areas and migration dynamics, such as COVID-19, represents a credible threat to the PTA, the study relies on Rambachan and Roth (2023)'s framework to assess the sensitivity of the estimates to increasing deviations from parallel trends. More precisely, the study, rather than assuming a binary judgment on parallel trends (assumption holds vs. fails), shifts the focus towards a continuous assessment of how robust the estimates are to increasingly severe departures from the identifying assumption.

4. Preliminary Analysis and Identification Checks

SNAI operates within a highly heterogeneous territorial context, where inner areas differ markedly in structural, demographic, and economic terms. This heterogeneity is documented in Table 1, which reports summary statistics comparing “not eligible” (i.e., units not selected under SNAI), “eligible” (i.e., units selected but not yet treated until 2023), and treated units.

Table 1 – *Summary Statistics by Treatment Cohort and Eligibility Status (2012-2016).*

Variable	Not Eligible	Eligible	Treated 2017	Treated 2018	Treated 2019
Obs	32,995	2,720	130	240	555
Emigr. Rate	30.31 (14.58)	29.88 (17.22)	26.86 (9.21)	28.40 (11.84)	29.27 (15.11)
Mean Pop.	3,689.28 (6,102.24)	1,674.46 (1,876.24)	2,977.66 (2,770.90)	2,564.33 (3,973.57)	1,609.20 (1,438.34)
Pop +75	12.68 (2.65)	16.54 (5.11)	15.51 (4.57)	15.22 (6.74)	16.35 (4.87)
Pop. Forgn.	5.82 (4.52)	3.99 (3.29)	6.62 (3.25)	5.11 (3.84)	5.88 (3.99)
Mean Income	17,122.75 (3,552.69)	14,929.92 (2,859.58)	17,260.47 (3,020.98)	17,464.26 (3,423.26)	17,604.49 (3,272)

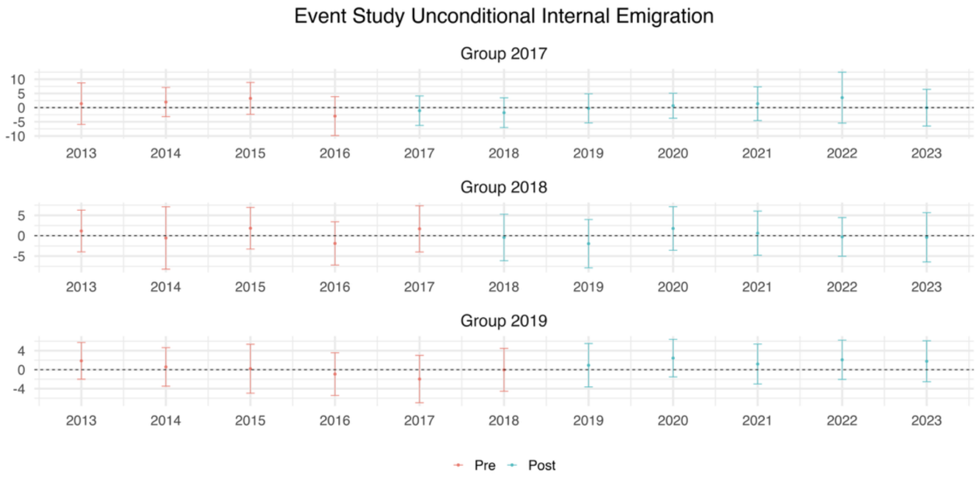
Mean and, in parentheses, Standard Deviations. Own elaboration from the ISTAT population census

Treated units exhibit lower internal emigration rates, less pronounced population ageing, and higher per capita income compared to eligible units. Although eligible units were formally included in the SNAI framework, they appear more structurally vulnerable and underperforming across most indicators than other groups. In contrast, treated municipalities tend to be smaller than non-eligible ones and display more pronounced ageing dynamics, though they report slightly higher income levels and lower emigration rates. Accounting for these differences, the selection of a valid control group was guided by alignment with the strategy’s selection logic. Non-eligible municipalities were excluded, as their characteristics placed them outside the scope of the SNAI. In contrast, eligible units were subject to the same selection mechanism as treated municipalities and were chosen based on their territorial relevance, vulnerability, and internal coherence with the characteristics of their respective pilot areas. Thus, eligible units were considered as an acceptable counterfactual group.

As a second step, the plausibility of parallel trends between treated and eligible units was assessed via an Event-Study plot estimated through the DR-DiD estimator.

Figure 1 presents event-study estimates on the full time series, including the COVID-19 period. Pre-treatment coefficients fluctuate around zero and are not statistically significant, suggesting no evidence of divergence prior to the treatment. However, the Wald test computed on joint pre-treatment coefficients rejects the null hypothesis of equal pre-trends (p-value=0.027).

Figure 1 – *Event Study estimates for the Unconditional Parallel Trend Assumption.*



ATT(g,t) estimated using Callaway and Sant'Anna (2021)'s Doubly Robust Estimator.

A clear discrepancy emerges between the event-study estimates and the Wald test: single-coefficient pre-treatment estimates are imprecise and therefore have limited power to detect small deviations, whereas the joint Wald test pools information across all pre-treatment periods and is more sensitive to systematic differences. The rejection of the Wald test should be read as circumstantial evidence that unconditional parallel trends are unlikely to hold in this context, plausibly due to residual covariate imbalance between groups. To further explore this, covariate imbalance was assessed by computing *Standardised Mean Differences (SMD)* between each treated cohort and the control group, limiting the sample to the 2012-2016 period and employing in the denominator the treated group's standard deviation. Table 2 reports the computed SMD, highlighting substantial imbalance at the 0.1 threshold level between groups and across all the variables considered. This evidence suggests that systematic differences in observable characteristics may have contributed to the violation of the unconditional PTA and further supports the need to consider covariate-adjusted estimators in the main analysis.

Table 2 – *Standardised Mean Difference Pre-treatment Covariates (2012-2016).*

Variable	Control	Treated 2017	SMD	Treated 2018	SMD	Treated 2019	SMD
Obs	2,720	130		240		555	
Mean Pop.	1674.46 (1876.15)	2977.66 (2770.90)	0.47	2564.32 (3973.57)	0.22	1609.20 (1438.34)	-0.05
Pop +75	16.54 (5.11)	15.51 (4.57)	-0.22	15.22 (6.74)	-0.20	16.35 (4.87)	-0.04
Pop. Forgn	3.99 (3.29)	6.62 (3.25)	0.81	5.11 (3.84)	0.29	5.88 (3.99)	0.47
Mean Income	14929.92 (2859.58)	17260.47 (3020.98)	0.77	17464.26 (3423.26)	0.74	16604.49 (2766.12)	0.61

Mean and, in parentheses, Standard Deviations. Own elaboration from ISTAT pop. Census.

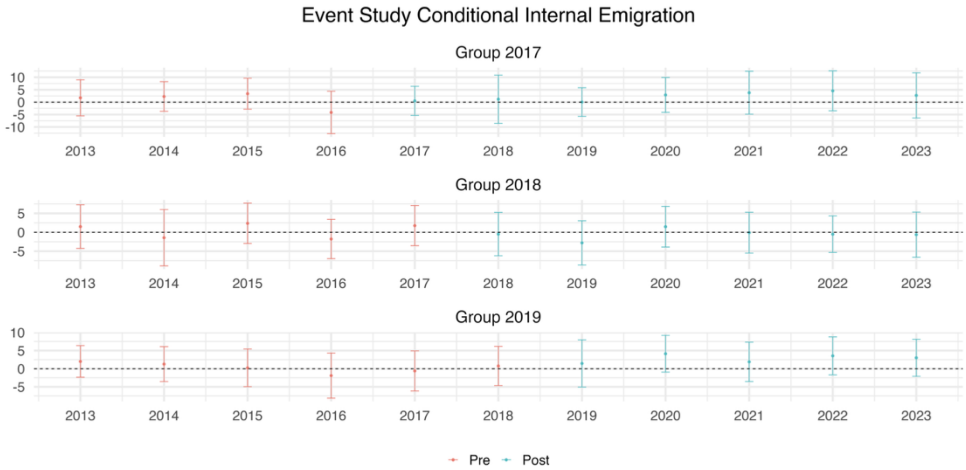
5. Results

This section presents the estimated effects of SNAI interventions on internal emigration rates employing the Doubly Robust Difference-in-Differences estimator (Callaway and Sant'Anna, 2021).

Identification checks highlighted the necessity to account for observable differences between treated and control units, as the unconditional PTA is implausible in this context. The estimates, in this respect, are adjusted for imbalanced covariates to re-establish parallelisms, among these: average population, percentage of residents aged +75, percentage of foreign residents, and average per-capita income. Figure 2 displays ATT(g,t) for 2017, 2018, and 2019 treated cohorts. Post-treatment coefficients are not statistically significant across periods and cohorts. However, the direction and magnitude of the estimated effects differ across groups and time horizons.

Regarding the conditional PTA, no individual pre-treatment ATT(g,t) is statistically significant, and the Wald test ($p = 0.246$) provides evidence of parallelism in the pre-treatment periods. This reflects improved comparability after controlling for covariates. However, pre-treatment diagnostics alone remain uninformative about the PTA in current periods (Callaway and Sant'Anna, 2021). Thus, while the conditional specification mitigates imbalance-driven discrepancies and provides some evidence of parallelism in the pre-treatment period, it does not eliminate uncertainty about whether the conditional PTA holds.

Figure 2 – *Conditional ATT(g,t) Estimates.*

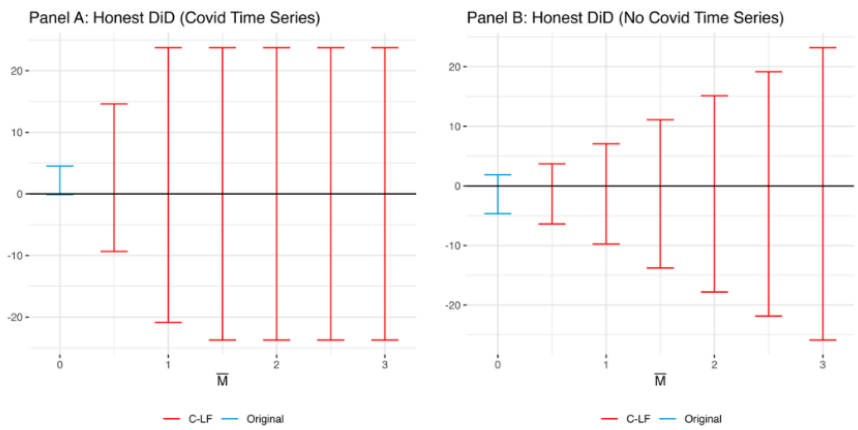


Covariates controlled: average population, percentage of residents aged +75, percentage of foreign residents, and average per-capita income.

5.1. *Sensitivity Analysis*

Given the limited informativeness of the pre-treatment diagnostics, it remains plausible that the conditional PTA may not hold exactly. For this reason, the analysis turns to the partial-identification framework of Rambachan and Roth (2023), which provides a principled way to evaluate how conclusions about the treatment effect change under increasingly conservative departures from the conditional PTA, while accounting for both statistical and identification uncertainty. The sensitivity analysis is performed on both the full time series, including COVID-19 years, and a trimmed time series, restricted to the 2012-2019 period. This allows us to remove a major source of bias and assess whether the evidence obtained in the full time series reflects structural identification issues rather than the confounding effect of COVID-19 alone. In this setting, relative magnitude restrictions are employed. Through the parameter $\bar{M}$, post-treatment deviations from parallel trends are allowed to scale proportionally to the largest violation observed between any two consecutive pre-treatment periods. Specifically, robustness was tested in the fourth year after implementation in the full time series and the first year after treatment in the trimmed series. The fourth period was preferred since all groups were exposed to the strategy for at least 4 years. Figure 3 presents the results obtained using the HonestDiD package separately for the full sample (Panel A) and the restricted sample (Panel B).

Figure 3 – *Honest DiD Sensitivity Analysis.*



The first coefficient derives from the original ATT computed via dynamic effects and a universal base period.

In Panel A, the baseline ATT(g,t) at event time 4 is 2.20 (SE=1.21), with a CI of [-1.01; 5.41]. Using the largest pre-treatment deviation (1.55) to calibrate M, the robust CI widens substantially even under mild deviations form parallelism: at M=1 (i.e. post-treatment deviations are as large as the deviations observed in the pre-treatment period) it spans [-20.85; 23.73], and at M=2 and M=3 (i.e. deviations twice as large and three times greater) is [-23.73;23.73]. Importantly, setting the M parameter to 2 or above appears empirically plausible, given that the mean variation in the outcome between 2019 and 2020 was -3.7 per thousand.

Panel B, focusing on short-term effects at event time 1 and excluding pandemic years, produces similarly wide intervals. Starting from an ATT(g,t) of -1.40 (SE=1.65) [CI: -5.61; 2.81], and a maximum pre-treatment deviation of -2.13, honest confidence intervals span from [-9.74;7.05] at M=1, to [-17.81; 15.12] at M=2.

While these results confirm the non-significance of the treatment effects in both specifications, HonestDiD estimates suggest that even moderate deviations from the conditional PTA are enough to render causal conclusions fragile, highlighting the limited informativeness of the estimates. In this regard, honest intervals include potentially economically significant effects in both directions that can be considered of high relevance for the strategy. Crucially, this fragility is not limited to the specification including COVID-19 years but persists in the trimmed series, where parallelism appeared more plausible. This pattern indicates that the available data do not provide enough information to distinguish between a truly null effect and policy effects that may exist but can not be reliably identified once even modest deviations from conditional PTA are allowed. From a substantive standpoint, the absence of detectable changes after almost a decade remains consistent with an effect that is either genuinely limited in magnitude or insufficiently identified under the constraints of the available design

6. Conclusions

This study examines the long-term impact of the National Strategy for Inner Areas on internal emigration rates, using a staggered Doubly Robust Difference-in-Differences design combined with a sensitivity analysis grounded in the HonestDiD partial identification framework. Results reveal estimates that are not statistically significant across groups up to four years of exposure to SNAI-related projects. This result is in line with Monturano *et al*. (2025)'s findings on the lack of significant effect of the SNAI on demographic dynamics. However, this null result must be interpreted with caution since the non-significance may not represent conclusive evidence of policy ineffectiveness. Indeed, HonestDiD results suggest that even modest deviations from the conditional PTA generate robust confidence intervals including economically significant effects in both directions. This was expected in light of the COVID-19 confounding effect. However, this fragility persists even when excluding the COVID-19 period. This reinforces the interpretation that issues lie not solely in pandemic-related confounding effects, but in structural limitations of the data and the context under study, which do not provide sufficient information to draw robust causal conclusions under incrementally relaxed identifying conditions.

Taken together, these results suggest that relying solely on visual inspection of pre-treatment coefficients or joint significance tests to assess PTA plausibility may be misleading. In the context of policy evaluation, this poses a tangible risk: it could lead to premature termination or misguided continuation of interventions based on fragile and weakly informative evidence.

Future research on SNAI would benefit from designs that are less vulnerable to violations of parallel trends or that rely on different identifying conditions altogether. Promising avenues include: (i) the use of individual-level microdata, which would allow for capturing behavioural responses obscured in aggregate flows; (ii) leveraging natural experiments; and (iii) adopting alternative methodologies such as Synthetic DiD, panel matching, or augmented synthetic controls to construct more credible counterfactuals based on pre-treatment dynamics. These approaches could substantially improve the ability to capture small but policy-relevant effects in demographically fragile areas.

References

BARCA F., CASAVOLA P., LUCATELLI S., 2014. Strategia nazionale per le aree interne: definizione, obiettivi, strumenti e governance, *Materiali ULAV,* Vol. 31, No. 10.

BENASSI F., BUSETTA A., GALLO G., STRANGE M., 2023. Neighbourhood effects and determinants of population changes in Italy: A spatial perspective. *Vienna Yearbook of Population Research,* Vol. 21, pp. 311–338.

CALLAWAY B., SANT'ANNA P.H.C., 2021. Difference-in-Differences with multiple time periods, *Journal of Econometrics,* Vol. 225, No. 2, pp. 200-230.

COPUS A., KAHILA P., DAX T., KOVACS K., TAGAI G., WEBER R., GRUNFELDER J., MEREDITH D., ORTEGA-REIG M., PIRAS M., LOFVING L., MOODIE J., FRITSCH M., FERRANDIS A. 2021. European shrinking rural areas: Key messages for a refreshed long-term vision, *TERRA: Revista de Desarollo Local,* Vol. 8, No. 8.

ELSHOF H., VAN WISSEN L., MULDER C.H., 2014. The self-reinforcing effects of population decline: An analysis of differences in moving behaviour between rural neighbourhoods with declining and stable populations, *Journal of Rural Studies*, Vol. 36, pp. 285-299.

GHANEM D., SANT'ANNA P.H., WÜTHRICH K., 2022. Selection and parallel trends. *arXiv preprint arXiv:2203.09001*.

LUCATELLI S., LUISI D., TANTILLO F., 2022 *L'Italia Lontana. Una politica per le aree interne.* Roma: Donzelli Editore.

MONTURANO G., RESCE G., VENTURA M., 2025. Short-term Impact of Financial Support to Inner Areas, *Italian Economic Journal.*

NEWSHAM N., ROWE F. 2025. The Demographic Causes of European Sub-National Population Declines, *European Journal of Population*, Vol. 41, No. 1, pp. 10.

PCM, March 2025. Piano Strategico Nazionale per le Aree Interne (PSNAI).

RAMBACHAN A., ROTH J., 2023. A More Credible Approach to Parallel Trends, *Review of Economic Studies,* Vol. 90, No. 5, pp. 2555–2591.

SONZOGNO G. V., URSO G., FAGGIAN A., 2022. Migration propensity of peripheral youth: insights from Italy. *Regional Studies, Regional Science*, Vol. 9, No. 1, pp. 709–726.

Gabriele DORE, University of Florence and Sassari, gabriele.dore@unifi.it