



Functional Inference for Assimilation Effects: Evidence from Migrants' Mental Health in Italy

Marta Pittavino¹(✉), Angela Andreella², Gaia Bertarelli²,
and Stefano Campostrini²

¹ Venice School of Management (VSM), Ca' Foscari University of Venice,
San Giobbe - Cannaregio 873, 30121 Venice, Italy
marta.pittavino@unive.it

² Department of Economics, Ca' Foscari University of Venice,
San Giobbe - Cannaregio 873, 30121 Venice, Italy
<https://www.unive.it/data/persona/28067863>

Abstract. We analyze the assimilation effect in migrants' mental health using data from the Italian health surveillance system PASSI for the period 2010-2023. The large sample size enables a detailed investigation of duration-of-stay dynamics across areas of origin and socio-demographic profiles. Assimilation is defined as the difference in the probability of depressive symptoms between migrants and natives with the same observable socio-economic characteristics, evaluated as a function of length of stay. To assess convergence patterns over time, we estimate a sampling-weighted semi-parametric model and construct stratified bootstrap-based simultaneous confidence bands and related inverse set methods to control the family-wise error rate. Assimilation is interpreted as a dynamic process in which the migrant-native gap is initially negative and converges to zero. Preliminary results do not support a convergence pattern. Instead, after several years of arrival, migrants exhibit a higher probability of depressive symptoms than comparable natives, particularly among socio-economically vulnerable individuals.

Keywords: Depression · length of stay · simultaneous confidence bands · vulnerability · PASSI data

1 Introduction

International migration has become a structural component of European societies, reshaping demographic profiles, labor markets, and welfare systems [5]. Italy, traditionally a country of emigration, has experienced substantial immigration flows over the past three decades, leading to increasing social and ethnic diversity [3, 8]. The health of migrants has been widely studied in terms of mortality and physical diseases [9]; comparatively less attention has been devoted to mental health, despite its central role in population well-being and health service

planning. Depression, in particular, represents both a major health outcome and a proxy for behavioural and psychosocial risk.

An established finding in migration epidemiology is the so-called assimilation effect [4]. Classical assimilation theory suggests that convergence toward the host society should promote well-being through improved socioeconomic opportunities, institutional access, and social inclusion [1, 6]. However, in the context of behavioural risk factors, assimilation does not simply denote convergence between migrants and natives. Rather, it refers to the empirical regularity whereby migrants often arrive with more favourable health behaviours and lower prevalence of selected risk factors, but progressively adopt the behavioural patterns of the host population as length of stay increases. Understanding whether assimilation enhances or undermines migrants' psychological well-being is crucial not only for theory development but also for the design of inclusive social and health policies in contemporary European societies [10].

In Italy, evidence based on data from the behavioural risk surveillance system PASSI [2] suggests that migrants may initially exhibit favourable health profiles, which are difficult to maintain over time [4]. Another study on migrants, focusing on the duration of stay, analyzes the gender health differences between nonmigrants and migrants [12]. However, due to a lack of data, most analyses remain descriptive and do not formally assess the temporal dynamics of the migrant-native gap under a unified inferential framework.

This study investigates the association between the length of stay and the probability of having depressive symptoms among migrants in Italy using PASSI data, building on [4] with a more detailed inferential approach and an extended time span. We then define the assimilation effect as the difference in the probability of being depressed between migrants and Italians, conditional on selected socio-demographic covariates. The inferential strategy relies on recent developments in inverse set estimation and the inversion of simultaneous confidence intervals [11] implemented within a sampling-weighted semi-parametric model and based on a stratified bootstrap procedure.

2 Data and Exploratory Data Analysis

We analyze data from PASSI [2] for the period 2010-2023. PASSI is an Italian surveillance system monitoring behavioral risk factors and preventive health measures among adults aged 18-69 residing in Italy. Its large sample size (384,409 observations with $\approx 5\%$ migrants) allows the investigation of health differentials across socio-demographic sub-populations, making it particularly suitable for studying assimilation dynamics over time. The analysis focuses on Italian-speaking migrants registered with local health units, thereby capturing individuals engaged in the formal healthcare system and an ongoing integration process.

Depressive symptoms are assessed through two questions referring to the previous two weeks: (1) days with little interest or pleasure in doing things, and (2) days feeling down, depressed, or hopeless. Following [7], we construct a binary indicator of depression equal to one if the total number of reported days

across the two items exceeds four, and zero otherwise. The socio-demographic variables included in this analysis are detailed in Table 1.

Table 1. Description of the socio-economic variables included in the analysis.

Variable	Description
Sex	Male; Female
Educational level	High if above high school, and Low otherwise
Economic status	High if individuals can easily make ends meet by the end of the month, and Low otherwise
Marital status	Single; Married; Previously Married
Area of origin	America (not HDC); Asia; EU not HDC; Italy; North Africa; Other Europe; Sub-Saharan Africa

To account for socio-economic heterogeneity in these initial exploratory plots, we classify the individuals as *vulnerable* if they simultaneously present low economic status, low educational attainment, and are unmarried; all others are categorized as *non-vulnerable*. Figure 1 reports the proportion of vulnerable individuals by area of origin. Higher proportions are observed among migrants from Sub-Saharan Africa and America (not HDC), whereas lower shares characterize migrants from Asia and Italy. This uneven distribution of structural vulnerability highlights the socio-economic heterogeneity of migrant populations that should be carefully considered when analyzing differences in depression prevalence.

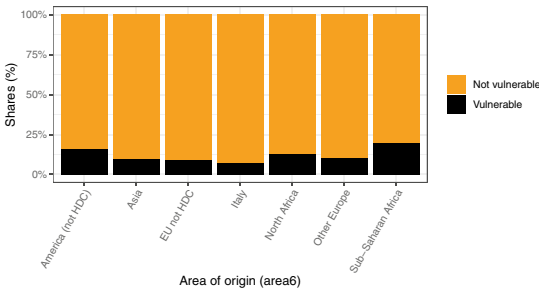


Fig. 1. Distribution of vulnerability status by area of origin.

Figure 2 shows that vulnerable individuals consistently have a higher prevalence of depression than non-vulnerable individuals across most areas of origin. In several groups, especially migrants from America (not HDC) and Sub-Saharan

Africa, this gap tends to widen with longer length of stay, suggesting cumulative disadvantage. However, both the size and trajectory of these differences vary across origin groups, indicating heterogeneous assimilation dynamics.

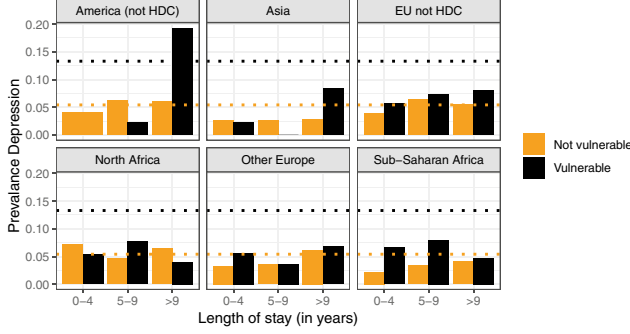


Fig. 2. Prevalence of depression by length of stay in Italy (in years) and vulnerability status, stratified by area of origin.

3 Assimilation Effect

Let Y_i be a Bernoulli random variable with parameter π_i indicating the probability of being depressed for the individual $i \in \{1, \dots, n\}$. We consider two sub-populations defined by $M_i \in \{0, 1\}$ where $M_i = 1$ stands for the individual i is migrant and $M_i = 0$ is Italian native. We then analyze the covariate $T_i \in \mathcal{T} \subseteq \mathbb{R}^+$, i.e., the length of stay in Italy, and a set of socio-demographic covariates $\mathbf{w}_i$, i.e., sex, education, economic status, and marital status defined in Table 1. For migrants only, let G_i denote the macro-area of origin.

We define the probabilities of migrant and native depressive symptoms in a fixed covariate profile $\mathbf{w}$ and $t \in \mathcal{T}$ as

$$\begin{aligned}\pi_M(t \mid \mathbf{w}, g) &= \text{pr}(Y = 1 \mid M = 1, T = t, \mathbf{W} = \mathbf{w}, G = g), \\ \pi_I(\mathbf{w}) &= \text{pr}(Y = 1 \mid M = 0, \mathbf{W} = \mathbf{w}).\end{aligned}$$

The object of interest is the assimilation function:

$$D(t \mid \mathbf{w}, g) = \pi_M(t \mid \mathbf{w}, g) - \pi_I(\mathbf{w}) : \mathcal{T} \rightarrow [-1, 1],$$

which compares migrants from the origin group g with the length of stay t with Italian natives who have the same socio-demographic profile $\mathbf{w}$. So, in our setting, assimilation is interpreted as a duration-of-exposure effect conditional on current age, thereby separating it from pure aging differences.

The assimilation hypothesis is formulated as a functional property of $D(\cdot \mid \mathbf{w}, g)$ over $t \in \mathcal{T}$. For a given $(\mathbf{w}, g)$, we investigate three components:

1. $\exists t_0 \in \mathcal{T}$ such that $D(t_0 | \mathbf{w}, g) < 0$.
2. For a tolerance $\epsilon > 0$, $\exists \mathcal{T}^0 \subset \mathcal{T}$, $\mathcal{T}^0 \neq \emptyset$ and $t_0 < \inf \mathcal{T}^0$, such that $|D(t | \mathbf{w}, g)| < \epsilon \forall t \in \mathcal{T}^0$.
3. $\exists [t_a, t_b] \subset \mathcal{T}$ with $t_b < \inf \mathcal{T}^0$ such that $\frac{d}{dt}D(t | \mathbf{w}, g) > 0 \forall t \in [t_a, t_b]$.

In practice, $D(t | \mathbf{w}, g)$ is estimated on a grid $\{t_1, \dots, t_L\} \subset \mathcal{T}$ from the predicted probabilities obtained under a semi-parametric model. To control the family-wise error rate (FWER) over t , we construct stratified bootstrap-based simultaneous confidence bands (SCBs) for $D(t | \mathbf{w}, g)$, such that

$$\text{pr}(\forall t \in \mathcal{T} : \hat{D}_l(t | \mathbf{w}, g) \leq D(t | \mathbf{w}, g) \leq \hat{D}_u(t | \mathbf{w}, g)) = 1 - \alpha$$

where $\alpha \in [0, 1]$ is the prespecified type I error.

The intervals of $\mathcal{T}$ to check the first two points that define the assimilation effect are estimated by inversion of the estimated SCB:

1. $\hat{\mathcal{T}}^+(\mathbf{w}, g) = \{t \in \mathcal{T} : \hat{D}_u(t | \mathbf{w}, g) < 0\}$
2. $\hat{\mathcal{T}}^0 = \{t \in \mathcal{T} : \hat{D}_l(t | \mathbf{w}, g) > -\epsilon \cap \hat{D}_u(t | \mathbf{w}, g) < \epsilon\}$

while the third one is checked analyzing the sign of the derivative of $D(\cdot)$.

Using this approach, we can determine when the assimilation process starts, identify potential change points, and assess for how long migrants experience an advantage or a disadvantage relative to natives, possibly across different subpopulations (e.g., vulnerable versus non-vulnerable groups). In particular, we characterize the time periods during which migrants are significantly disadvantaged by inverting the lower SCB at threshold 0:

$$\hat{\mathcal{T}}^-(\mathbf{w}, g) = \{t \in \mathcal{T} : \hat{D}_l(t | \mathbf{w}, g) > 0\},$$

which provides an interpretable summary of the timing and persistence of migrant-native gaps across origin groups, with direct policy relevance.

By [11], the SCB guarantees $\text{pr}(\hat{\mathcal{T}}^-(\mathbf{w}, g) \subseteq \mathcal{T}^-(\mathbf{w}, g)) \geq 1 - \alpha$ where $\mathcal{T}^-(\mathbf{w}, g) = \{t \in \mathcal{T} : D(t | \mathbf{w}, g) > 0\}$. Figure 3 is an example for the subpopulation of single females aged 18–24 with economic difficulties and a low education level from the Americas (not HDC). So, we can state with 90% confidence that, after 12 years, the probability of being depressed is higher for migrants than for Italians with the same socio-demographic profile described above.

4 Further Research

In the present analysis, heterogeneity is captured only through country of origin, without explicitly accounting for temporal (survey year) or spatial (place of residence) contextual variation. Future work could incorporate these additional dimensions. The framework can also be extended to a multivariate setting by jointly analyzing multiple behavioural risk factors, as in [4].

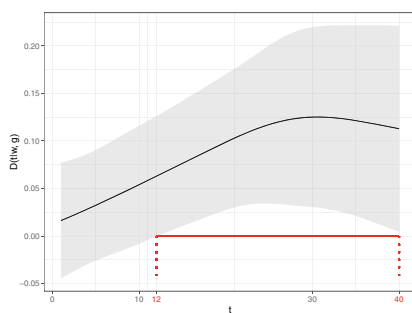


Fig. 3. Simultaneous confidence bands for $D(t \mid \mathbf{w}, g)$ for a single female aged 1824 with economic difficulties and a low level of education from the Americas (not HDC).

References

1. Alba, R., Nee, V.: *Remaking the American mainstream: assimilation and contemporary immigration*. Harvard University Press (2003)
2. Baldissera S, Ferrante G, Quarchioni E, Minardi V, Possenti V, Carrozzi G, Masocco M, Salmaso S, Group TP: Field substitution of nonresponders can maintain sample size and structure without altering survey estimates-the experience of the Italian behavioral risk factors surveillance system (PASSI). *Ann. Epidemiol.* **24**(4), 241–245 (2014)
3. Bonifazi, C., Strozza, S., Conti, C.: *The Italian transition from emigration to immigration country*. IRPPS Working Papers (2008)
4. Campostrini, S., Carrozzi, G., Severoni, S., Masocco, M., Salmaso, S.: WHO migration health programme, office of the regional director, who regional office for europe balestra Federica Severoni Santino, and PASSI national coordinating group Baldissera Sandro Bertozzi Nicoletta Bolognesi Lara Moghadam Pirous Fateh Ferrante Gianluigi Quarchioni Elisa Sampaolo Letizia Masocco maria. migrant health in Italy: a better health status difficult to maintain-country of origin and assimilation effects studied from the Italian risk factor surveillance data. *Popul. Health Metrics* **17**(1), 14 (2019)
5. Dustmann, C., Frattini, T., Preston, I.: The effect of immigration along the distribution of wages. *Rev. Econ. Stud.* **80**(1), 145–173 (2013)
6. Gordon, M.M.: *Assimilation in American life: the role of race, religion, and national origins*. Oxford University Press (1964)
7. <https://www.epicentro.iss.it/passi/indicatori/approfondimentoDepressione#>
8. ISTAT. *Rapporto annuale: La situazione del Paese*. Rome: Istituto Nazionale di Statistica (2023)
9. Kunst, A., Stronks, K., Agyemang, C.: Non-communicable diseases. In: Rechel, B., et al. (eds.) *Migration and health in the European Union*, pp. 101–120. McGraw Hill Open University Press, New York (2011)
10. Penninx, R., Garcés-Masareñas, B.: The Concept of integration as an analytical tool and as a policy concept. In: Garcés-Masareñas, B., Penninx, R. (eds) *Integration Processes and Policies in Europe*. IMISCOE Research Series. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-21674-4_2

11. Ren, J., Telschow, F.J., Schwartzman, A.: Inverse set estimation and inversion of simultaneous confidence intervals. *J. R. Stat. Soc.: Ser. C: Appl. Stat.* **73**(4), 1082–1109 (2024)
12. Trappolini, E., Giudici, C.: Gendering health differences between nonmigrants and migrants by duration of stay in Italy. *Demogr. Res.* **45**, 221–258 (2021)