The Social Connectedness Index

Drew Johnston, Meta Platforms
Theresa Kuchler, NYU Stern, NBER
Manas Kulkarni, NYU Stern
Johannes Stroebel, NYU Stern, NBER

We describe a new release of the Social Connectedness Index (SCI), a large-scale measure of social ties across geographic locations constructed from Facebook friendship links. The newly available SCI data has substantially expanded global coverage and finer geographic resolution than prior versions, and provides measures of social connectedness for national and subnational units across 178 countries.

Social connections between individuals and places play a central role in shaping economic, social, and political outcomes (Bailey et al. 2026; Chetty et al. 2022a; Jackson et al. 2008; Bailey et al. 2021). Yet despite their importance, comprehensive measures of social networks have historically been difficult to construct. Survey-based approaches are limited in scale, while administrative datasets typically capture only narrow dimensions of social ties.

To address some of these measurement challenges, Bailey et al. (2018a) introduced the Social Connectedness Index (SCI). Based on data from Facebook friendship networks, the SCI quantifies the intensity of social ties across geographic areas. Since its initial release, the SCI has been widely used in academic research and policy analysis, informing studies of migration, trade, product diffusion, and access to capital, among other topics (Bailey et al. 2020a; Bailey et al. 2020b; Bailey et al. 2021; Wilson 2022; Kuchler et al. 2022b; Kuchler et al. 2022a; Minora et al. 2023; Rehbein et al. 2025).

This note describes a new data release via the Humanitarian Data Exchange, which provides an updated version of the Social Connectedness Index as of January 25, 2026. Relative to earlier releases, the new SCI data features substantially improved coverage and finer geographic resolution.

1 Conceptual Framework

The Social Connectedness Index (SCI) measures the intensity of social connections between pairs of geographic locations using friendship links on Facebook as a proxy for real-world social ties.

Facebook friendships are undirected connections established with the mutual consent of both users, with each user able to maintain up to 5,000 friends. Because these connections typically represent real-world social relationships and require active confirmation by both parties, the resulting networks offer a meaningful proxy for offline social ties. While no single platform captures all forms of social interaction, the scale and broad geographic coverage of Facebook's user base allow the SCI to provide a uniquely comprehensive view of social connectedness across regions around the world. Facebook friendship data have therefore been widely used by researchers to study social connectedness (Chetty et al. 2022a; Chetty et al. 2022b; Jones et al. 2013; Bailey et al. 2018b; Bailey et al. 2024; Bailey et al. 2019; Bailey et al. 2022; Johnston et al. 2025a; Johnston et al. 2025b).

The Social Connectedness Index between locations i and j is formally defined as:

$$SCI_{i,j} = \frac{Friendships_{i,j}}{Pop_i \times Pop_j},$$

where $Friendships_{i,j}$ denotes the total number of Facebook friendship links between users assigned to location i and users assigned to location j, and Pop_i and Pop_j denote the number of Facebook users in locations i and j, respectively. Each friendship link is weighted equally, and where $i = j$, a small adjustment is made to the denominator to account for the fact that users cannot form friendships with themselves.

The SCI is scaled within each dataset to lie between 1 and 1,000,000,000. It is thus best interpreted as a relative measure of friendship intensity. For example, if the SCI between location A and location B is twice as large as the SCI between location A and location C, then a randomly chosen Facebook user in location A is twice as likely to be friends with a randomly chosen Facebook user in location B than with a user in location C.

2 Data Sample

The SCI is constructed using data from Facebook, a global online social networking platform. The underlying sample consists of Facebook users from 178 countries who were active on the platform in the 30 days prior to January 25, 2026 and who had more than 20 Facebook friends. As context for the scale of the platform, in September 2025 Meta reported 3.54 billion daily active people across its family of apps (Meta Platforms, Inc. 2025).

For each user, we observe an estimated home location derived from a proprietary model that combines multiple signals, including self-reported location information, IP addresses used to access Facebook, and patterns of on-platform activity. These estimated locations are used to assign users to countries and subnational regions.

Throughout the construction of the SCI, we implement safeguards to protect user privacy. Geographic units with fewer than 500 users are excluded from the data release. In addition, all published statistics are protected using μ-Gaussian Differential Privacy (Dong et al. 2022).

3 Geographic Coverage

The new SCI data release includes measures of social connectedness across the following geographic levels—many of which are available to researchers for the first time—allowing the study of social networks at local, national, and international scales:

- Country–to–country pairs
- Region–to–region pairs:
  - GADM (Levels 1 and 2)
  - GeoBoundaries (Levels 1 and 2)
  - European NUTS (Levels 1, 2, and 3)
  - US Counties and ZCTAs
- Country–to–region pairs:
  - GADM (Levels 1 and 2)
  - GeoBoundaries (Levels 1 and 2)
  - European NUTS (Levels 1, 2, and 3)
  - US Counties and ZCTAs

4 Interpretation & Data Patterns

Prior work has established several empirical patterns using the SCI (Bailey et al. 2018a), all of which replicate in the new data release.

First, the SCI declines strongly with geographic distance. Locations that are closer to one another tend to exhibit substantially higher levels of social connectedness than distant locations.

Second, political and administrative boundaries play a significant role in shaping social networks. Social connectedness often drops sharply at state or national borders, even conditional on geographic distance.

Third, the SCI captures persistent historical relationships. Measured social connectedness today reflects past migration flows, settlement patterns, and trade relationships.

Fourth, the geographic concentration of social networks varies substantially across locations. Some counties or regions are characterized by highly localized networks, while others maintain social ties with a geographically dispersed set of places. These differences are systematically related to socioeconomic characteristics such as income, educational attainment, and geographic mobility.

Figure 1 highlights that many of these patterns extend to studying data from South Asia, South America, and Africa, regions that have not been the focus of prior work with the SCI.

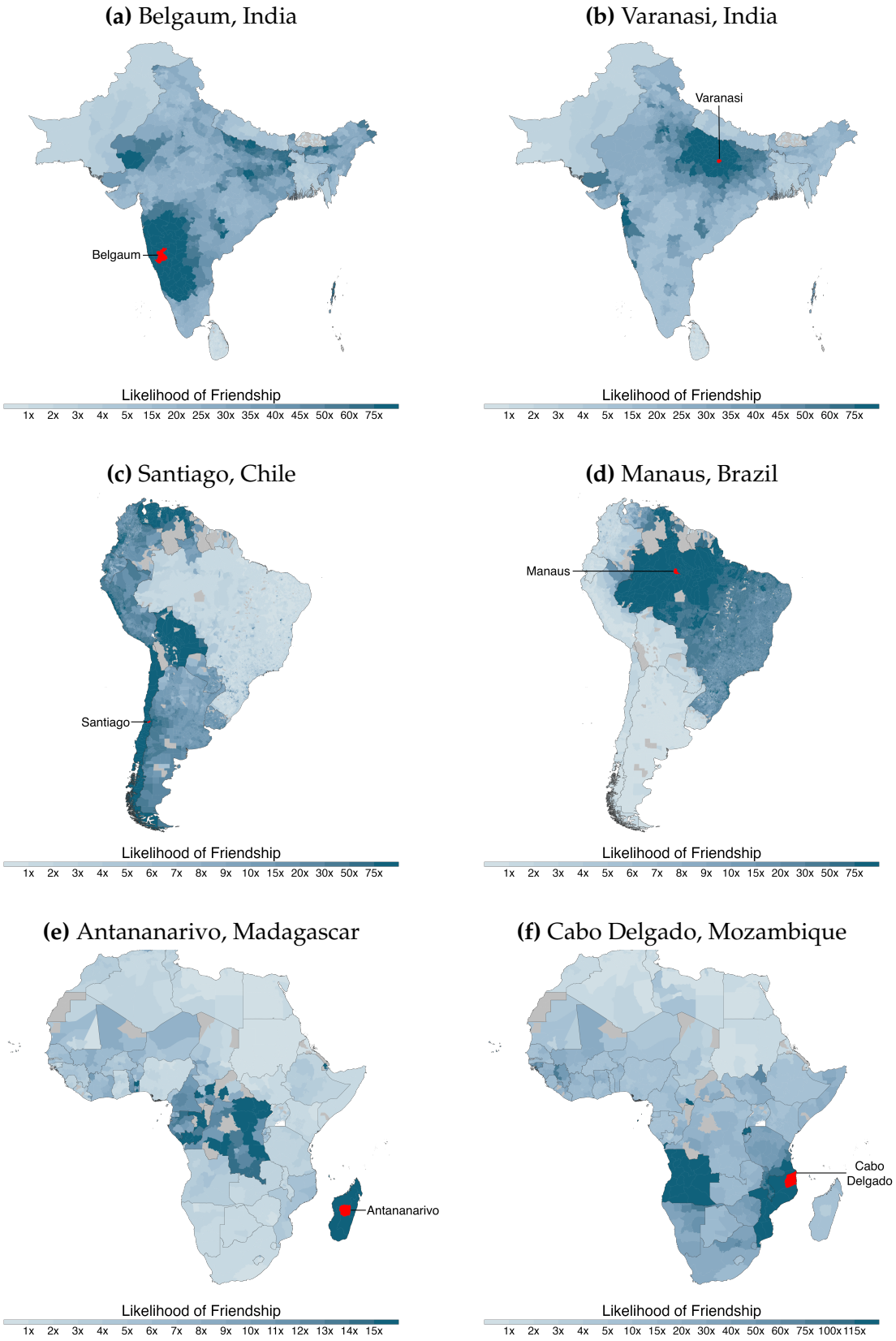

Fig. 1: Social Connectedness in South Asia, South America, and Africa

Note: Each panel shows the relative likelihood of friendship implied by the Social Connectedness Index (SCI), obtained by normalizing SCI values for the focal region by a low-connectivity reference level (the 10th percentile among regions in the countries shown). The focal region is highlighted in red.

5 Stability Over Time

Although the underlying Facebook data evolve continuously, the SCI measures a relatively stable underlying object: the geographic structure of real-world social networks.

Prior research has shown that social connectedness measured at a single point in time predicts outcomes such as trade flows many years into the past (Bailey et al. 2021), suggesting that the underlying object—how strong the social connections between different countries or regions are—does not vary at high frequency.

Consistent with this interpretation, the SCI between location pairs in the newly released data is highly correlated with the SCI in an earlier data release from 2021, with a correlation of $\rho = 0.97$ at the country level and a correlation of $\rho = 0.97$ at the U.S. county level. These findings also suggest that the 2026 release of the SCI should be useful to researchers for many years to come.

6 Geographic Aggregation

The SCI can be aggregated to higher levels of geographic aggregation using a formula introduced by Bailey et al. (2021). Specifically, the SCI between two locations i and j is defined as the probability that a representative Facebook user in location i is friends with a representative Facebook user in location j. This measure is equivalent to a population-weighted average of the SCI across all subregions within the two location.

For example, let $R(i)$ denote the set of regions in country i, and $R(j)$ the set of regions in country j. Let $Friendships_{r_i,r_j}$ be the total number of Facebook friendship links between individuals in regions $r_i \in R(i)$ and $r_j \in R(j)$. Let Pop_{r_i} and Pop_{r_j} denote the total (Facebook) population in regions r_i and r_j, respectively, and $PopShare_{r_i}$ and $PopShare_{r_j}$ denote the share of the population in those regions within their respective countries, such that $\sum_{r_i \in R(i)} PopShare_{r_i} = 1$ and $\sum_{r_j \in R(j)} PopShare_{r_j} = 1$. The SCI between countries i and j is then given by:

$$SCI_{i,j} = \frac{Friendships_{i,j}}{Pop_i \times Pop_j} = \frac{\sum_{r_i \in R(i)} \sum_{r_j \in R(j)} Friendships_{r_i,r_j}}{\left(\sum_{r_i \in R(i)} Pop_{r_i}\right) \times \left(\sum_{r_j \in R(j)} Pop_{r_j}\right)}$$

$$= \sum_{r_i \in R(i)} \sum_{r_j \in R(j)} \frac{Pop_{r_i}}{\sum_{r_i \in R(i)} Pop_{r_i}} \times \frac{Pop_{r_j}}{\sum_{r_j \in R(j)} Pop_{r_j}} \times \frac{Friendships_{r_i,r_j}}{Pop_{r_i} \times Pop_{r_j}}$$

$$= \sum_{r_i \in R(i)} \sum_{r_j \in R(j)} PopShare_{r_i} \times PopShare_{r_j} \times SCI_{r_i,r_j}$$

This formula provides a flexible framework for aggregating the SCI to higher geographic levels. In the absence of access to Facebook user counts for each region, population counts and shares from administrative data can be used as proxies.

7 Limitations

While the SCI provides a measure of social connectedness at unprecedented scale, it is important to recognize its limitations.

First, the index only reflects social ties among Facebook users. Although Facebook usage is widespread and relatively representative along many dimensions, this representativeness might vary across locations, for example when internet access is not ubiquitous.

Second, the SCI measures the prevalence of friendship links but not their intensity, frequency of interaction, or context. Two pairs of locations may exhibit similar SCI values even if the nature of the underlying social relationships differs somewhat, for example because countries might differ in the norms around adding friendship links to individuals not known in real life.

8 Conclusion

The Social Connectedness Index provides a freely available measure of social connectedness across geographies. By leveraging social network data, the SCI overcomes longstanding measurement challenges and enables new lines of research on the role of social networks in economic, social, and political outcomes. This updated release expands on

prior versions of the SCI by offering a richer set of geographic granularities, thereby enabling analysis at finer spatial scales. It is also based on a more recent snapshot of the underlying data. Together, these improvements enhance the relevance and usefulness of the SCI for contemporary research and policy analysis.

9 Suggested Citations and BibTex

Please use the following bibtex entries to cite Bailey et al. (2018a) and this note as sources for the data:

```
@article{bailey2018,
  title = {Social Connectedness: Measurement,
    Determinants, and Effects},
  author = {Bailey, Michael and Cao, Rachel and
    Kuchler, Theresa and Stroebel, Johannes and
    Wong, Arlene},
  journal = {Journal of Economic Perspectives},
  volume = {32},
  number = {3},
  pages = {259--280},
  year = {2018},
  publisher = {American Economic Association}
}

@unpublished{johnston2026,
  title  = {The Social Connectedness Index},
  author = {Johnston, Drew and Kuchler, Theresa
    and Kulkarni, Manas and Stroebel, Johannes},
  year   = {2026},
  note   = {Working paper}
}
```

References

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018a). "Social Connectedness: Measurement, Determinants, and Effects". *Journal of Economic Perspectives* **32**.(3), 259–280.

Bailey, M., Cao, R., Kuchler, T., and Stroebel, J. (2018b). "The Economic Effects of Social Networks: Evidence from the Housing Market". *Journal of Political Economy* **126**.(6), 2224–2276.

Bailey, M., Dávila, E., Kuchler, T., and Stroebel, J. (2019). "House Price Beliefs and Mortgage Leverage Choice". *The Review of Economic Studies* **86**.(6), 2403–2452.

Bailey, M., Farrell, P., Kuchler, T., and Stroebel, J. (2020a). "Social Connectedness in Urban Areas". *Journal of Urban Economics* **118**, 103264.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., and Stroebel, J. (2021). "International Trade and Social Connectedness". *Journal of International Economics* **129**, 103418.

Bailey, M., Johnston, D., Koenen, M., Kuchler, T., Russel, D., and Stroebel, J. (2024). "Social Networks Shape Beliefs and Behavior: Evidence from Social Distancing During the COVID-19 Pandemic". *Journal of Political Economy Microeconomics* **2**.(3), 463–494.

– (2026). "The social integration of international migrants: Evidence from the networks of Syrians in Germany". *Journal of Political Economy*.

Bailey, M., Johnston, D., Kuchler, T., Russel, D., State, B., and Stroebel, J. (2020b). "The determinants of social connectedness in Europe". *International Conference on Social Informatics*. Springer, pp. 1–14.

Bailey, M., Johnston, D., Kuchler, T., Stroebel, J., and Wong, A. (2022). "Peer Effects in Product Adoption". *American Economic Journal: Applied Economics* **14**.(3), 488–526.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022a). "Social Capital I: Measurement and Associations with Economic Mobility". *Nature* **608**.(79217921), 108–121.

– (2022b). "Social Capital II: Determinants of Economic Connectedness". *Nature* **608**.(79217921), 122–134.

Dong, J., Roth, A., and Su, W. J. (2022). "Gaussian differential privacy". *Journal of the Royal Statistical Society Series B: Statistical Methodology* **84**.(1), 3–37.

Jackson, M. O. et al. (2008). *Social and economic networks*. Vol. 3. Princeton university press Princeton.

Johnston, D., Kuchler, T., Kumar, A., Stroebel, J., Kulkarni, M., and Bailey, M. (2025a). "Social Capital Around the World: Measurement and Correlates". Working Paper.

– (2025b). "Social Capital Around the World: Predictors of Cross-Class Connectedness". Working Paper.

Jones, J. J., Settle, J. E., Bond, R. M., Fariss, C. J., Marlow, C., and Fowler, J. H. (2013). "Inferring Tie Strength from Online Directed Behavior". *PloS One* **8**.(1), e52168.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). "Social Proximity to Capital: Implications for Investors and Firms". *The Review of Financial Studies* **35**.(6), 2743–2789.

Kuchler, T., Russel, D., and Stroebel, J. (2022b). "The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook". *Journal of Urban Economics* **127**, 103314.

Meta Platforms, Inc. (2025). *Form 10-Q For the Quarterly Period Ended September 30, 2025*.

Minora, U., Johnston, D. M., Bosco, C., Sermi, F., Giraudy, E., Belmonte, M., and Iacus, S. M. (2023). *The War in Ukraine and the Potential of Facebook's Social Connectedness Index to Anticipate Human Displacement*. International Organization for Migration.

Rehbein, O. and Rother, S. (2025). "Social connectedness in bank lending". *The Review of Financial Studies*, hhaf014.

Wilson, R. (2022). "The impact of social networks on EITC claiming behavior". *Review of Economics and Statistics* **104**.(5), 929–945.