BUILDING PROXY INDICATORS OF NATIONAL WELLBEING WITH POSTAL DATA

PARTNERS: UNIVERSITY OF CAMBRIDGE, UNIVERSAL POSTAL UNION
PROGRAMME AREA: ECONOMIC WELLBEING

SUMMARY
This study investigated for the first time the potential of using the network of international postal flows to approximate socioeconomic indicators typically used to benchmark national wellbeing. The research used aggregated electronic postal records from 187 countries collected by the Universal Postal Union from 2010 to 2014 as a proxy indicator for real-world conditions. The indicators gathered from the international postal network were correlated with fourteen widely used socioeconomic indicators, demonstrating new approaches to approximating indicators such as the Human Development Index (HDI) and Gross Domestic Product (GDP). In addition, the postal data was combined with data from other global networks – trade, migration, international flights, Internet protocol (IP addresses) and digital communications – to produce novel multidimensional connectivity indicators.

BACKGROUND
Timely information on the socioeconomic wellbeing of communities is essential to ensure they receive effective provision of services. Many sources of big data, from mobile data to social media data, have previously been shown to yield useful proxies for socioeconomic indicators (i.e., Eagle et al., 2010; UN Global Pulse, 2015). However, digital data sources typically exhibit deeper penetration in developed economies than in developing ones, and they often rely on expensive technologies such as smartphones and a robust communications infrastructure.

This study analyzed a data source that is undoubtedly “big,” yet represents one of the most established and pervasive long-distance communications networks in the history of mankind. The postal network is the world’s largest physical network with some 680,000 post offices. Every time a letter or a parcel is sent, it leaves many digital traces including the place it was sent from (the origin) and where it was delivered (destination). The Universal Postal Union (UPU), established in 1874, collects electronic data records, or traces, each time a mail item, such as a letter or parcel, is sent across borders. In many cases, these traces come from post offices disconnected from any digital or mobile network.

This research investigated for the first time the potential of the network structure of the international postal network (IPN) to produce proxy indicators for countries’ socioeconomic profiles, analyzing 14 million records of dispatches sent between 187 countries over a four-year period, from 2010 to 2014. This study also examined the relationship between the postal network and several other global networks: the global migration network (Abel and Sander, 2014), the international flights network (Guimera et al., 2005), the IP traceroute network (Shavitt and Shir, 2005), the world trade network (Simoes and Hidalgo, 2011) and the social media density network (State et al., 2015).

DERIVING PROXIES FOR SOCIO-ECONOMIC INDICATORS FROM GLOBAL NETWORKS
This study produced aggregate measures of each country’s connectivity based on the postal network structure and flows, and then correlated different measures of connectivity with fourteen socioeconomic indicators commonly used by international organizations (including a country’s GDP per capita, human development index (HDI), Gini coefficient, consumer price index (CPI), and poverty rate).

For each network, a country’s degree of connectivity for incoming and outgoing flows was quantified. Furthermore, the study calculated the weighted degree of connectivity of a country based on the postal flows of countries to which it is linked.

HOW TO CITE THIS DOCUMENT:
Finally, all of the networks were combined into a multidimensional model using a multiplex network model (Kivelä et al., 2014). Multiplexing is the process of placing one network on top of another, so that if one network partially covers a region, then another can help plug the gaps in the corresponding socioeconomic understanding. Each separate network was treated as a different “layer” in this model, and connections between countries were considered for all of the networks in combination.

This allowed for the calculation of multiple metrics of country connectivity across all of the networks, including the global degree of connectivity of each country (the number of countries one country exchanges goods and information with across all six networks).

**INSIGHTS & OUTCOMES**

Exploring how countries interact through flows of goods, people and information can reveal important aspects of how countries relate to each other, including the strength of international ties and the network communities that countries form.

Strong correlations were identified between connectivity measures derived from all of the different global networks and socioeconomic indicators, showing the potential of such networks to serve as proxy indicators of national well-being.

The study showed that economic indicators such as GDP per capita and the Human Development Index closely correlate with indicators derived from the postal network ($r = 0.79$ for GDP and postal weighted outflows) and ($r = -0.77$ for HDI and postal weighted outflow).

This analysis also showed that postal activity has been steadily increasing since 2010, likely due to the parallel growth of e-commerce. This consistent growth and the rise of e-commerce suggest that postal flows will remain a sustainable indicator of economic activity.

![Figure 2: Spearman rank correlation coefficients between measures of connectivity in networks (horizontal) and socioeconomic indicators (vertical).](image)

Every country analyzed by the study was found to connect with an average of 110 other countries in two or more networks. The global network degree – measuring a country’s connectivity across all six networks in combination – served as the best proxy for half of the indicators considered when compared to the individual networks.

GDP per capita and life expectancy were found to be the most closely correlated with the global network degree, followed by the postal, trade and IP weighted degrees. This shows a relationship between national wealth and the flow of goods and information. Examining multiple flows together may therefore be an important approach to generate robust and multidimensional connectivity indicators that can serve as proxies for well-being.

Furthermore, partitioning the network structure, it is possible to segment clusters of countries that share flows of goods, persons and information, yielding a high-level, global, and nearly real-time picture of how different international communities evolve over time.

**IMPLICATIONS & RECOMMENDATIONS**

- This study showed that postal flows can reveal important aspects of international dynamics, and can act as a proxy indicator for many socioeconomic indicators such as GDP, poverty rate, and HDI.
- Combining the postal network with other networks generated robust multidimensional indicators that were highly correlated with the diverse socioeconomic indicators examined.
- The postal network proxies may also be useful at the subnational level: measures of postal connectivity at the community level could potentially provide a view of local wellbeing. Further research could explore such potential subnational applications.
- This project demonstrated the value of static snapshots of global networks. Further research is warranted to explore whether the dynamics of international flow networks might act as real-time proxies for economic development and wellbeing.

**REFERENCES**

UN Global Pulse. ‘Using mobile phone data and airtime credit purchases to estimate food security’, Global Pulse Project Series no. 14, 2015.


Simoes AJG, Hidalgo CA. The Economic Complexity Observatory: An analytical tool for understanding the dynamics of economic development. Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence. 2011. Available at: https://atlas.media.mit.edu/about/data/sources/


**FULL TECHNICAL REPORT**

Hristova et al. “Estimating Food Consumption and Poverty Indices with Mobile Phone Data”, PLOS ONE 2016 http://dx.plos.org/10.1371/journal.pone.0155976