

Helpdesk Research Report

Early warning models for irregular migration

Elisa Lopez-Lucia

16.07.2015

Question

What possible early warning mechanisms are in place or could be developed to help assess large scale irregular migration between developing countries, or on to developed countries? Have models been developed to help predict flows, if so have they been effective at predicting irregular migration from developing to developed countries?

Contents

1. Overview
2. Estimating and modelling (irregular) migration
3. Early warning models of forced migration
4. References

1. Overview

Estimating and predicting migrations has been a growing issue on the agenda of scholars and policymakers in the last decades. Forecasting irregular migration is of particular interest to policymakers as a tool enabling them to adapt policy to future trends.

The **meaning of irregular migration** is not always clear as there is no universally accepted definition. It is still often used interchangeably with 'illegal migration' even though 'illegal migration' is increasingly restricted to cases of smuggling and trafficking of persons. For the International Organisation for Migration (IOM), irregular migration is movement of people that takes place outside the regulatory norms of the sending, transit and receiving countries.¹

¹ See IOM key definitions: <https://www.iom.int/key-migration-terms>.

However, **forecasting irregular migration** poses various problems. While data collection on international migration has improved, it is still extremely difficult to estimate the flow or stock of irregular migrants. Moreover, most models focus on regular migration and conceptualise migration as a voluntary act which excludes forced migration. The lack of data (in particular on irregular migration) but also their lack of flexibility to integrate change in the context led to new initiatives such as the 'Global Migration Futures' project which uses scenario-building instead of forecasting.

Researchers looking at **forced migration** have developed forecasting models. Attempts have been made to design early warning models to predict refugee flows. However, scholars tend to focus more on early warning models which predict conflict and political crises that are considered the root causes of forced migration. Recently, a project was launched at the University of Georgetown to build an early warning system for detecting forced population displacement.

2. Estimating and modelling (irregular) migration

There have recently been many efforts to estimate and model migration movements. Estimations seek to improve the precision and reliability of data, while modelling aims to ascertain the factors driving migration mostly through statistical modelling (time series, multiple regression analysis, etc.) and computer simulations (Hristoski & Sotiroski 2012).

Estimating (irregular) migration

There has been an increase in the past three decades of databases on migration, funded by international organisations, think tanks and research groups (Hristoski & Sotiroski 2012).

Some of the main databases are:

- The United Nations Global Migration Database (UNGMD).²
- The Organisation for Economic Cooperation and Development (OECD) International Migration Database that stores data for OECD countries.³
- The Migration Policy Institute (MPI) Database.⁴
- The International Labour Organization (ILO) Migration Database.⁵
- The United Nations High Commissioner for Refugees (UNHCR) has compiled annual data on refugee stocks, flows, and characteristics covering more than 150 countries.⁶ UNHCR also maintains a database on monthly asylum applications submitted in 37 industrialised countries by country of citizenship.
- The UN Economic Commission for Latin America and the Caribbean has developed a regional database on the international migrant stock.⁷

In spite of these databases, the estimation of international migration flows is still very uncertain. The most important problems are: a lack of current data within and across countries; data on the volume of

² <http://www.unmigration.org>.

³ <http://www.oecd.org/els/mig/keystat.htm>

⁴ <http://www.migrationinformation.org/DataHub>

⁵ http://www.ilo.org/ilostat/faces/help_home/data_by_subject?_afzLoop=716521715438574&_adf.ctrl-state=8my0zr892_4#%40%3F_afzLoop%3D716521715438574%26_adf.ctrl-state%3D24qxctwkt_4

⁶ <http://www.unhcr.org/statistics.html>

⁷ <http://www.eclac.cl/Celade/proyectos/migracion/IMILA00e.html>

migration can be reported differently by the receiving country and the sending country; different data collection procedures and methodologies across countries and organisations complicate comparing data (Bijak & Wisniowski 2011; Hristoski & Sotiroski 2012; Santo Tomas, Summers & Clemens 2009). There are thus important gaps in knowledge on the basic stocks and flows of international migrants. According to Tomas, Summers & Clemens (2009, p. 5), the ‘nonexistence or inaccessibility of detailed, comparable, disaggregated data on migrant stocks and flows is the greatest obstacle to the formulation of evidence-based policies.’ Moreover, a major dimension of this problem is the difficulty to gather data on irregular migrants across the world (Hristoski & Sotiroski 2012; Santo Tomas, Summers & Clemens 2009).

Some specific projects such as CLANDESTINO (2007-2009), coordinated by the Hellenic Foundation for European and Foreign Policy (ELIAMEP) and funded by the European Commission (EC), have looked at irregular migration. CLANDESTINO’s aim was to collect, classify data and estimate irregular migration in 12 European Union (EU) countries. The researchers (Jandl, Vogel & Iglicka 2008) argue that while the degree of accuracy will never reach the accuracy of estimating the size of regular migrants, there is still a lot that social scientists can do to estimate irregular migrants. CLANDESTINO was therefore launched to fill this gap and resulted in an online database in 2009 that is still occasionally updated. Jandl (2004; 2008, 2011) argues that even though estimation is always problematic because by its very nature irregular migration is undocumented and unobservable, a variety of statistical methods are available. The problem is that they are almost never used in practice. He also stresses the need to undertake more qualitative research to understand irregular migration processes.

Modelling (irregular) migration

Modelling migration refers to assessing the factors that underlie decisions to migrate. For Greenwood (2005, p. 725), the ‘models are abstract depictions and simplifications of complex real-world processes that may or may not be expressed mathematically.’ Hence, the aim of empirical models is to assess the relative importance of various determinants of migration. More generally, it seeks to answer the question: why does a certain migration phenomenon occur? According to Greenwood (2005), migration models could also be used to predict migration. However, they have not frequently been used for this.

A particularity of most of these models is that they start with the assumption that migration is a voluntary human act. D’Artis’ (2005) elaboration of a model to predict migration flows from the three Baltic States to the EU after their accession is a good example of this. The last two enlargements of the EU generated a large body of literature attempting to predict the size of possible labour migration in the years following the accession. Usually, the theoretical bases for the empirical study are simple economic arguments that relate migration to gains and losses in terms of human capital and costs of migration (D’Artis 2005). Hence, this limits the use of these models for irregular migration (Greenwood 2005). Indeed, since they do not take into account migration as an involuntary act, they cannot account for migration as a result of a war or a political crisis which is often a driver of migration from developing to developed countries, and between developing countries.

More generally, it seems that no model has been developed to forecast irregular migration (expert comment). However, another expert argues that it is possible to forecast irregular migration as data tracking is more sophisticated than before. The point would be not only to understand the root causes, but also the options available inside the country, and in neighbouring countries - as these can be facilitators or obstacles. In the end, people who wish to leave (or have to) will find a way out - unless there are obstacles from keeping them to do so (expert comment).

Paoletti *et al.* (2010) also point to another problem that undermines these attempts at modelling migration. They show that statistical data is used to analyse previous trends, make assumptions about the main relationships with key variables and forecast migration flows into the future. To do so, studies rely on the use of historical and current migration data such as in the databases listed previously. For them, this is highly problematic since it ignores the uncertain context that affects migration trends and policies. They claim that forecasting models face several problems:

- They assume that the structure of the model remains the same across time: it implies that the relationship between migration and its main determinants – the structure – is going to be the same in the future.
- The lack of data: in some cases this means that parameters of projection models are estimated using historical migration data for other countries – thus assuming the same structure across countries.
- Official statistics generally do not capture irregular migration.

Hence, ‘as a result of these limitations, too often future migration trends are explored using extrapolations of current trends without taking into consideration future cultural, economic and environmental change on the global level. Such structural changes are likely to lead to fundamental changes in the direction, volume and complexity of world migration, but cannot be easily accounted for using traditional forecasting methods’ (Paoletti *et al.* 2010, p. 3).

The research project ‘the Global Migration Futures’ based at the University of Oxford and implemented by the International Migration Institute (IMI) was partly set up to respond to the problems of traditional forecasting methods. It has been trying to use an innovative scenario methodology to identify key uncertainties and relative certainties that may drive future migration. It seeks to examine future migration patterns and trends by discussing relevant variables. The methodology is both exploratory and participatory. Indeed, and this is one of its central components, it integrates migration experts, scholars from various fields and stakeholders into the process of scenario-building. Scenarios are defined as ‘stories that describe different futures that are developed using methods that systematically gather perceptions and data about certainties and uncertainties’ (Paoletti *et al.* 2010, p. 8). Hence, migration scenarios are not forecasts, the point is not so much to get it right as to have ‘a set of scenarios that illuminates the major forces driving the system, their interrelationships and the critical uncertainties’ (Paoletti *et al.*, 2010, p. 10). The process of scenario-building is based on the following factors that are assumed to be the main drivers of migration:

- Demographic factors: fertility, longevity and gender.
- Economic context: labour market structure, relative wages, poverty traps.
- South–North flows and transnational networks.
- Institutional and political environment.
- Effects of climate change and environmental degradation.

A series of policy briefing papers is now available on the ‘the Global Migration Futures’ website which draws on this methodology and explores future scenarios in various geographical areas (Pacific, Horn of Africa, North Africa, etc.).

3. Early warning models of forced migration

There have been attempts to devise early warning mechanisms in forced migration, in particular refugee flows. However, Shellman & Stewart (2007) warn that while early warning models are successful in forecasting natural disasters like droughts and storms, models to forecast humanitarian disasters like refugee movements are not as successful. Schmeidl (2003) also points to the lack of details, precision and flexibility across cases and time of these models even though the importance of 'conditional probabilities' should not be underestimated when making predictions.

The first seminal work on early warning of refugee flows was done by Lance Clark (1989). According to Clark an early warning system for refugee flows serves two functions: 1) the alleviation of the causes of refugee flows; 2) the provision of more adequate and timely refugee relief. He argues that early warning should discuss a range of scenarios, especially mass influx emergencies. For Clark, the three questions for the influx scenarios are:

- What are the major humanitarian crises which people face in the potential sending country?
- What determines which people among the affected population will respond to these problems by leaving the country?
- What events or changes are the last critical ones that might lead directly to an international refugee flow?

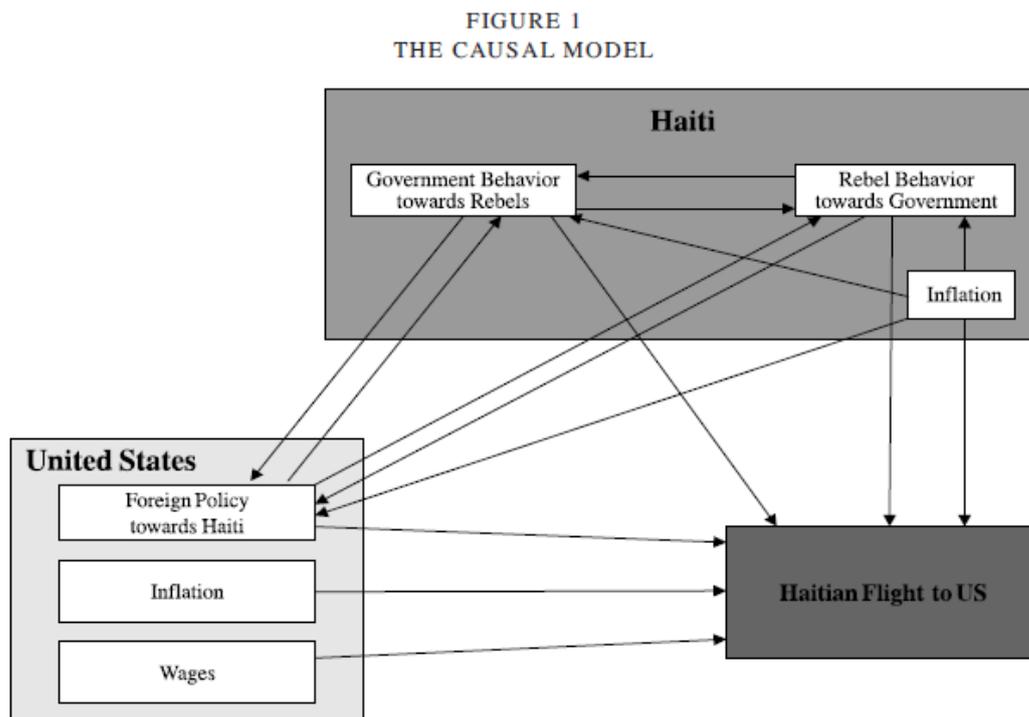
To answer these questions, the early warning system should be based on the analyses of three groups of factors: 1) push factors (root causes and proximate events); 2) intervening factors; 3) triggering factors.

Following up this seminal work, Schmeidl (1997) published a widely referenced work on early warning. In her paper she develops a theoretical model of refugee migration. She empirically tests the model to assess the role played by generalised structural factors in driving forced migration. Methodologically, she regresses the number of refugees on several political, economic, and intervening variables, using pooled time-series analysis over a 20 year period (1971-1990). Drawing on Clark, she distinguishes between three groups of factors (root causes, proximate conditions and intervening factors). She concludes that: measures of institutional human rights violations have weaker predictive power than do measures of generalised violence; civil wars with foreign military interventions are more important in producing large refugee populations and prolonged migrations than are civil wars without outside influence; ethnic rebellion is important as a cause of small refugee migrations but cannot significantly predict mass exodus; and economic and intervening variables have little impact on predicting refugee migration. She finishes by claiming that future research should seek to clarify the role of economic factors.

Later works on early warning of forced migration usually draw on Clark (1989) and Schmeidl (1997). One of these works is Davenport, Moore & Poe (2003) which conducted statistical analyses on a pooled cross-section time-series data set consisting of data from 129 countries for the years 1964–1989. Their findings support the conclusion that threats to personal integrity are of primary importance in leading people to abandon their homes. Moore & Shellman (2004) also explore the determinants of forced migration. They focus specifically on the push factors that cause people to flee their homes and ask: what characteristics of countries help explain the variation, over time and across countries, in forced migration movements? Their conclusions show that the violent behaviour of governments and dissidents (and their interaction) are the primary determinants of forced migration flows. They also suggest that high levels of dissident violence might be the strongest indicator. Finally, they argue that this model could be useful to serve an early warning capacity and develop contingency plans.

Shellman & Stewart (2007) use a case study to build a model. Their model aims to predict forced migration of Haitians fleeing to the US based on factors such as civil violence, poor economic conditions and foreign interventions. The rationale is that the model should predict the risk factors which predict migration. Figure 1 shows how these risk factors are causally related to each other. At the end, they argue they have predicted weekly flows as opposed to annual flows which enable policy-makers to better plan for humanitarian crises. They thus claim that given the model's performance, it should be reproduced in additional countries to build more precise early-warning models of forced migrant events.

Figure 1: causal model of forced migration from Haiti to the US



Source: Shellman & Stewart (2007, p. 176).

Overall very little has been done in the area of early warning of forced migration. According to Schmeidl (2003) this is because attention has shifted to the early warning of conflicts and state failure which are considered to be at the origin of forced migration.

However, recently a project has been launched at the University of Georgetown named 'Forecasting the Break: Building Community and Capacity for Large-Scale Data-Intensive Research in Forced Migration Studies.' This ongoing project includes a multidisciplinary community of scholars and practitioners to create a pilot of a large-scale, data-intensive early warning system for detecting forced population displacement. The project has five components (Forecasting the Break 2015):

- The development and validation of a theoretical model of forced migration that captures the complexity and dynamism of the phenomenon.
- The development of algorithms and methodologies that take disparate forms of data with varying degrees of reliability and completeness, and extract meaningful timely patterns that can be used in conjunction with theoretical models of forced migration as evidence for identifying potential regions of concern.

- The creation of a platform that provides timely warning of displacement using integrated knowledge from both a theoretical model of forced migration and a data-driven model based of relevant patterns of concepts, events, and other signs of displacement that exist in large open-source data repositories.
- The development of analytic tools that enable policymakers and practitioners to test different scenarios to respond to forecasted movements.
- The development of a framework that embeds ethical and privacy considerations on the use of big data into the early warning system.

4. References

- Bija, J., & Wisniowski, Arkadiusz (2011). *Statistical modelling of international migration flows*. 58th World Statistics Congress of the International Statistical Institute (ISI2011). Retrieved from: http://eprints.soton.ac.uk/208435/1/ISI2011_Bijak_STS018.pdf
- Clark, L. (1989). *Early Warning of Refugee Flows* (Refugee Policy Group Paper). Washington DC: RGP. Retrieved from: http://repository.forcedmigration.org/show_metadata.jsp?pid=fmo:3342
- D'Artis, K. (2005). Can we use NEG Models to Predict Migration Flows? An Example of CEE Accession countries. *Migration Letters*, 2:1, 32-63.
- Davenport, C., Moore, W., & Poe, S. (2003). Sometimes you just have to Leave: Domestic Threats and Forced Migration, 1964-1989. *International Interactions: Empirical and Theoretical Research in International Relations*, 29:1, 27-55. <http://dx.doi.org/10.1080/03050620304597>
- Forecasting the Break. (2015). *Forecasting the Break: Big Data Analytics for Assessing the Dynamics of Forced Migration. Progress Report and Proposed Next Steps*. Unpublished.
- Greenwood, M. J. (2005). Modeling Migration. *Encyclopedia of Social Measurement*, 2, 725-734. Retrieved from: <http://www.colorado.edu/econ/courses/spring12-4292-001/Modeling%20Migration.pdf>
- Hristoski, L., & Sotiroski, K. (2012). Conceptual Data Modelling of Modern Human Migration. *Management Information Systems*, 7:4, 3-12. Retrieved from: <http://www.ef.uns.ac.rs/mis/archive-pdf/2012%20-%20No4/MIS2012-4-1.pdf>
- Jandl, M. (2004). The Estimation of Illegal Migration in Europe. *Migration Studies*, XLI: 153, 141-155. Retrieved from: <http://www.net4you.com/jandlftp/Estimation-2004.pdf>
- Jandl, M. (2008). *Methodologies for the estimation of stocks of irregular migrants*. Joint UNECE/Eurostat/UNFPA/MEDSTAT II Work Session on Migration Statistics. Retrieved from: <http://www.docstoc.com/docs/41869680/Methodologies-for-the-estimation-of-stocks-of-irregular>
- Jandl, M. (2011). Methods, Approaches and Data Sources for Estimating Stocks of Irregular Migrants. *International Migration*, 49:5, 53-77. <http://dx.doi.org/10.1111/j.1468-2435.2011.00701.x>
- Jandl, Vogel & Iglicka (2008): Report on methodological issues. Brussels: EC. Retrieved from: http://clandestino.eliamep.gr/wp-content/uploads/2009/10/clandestino_report-on-methodological-issues_final12.pdf

Moore, W. H., & Shellman, S. M. (2004). Fear of Persecution. Forced Migration, 1952-1995. *Journal of Conflict Resolution*, 40, 5, 723-745. <http://dx.doi.org/10.1177/0022002704267767>

Paoletti, E., de Haas, H., & Vargas-Silva, C. (2010). *The concept and theory of migration scenarios*. Oxford: IMI. Retrieved from:
http://www.heindehaas.com/Publications/Concept_theory_of_migration_scenarios.pdf

Santo Tomas, P. A., Summers, L. H., & Clemens, M. (2009). *Migrant Counts. Five Steps Towards Better Migration Data*. Washington DC: Center for Global Development. Retrieved from:
http://www.cgdev.org/files/1422146_file_CGD_migration_FINAL_web.pdf

Schmeidl, S. (1997). Exploring the Causes of Forced Migration: A Pooled Time-Series Analysis, 1971-1990. *Social Science Quarterly*, 78,2, 284-308. Retrieved from: <http://www.popline.org/node/269979>

Schmeidl, S. (2003). The Early Warning of Forced Migration: State or Human Security? In Newman, E., & Van Selm, J. *Refugees and Forced Displacement. International Security, Human Vulnerability, and the State*, 130-155. Tokyo: UNU Press. Retrieved from:
http://www.swisspeace.ch/fileadmin/user_upload/Media/Publications/Journals_Articles/Publications_by_staff/Schmeidl__Susanne__The_early_warning_of_forced_migration.pdf

Shellman, S. M., & Brandon, M. S. (2007). Predicting Risk Factors Associated with Forced Migration: An Early Warning Model of Haitian Flight. *Civil Wars*, 9:2, 174-199.
<http://dx.doi.org/10.1080/13698240701207344>

Key websites

- CLANDESTINO Database on Irregular Migration: <http://irregular-migration.net/>
- Forecasting the Break, University of Georgetown: <https://isim.georgetown.edu/forecasting>
- Global Migration Futures, University of Oxford: <http://www.imi.ox.ac.uk/projects/gmf>
- IOM: <http://www.iom.int/>
- OECD International Migration Database: <http://www.oecd.org/els/mig/keystat.htm>
- MPI: <http://www.migrationinformation.org/DataHub>
- UN Economic Commission for Latin America and the Caribbean database: <http://www.eclac.cl/Celade/proyectos/migracion/IMILA00e.html>
- UNGMD: <http://www.unmigration.org>
- UNHCR: <http://www.unhcr.org/statistics.html>

Expert contributors

Dita Vogel, University of Bremen

Susanne Schmeidl, University of New South Wales (UNSW) Australia

Suggested citation

Lopez-Lucia, E. (2015). *Early warning models for irregular migration* (GSDRC Helpdesk Research Report 1241). Birmingham, UK: GSDRC, University of Birmingham.

About this report

This report is based on four days of desk-based research. It was prepared for the UK Government's Department for International Development, © DFID Crown Copyright 2015. This report is licensed under the Open Government Licence (www.nationalarchives.gov.uk/doc/open-government-licence). The views expressed in this report are those of the author, and do not necessarily reflect the opinions of GSDRC, its partner agencies or DFID.

The GSDRC Research Helpdesk provides rapid syntheses of key literature and of expert thinking in response to specific questions on governance, social development, humanitarian and conflict issues. Its concise reports draw on a selection of the best recent literature available and on input from international experts. Each GSDRC Helpdesk Research Report is peer-reviewed by a member of the GSDRC team. Search over 400 reports at www.gsdrc.org/go/research-helpdesk. Contact: helpdesk@gsdrc.org.