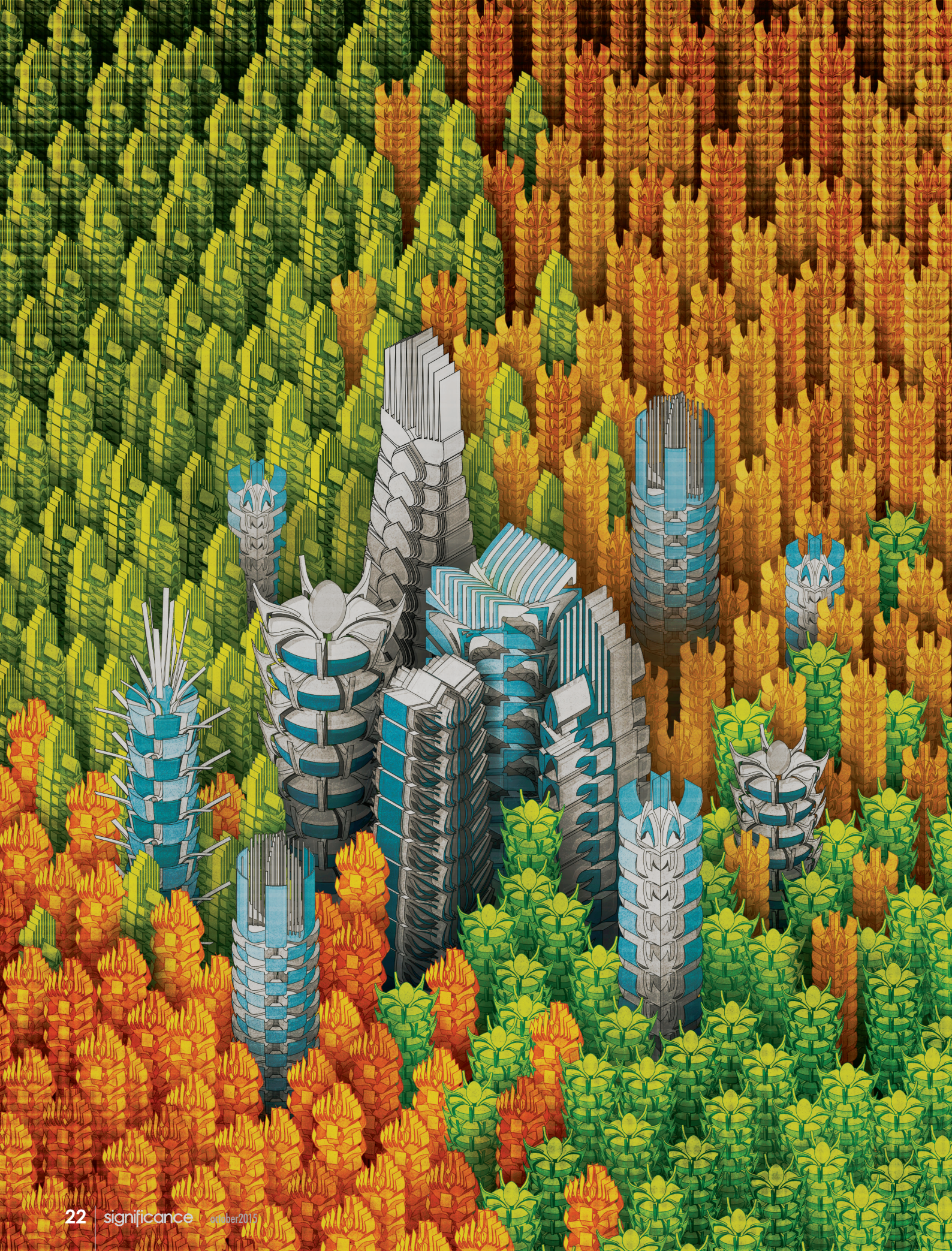


Fuel for growth and development

World Statistics Day is the ideal time to showcase diverse applications of data, and how the work of statisticians is helping improve the lives of people around the world. The next 20 pages offer merely a glimpse of the full range of problems to which statistical analysis is applied, despite touching on a wide range of issues: from economic development to trade and transport, data sharing, human rights and hunger. We begin, though, with a special introduction from **John Pullinger**, the UK National Statistician and current chair of the UN Statistical Commission, on the importance of World Statistics Day

I hope that, by now, it should come as no surprise to you that the theme of this year's World Statistics Day is "Better data. Better lives". Thanks to the website – worldstatisticsday.org – we can see versions of the logo in languages from Arabic to Vietnamese, and I'm sure that it will be a successful day for all the participating bodies around the world. This theme is vitally important for the work of the statistical community worldwide.

When I took over as UK National Statistician last year, I restated the aim of the Government Statistical Service as "Better Statistics, Better Decisions". A basic example I've already used of the power of better data to help people make better decisions is the illuminated sign on the bus stop showing when the next bus is coming. That gives us an informed choice whether to wait or start walking instead.

The history of official statistics in the UK being used for decision making is a long one; even if the Domesday Book was intended to help the King maximise his revenues, the regular publication of bills of mortality arose in late Tudor times out of the realisation that the rapid growth of London required better monitoring of the risk of plague in the public interest. More recently, Winston Churchill established the Central Statistical Office out of a perceived need to get better measures of the output of Britain's war economy.

Some of the challenges that we in the statistical community face if we are to continue to inform the decisions that will lead to better lives for everyone were examined earlier this year at the meeting of the UN Statistical Commission, which I had the honour to chair. Items on the agenda included how best to chart progress across the world on sustainable development, how to mobilise the power of the data revolution, and agreeing a plan to lift the art of census-taking to a new level.

However, given what has happened since, perhaps the most emotive area we discussed was how best to measure refugees and internally displaced people. I don't think any of us, having watched the television coverage that this issue has received this summer, can fail to recognise it as a key area where statisticians have a part to play: not only in informing decision making, but informing public debate. While much of this applies most strongly to those of us who work in official statistics, I believe that all professional statisticians have a contribution to make, whether through involvement in non-governmental organisations, making expert comment available through the media or instilling the highest standards among the students who will be the statisticians of tomorrow.

So I certainly wish every success to all those across the globe who, on 20 October and beyond, will be trying to spread the message of World Statistics Day that better data can indeed lead to better lives for all of us.

© 2015 The Royal Statistical Society

The revolution will be analysed

Two years after a United Nations report called for a "data revolution", **Claire Melamed** takes stock of how close we have come to achieving that goal

Data has never been so in demand. This may come as something of a surprise to the legions of statisticians, epidemiologists, demographers, econometricians and others who have been labouring away in the face of public and political disinterest for decades, but suddenly data is the hot political issue, at least in some circles.

The first inkling of this came with the calls for a "data revolution" in a United Nations (UN) report from May 2013. A UN panel that featured two presidents, a prime minister, the CEO of a Fortune 500 company and a Nobel Prize winner, among others, suggested this revolution was needed to "improve the quality of statistics and information available to citizens" by taking advantage of "new technology, crowd sourcing and improved connectivity". It was an idea whose time had come: this small phrase then unleashed a wave of thinking and agonising about data – and in particular, about how inadequate is the data on many of the world's gravest problems.

To the campaigners, the politicians and others who use such data, the fact that, for example, there are actual reported numbers for just 15% of all malaria deaths[1] and 16% of all births[2] was a bit of a shock. These percentages, of course, are only estimates. Around half of all the data used to monitor the Millennium Development Goals – internationally agreed targets for tackling extreme poverty – was estimated, and – over the life of the programme, which ends this year – there was no five-year period in which more than 70% of the data needed[3] to track the goals was actually available, and no one year in which more than 50% of the data was available (IEAG secretariat calculations, based on figures from UN Statistical Division).

The data used for monitoring global goals is only part of the story. A reliance on nationally representative surveys for much of the data that is used to track development targets means that for many indicators it is almost impossible to disaggregate at a sub-national level. A national figure on access to clean water, for example, is good for monitoring a country's progress on global goals, but very much less useful in planning the investments that will help a country to achieve those goals. Without knowledge of where water is short and who is lacking access, policy to improve access to water is unnervingly like guesswork. This is the situation that faces many governments as they try to plan investments and resource allocations for health services, for education and for infrastructure.

Waking up to the problem

The dawning awareness of the gaps in data has come together with other pressures to start to improve the quantity and quality of data available. Firstly, on the demand side, growing pressures from citizen groups and others for more information about what governments do in their name has made governments pay more attention to what they know and how accurate it is. Luckily, this has coincided with dramatic changes on the supply side, as new technology slowly improves every aspect of data collection, processing and use – from the use of tablets for collecting survey data to the use of mobile phone records to track disease epidemics.

So the stage is set for the data revolution. Indeed, it is happening all around us. The question is how this revolution can be mobilised to fill the gaps in data, to improve the information available to governments and citizens, and to galvanise both public and private sectors to improve services and give people more of what they want and need.

© 2015 The Royal Statistical Society

miroslav_1/iStock Editorial/Thinkstock

An Ethiopian woman sells crops in a crowded market. Farmers can now receive regular updates on the export price of key crops through the Ethiopian Commodity Exchange

Last year, the UN Secretary General asked a group of the world's experts on data, the Independent Expert Advisory Group on the Data Revolution for Sustainable Development (IEAG), to help provide some answers to that question. The final report, *A World That Counts*, which I authored, argued that a collective effort on data is needed to enable governments to both achieve and monitor the ambitious new Sustainable Development Goals, due to be agreed at a special session of the United Nations in September 2015.[3]

During the production of the report, we were offered glimpses of the world of data that awaits, if the investments are made. The Mtrac programme in Uganda, for example, supported by the children's rights organisation UNICEF, the World Health Organization and the US Agency for International Development, uses SMS surveys, completed by health workers, to alert officials to malaria outbreaks and to monitor stocks of the required drugs at health facilities. Before Mtrac, the Ministry of Health had almost no data on individual health facilities. Now, thanks to more than 20 000 health workers who contribute data, the government can collect data on thousands of health posts, and analyse the results in less than two days, at a minimal cost. The vast majority of reports are followed up within two weeks, and the number of facilities that are out of stock of critical drugs to treat malaria at any given time has fallen from 85% to 15%.[3]

It is this combination of old and new techniques, with appropriate provision for privacy and human rights, that offers the most exciting possibilities in the data revolution. It is not a question of established techniques being superseded by new methods, but of how they work together. Traditional data collection methods such as surveys can offer essential "ground truthing" to newer data sources: pictures collected by satellite, or mobile phone call records. But these new methods can offer the kind of real-time data that a survey cannot. It is by putting them together that the real, radical, improvements will be and are being made.

What next?

With so much activity already happening, it might seem that the wisest course of action is to step back and let it happen. It is certainly important that any action by governments, multilateral institutions and other players does not squash the huge creativity and innovation that has characterised the data revolution so far.

But there is a real risk of growing inequalities in access to data and in capacity to use and benefit from new sources of data. As always, it is the countries, companies and individuals with the most resources that are benefiting most from the new opportunities. To some extent this is inevitable. But it would be a tragedy if, starting at this low base, and knowing what we know about the risks of rising inequality, action were not taken to try and spread the benefits and to use the data revolution to improve data on the people and issues about which least is known – for example, the estimated 230 million children under 5 whose births are not registered.[4] This is where collective action – to encourage investment in areas of greater need, to provide incentives for filling the world's most pressing data gaps, and coming together to solve critical problems – can help to even the scales and help all to benefit.

The need is there, and the report of the IEAG has been followed by action on a range of fronts. The UN system has

been mobilised to continue the work, by committing to sponsor a "World Data Forum" in 2016. There continues to be a focus on how data can be improved as a key part of achieving the new Sustainable Development Goals – and a recent conference on Financing for Development confirmed the importance of new investments in data collection and use to make that possible. A group of governments, companies and civil society organisations – among them the USA, Kenya, Colombia and Mexico; Mastercard and Orange; and the ONE campaign and Civicus – are working together and with others to form a global partnership on data for sustainable development.

A revolution is not simply a report, a world forum or a global partnership. All of these things are just means to an end. That end is improvements in the lives of people

But a revolution is not simply a report, a world forum or a global partnership, important though they are. All of these things are just means to an end. That end is improvements in the lives of people. Human happiness can seem quite removed from analysis of satellite photos or thousands of call records on a spreadsheet, or the painstaking work of cleaning survey data. But the links are real.

The Ethiopian Commodity Exchange, for example, sends price information to farmers regularly throughout the day – and farmers' share of the export price for key crops doubled within four years of the exchange opening. As a trader commented, somewhat ruefully, at a public meeting: "Even if I wanted to cheat farmers I can't, because they know the prices before I do" (bit.ly/1NCyMtK). In Seoul, the use of mobile phone records tracking people's movement around the city at night enabled better planning and pricing of night bus routes, saving poorer residents an estimated $1.2 million overall (bit.ly/1Ikkbek).

The start of something bigger

This is just the beginning. As policy-makers, funders and campaigners wake up to how much is not known, and the ways in which this prevents action on things that they care about, the demand for more and better data can only grow. A host of new data providers, working together with established agencies, are poised to fill those gaps. And these groups are slowly being brought together, and learning how to work together for the collective good.

It may be surprising to find data in the international political spotlight. But everyone who thinks that it is important to have good data, to produce good information and ultimately good policy, should seize this moment and use it to raise the investments, the innovations and the collaborations that can create a lasting legacy.

References

1. World Health Organization (2013) *World Malaria Report 2013*. Geneva: WHO.
2. World Health Organization, UNICEF, UNFPA and World Bank (2012) *Trends in Maternal Mortality: 1990 to 2010*. Geneva: WHO.
3. Independent Expert Advisory Group on the Data Revolution for Sustainable Development (2014) *A World That Counts*. Report prepared for the UN Secretary General. bit.ly/1DcbDol
4. United Nations Children's Fund (2013) *Every Child's Birth Right: Inequities and Trends in Birth Registration*. New York: UNICEF.

Claire Melamed is the director for poverty and inequality at the Overseas Development Institute, an independent think tank based in London. From September to December 2014 she was seconded to the office of the UN Secretary General as the head of secretariat, and was lead author of the report of the Independent Expert Advisory Group on the Data Revolution for Sustainable Development

Shared learning

Collaborating on statistical initiatives is one way of plugging gaps in knowledge. **George Bentley** explains how Statistics Canada is helping Caribbean countries learn more about themselves

Canada owes its national statistics office in large part to a group of countries that came together in 1918 to help the Dominion Bureau of Statistics – as it was then known – find its feet. Now, as Statistics Canada, the agency will spend the next seven years offering support to 14 Caribbean countries that are looking to enhance their statistical systems. Through measuring their socioeconomic landscape, the hope is that these countries can develop policies to improve the well-being of their people.

The Project for the Regional Advancement of Statistics in the Caribbean (PRASC) focuses on four key areas for enhancement, the first being a country's national accounts – a statistical cornerstone that describes the basic elements of an economy, including national output, expenditure and income. PRASC will identify gaps in data and data collection methods and suggest solutions to address

© 2015 The Royal Statistical Society

them. It will also support the national statistical offices of the 14 participating countries as they transition to the new framework of the System of National Accounts 2008, the latest version of the international statistical standard for national accounts adopted by the United Nations Statistical Commission.

To support the national accounts, PRASC's second area of focus is on improving the infrastructure for economic surveys, particularly statistical methods and processes that result in an integrated approach to business surveys.

For example, work is already under way in Belize to help that nation build its economic statistical infrastructure, starting with its Business Register – something which is used as the principal framework for the country's economic statistics. In late June, representatives from Statistics Canada travelled south to meet with their Belizean counterparts. The objective of this first mission was to assist with the development phase of the Business Establishment Survey using the generic statistical business process model approach. The aim is to develop and produce a detailed set of supply and use tables for the country in early summer 2016.

PRASC's third key area, while similar to the second, shifts the focus from economic surveys to household surveys. The idea is to build a survey framework that can be adapted from one survey to another. It pays particular attention to disaggregating statistics by gender – an area of focus for the international community.

The fourth aim of the project is to develop communication and dissemination strategies to share information, thereby enhancing the first three areas. The Caribbean region has identified, as a priority, the need to improve dissemination tools and the exchange of data between partners, as well as strategies, tools and expertise to communicate with stakeholders and data users.

A staged approach to learning and implementing

PRASC takes a phased "learning-by-doing" approach, with a logical progression from knowledge sharing/training, to adaptation and testing of newly acquired best practices and tools, to national implementation, and finally, to regional sharing.

The need for this is real, according to the Hon. Oliver Joseph, Minister for Economic Development, Planning, Trade and Cooperatives in Grenada. Speaking at the project's first meeting in April, he explained how agencies in the region do the best they can with limited resources, but most struggle to provide the data that allow policy-makers to make informed decisions.

"Therein lies the need to measure accurately the magnitudes of economic indicators such as our gross domestic product, our external trade positions, and our rates of unemployment and poverty. No longer can we go on 'feelings, intuitions and premonitions' to make decisions that should affect the lives of our people in the region", said the minister.

Almost a century ago, the sharing of knowledge assisted Canada in shaping and building its first national statistical office. Through PRASC and other initiatives, Statistics Canada is committed to paying it forward.

George Bentley is senior writer and editor with Statistics Canada. A long-time journalist, he has been with the Canadian government since 2000, working with the departments of Finance and Agriculture, before joining the statistical agency in 2013

Inaugural meeting of PRASC, April 2015, Grenada

Signs of hunger

Julie Shortridge, **Stefanie Falconi**, **Ben Zaitchik** and **Seth Guikema** explain how statistical models can be used to predict undernourishment around the world based on climatic, social and economic factors

A surge in food prices between 2006 and 2008, and again in 2010, took many by surprise, and left many millions of people without the means to stave off hunger. A 2011 report by the United Nations Food and Agriculture Organization (FAO) identified several factors that led to the price rises, among which were weather shocks such as drought, which led to a decline in wheat production, and rising production costs and increased demand for meat (and animal feed) in large developing countries.[1]

Evaluated models

The following models were tested for their ability to predict undernourishment.

- generalised linear regression models (GLMs);
- generalised additive models (GAMs), which predict the response variable as the sum of multiple smoothing functions applied over each predictor variable;
- Gaussian mixture models (GMM), which assume that unidentified subpopulations exist within the sample population;
- a single classification and regression tree (CART), which partitions the data into binary subsets based on the value of predictor variables;
- a Bayesian additive regression tree (BART), which averages predictions over multiple regression trees whose structures are constrained by an underlying Bayesian probability model;
- a bagged CART (BC) model, which averages predictions over multiple regression trees, each of which is fitted on a bootstrap resampled version of the data;
- a random forest (RF) model, which averages predictions over multiple regression trees, each of which is fitted using a subset of predictor variables to reduce correlation between tree predictions;
- an ensemble mean model, which averaged predictions from the highest-performing models described above.

Two null models were included for comparison purposes: in evaluation 1 the null model assumed that undernourishment remained unchanged from the previous period, and in evaluation 2 the null model's prediction was equal to the mean prevalence of undernourishment in all countries.

The FAO forecast for the future of food prices was far from comforting. "Prices are generally expected to rise because continued population and economic growth will put upward pressure on demand, as will the anticipated increased use of biofuels", it said. "Natural resource constraints, especially climate change and the limited availability of productive land and water in some regions, pose substantial challenges to producing food at affordable prices."

Climate change might also make food prices more unstable, the FAO noted. "If the frequency of extreme weather events increases, production shocks will be more frequent, which will tend to make prices more volatile."

While extensive research has been conducted on the relationship between climate and agricultural yields, access to food is a multidimensional problem that cannot be explained by agricultural productivity alone. If the international community wishes to take steps to tackle the knock-on effects of higher and more volatile food prices – namely, increasing hunger for those people who are no longer able to afford to eat – predictive models with demonstrable accuracy are needed to identify conditions likely to lead to food-insecure situations, support interventions, and reduce suffering.[2] In the long term, an improved understanding of the physical, cultural, and socio-economic conditions that lead to food insecurity is needed to inform adaptation strategies based on strong forecasting indicators.

A complex problem

Statistical models could play an important role in identifying food-insecure situations as measurement of undernourishment currently relies on resource-intensive surveys that are typically only available once every few years. These models could also provide

© 2015 The Royal Statistical Society

insights into the relationship between climatic, social and economic conditions and food-security outcomes. Yet, statistical analyses of food security to date have been relatively limited.

One challenge is the many complex non-linear relationships that exist between climate, socio-economic conditions, and food-security outcomes. This complexity is likely to make a muddle of linear regression approaches, leading to limited success in developing models with good predictive accuracy. The use of more sophisticated non-linear models could therefore address these issues.

Recent work by our team of researchers at Johns Hopkins University assessed multiple regression and data-mining approaches for their ability to predict country-level prevalence of undernourishment based on socio-economic, climatic, and food-production covariates – that is, variables that might have predictive ability.

Two evaluations were conducted. In the first (evaluation 1), information on undernourishment from previous years was included as a covariate, resulting in predictive models that could complement survey-based estimates. In the second (evaluation 2), previous information on undernourishment levels was not included as a covariate, resulting in models that could be used to estimate food insecurity in countries without previous measurements or projections for many years into the future.

We evaluated the relationship between climatic and socio-economic conditions and national-level food insecurity in 144 countries during the period 2000–2002. FAO estimates of the prevalence of undernourishment, which represent the percentage of the population whose food consumption is continually below a minimum level, were chosen as a response variable to represent food insecurity in each country.

Three classes of variables were used to predict undernourishment. Socio-economic variables included per-capita gross domestic product and the percentage of crop areas equipped for irrigation. In the first evaluation only, this also included the prevalence of undernourishment for the period 1995–1997 (the closest previous period for which undernourishment estimates were available).

Food trade and production variables included the quantity of food imports, net food trade, and the percentage of protein and fat in dietary consumption. Latitude was included as a covariate to account for the observed relationship between geography and economic conditions.[3]

Climatic variables included measurements of growing degree days (GDD, a measure of heat accumulation available for plant growth), precipitation, relative humidity, and soil moisture. Because a cross-sectional evaluation of mean climate conditions would simply measure differences between countries (for example, Bangladesh is warm and wet compared to Canada), we used anomalies to assess climatic conditions during the response period relative to the prior decade. For instance, if historic annual rainfall in Ethiopia has a mean value of 1000 mm and a standard deviation of 250 mm, then a year with 750 mm of rainfall would have a precipitation anomaly of –1. Standard deviations were also included to represent the variability in climatic conditions over the growing season in each country – for example, whether rainfall was characterised by regular moderate events or heavy storms followed by long dry periods.

Multiple models were tested using 100 iterations of random hold-out cross validation. In each iteration, approximately 10% of samples were randomly selected and removed from the data set. The remaining data was

Table 1. Mean and standard deviation (SD) of MSE for models based on hold-out cross-validation. Bold italic results indicate models that outperform the null model based on mean squared error

	GLM1	*GLM2*	*GAM1*	*GAM2*	*GMM2*	*GMM3*	*BART*	*CART*	*RF*	*BC*	*Ensemble*	*Null*
Evaluation 1 (including data from undernourishment in previous years)												
MSE mean (all countries)	***22.15***	***22.14***	32.52	33.03	24.24	***20.2***	29.89	***22.32***	41.03	24.51	***17.96***	23.88
MSE SD (all countries)	14.32	13.15	25.81	24.87	17.27	11.28	18.77	16.18	29.14	15.42	10.4	15.36
MSE mean (high undernourishment countries)	***37.24***	***37.44***	54.32	55.51	40.72	***33.99***	50.29	***34.65***	66.20	***39.23***	***29.45***	40.03
MSE SD (high undernourishment countries)	25.14	25.77	44.52	43.91	29.88	18.96	34.12	24.73	46.88	24.11	17.31	27.9
Evaluation 2 (no data from undernourishment in previous years)												
MSE mean (all countries)	***134.46***	***98.83***	252.65	***166.96***	***117.37***	***108.45***	***95.92***	***108.08***	***96.66***	***86.70***	***87.15***	184.47
MSE SD (all countries)	94.93	52.71	459.82	282.07	65.38	53.07	52.47	57.24	57.15	48.09	48.65	72.65
MSE mean (high undernourishment countries)	***204.67***	***153.43***	363.77	237.57	***179.70***	***166.90***	***153.38***	***161.83***	***154.00***	***128.42***	***137.33***	230.50
MSE SD (high undernourishment countries)	147.08	84.47	636.78	383.22	109.60	85.81	85.50	86.18	92.63	71.17	79.41	127.88

used to fit the models and then predict undernourishment in the removed samples, providing a comparison of their out-of-sample predictive accuracy. In each iteration, the mean square error (MSE) between predicted and actual undernourishment was estimated for all countries in the hold-out sample, as well as for only those countries with actual undernourishment above 5% ('high undernourishment' countries).

Finding the right predictors

The results of our testing indicate that country-level undernourishment can be accurately predicted by using previous undernourishment information in combination with predictor variables related to socio-economic conditions, food production and trade, and climate.

We found that the highest-performing model when data on previous undernourishment was available (evaluation 1) was the ensemble model, which averaged

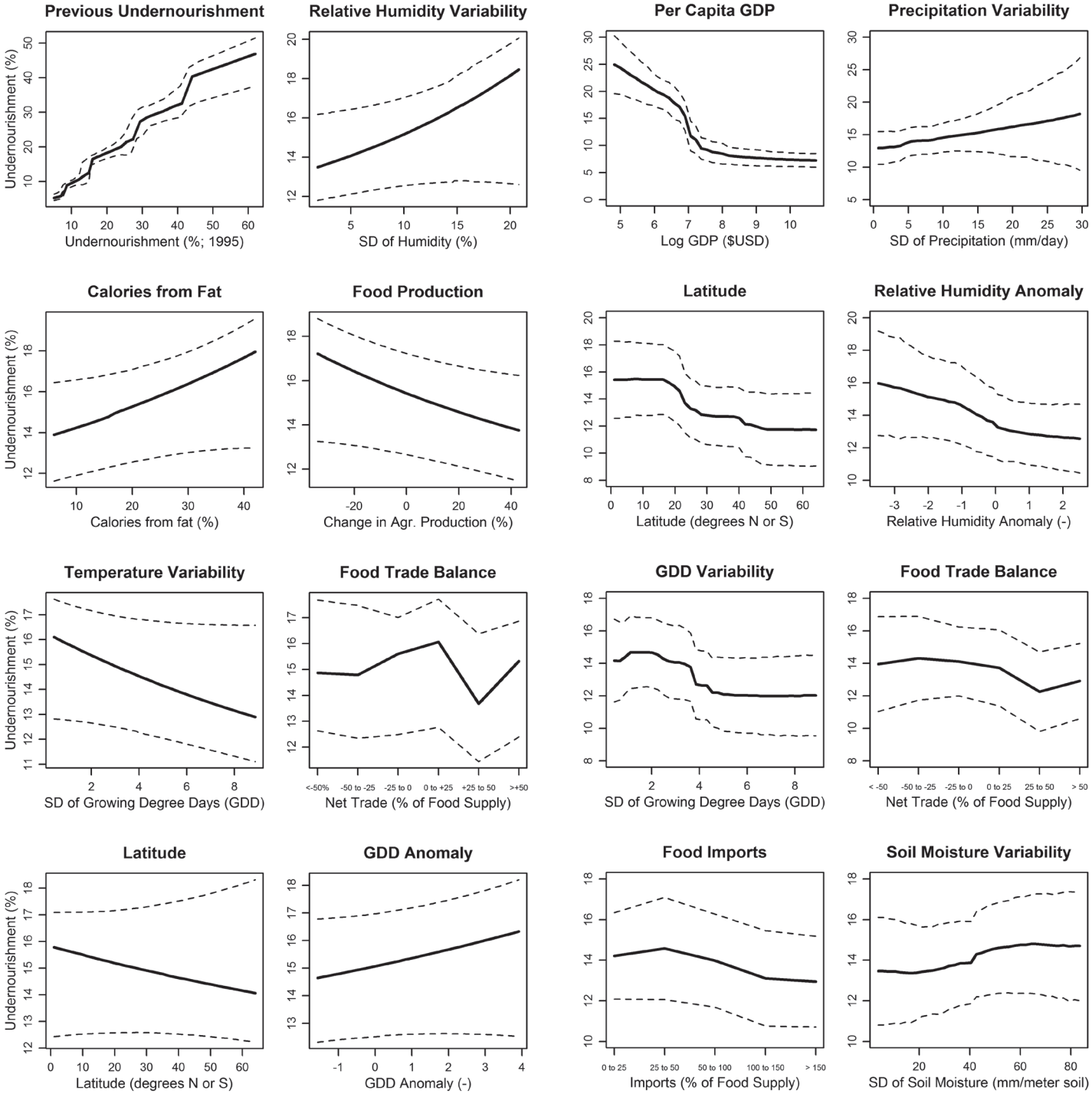


Figure 1. Relative covariate influence in the ensemble model from evaluation 1. Dashed lines represent 95% confidence bounds

Figure 2. Relative covariate influence in the ensemble model in evaluation 2. Dashed lines represent 95% confidence bounds

predictions from the LM, GMM, CART, and bagged CART models (see Table 1 and box on page 28 and 29 for description of models tested). In both evaluations, the ensemble models achieved statistically significant reductions in predictive error compared to the null models and exhibited a lower variance in MSE, indicating a reduced possibility for instances where MSE is very high.

Evaluation of the relationship between predictor variables and model predictions in the two ensemble models was conducted using partial dependence plots that demonstrate the marginal influence of changing a single predictor variable while all others remain constant. A relatively flat partial dependence plot indicates that the covariate of interest has little influence over the model's predictions, while a large change in response variable values indicates that the covariate has a large degree of influence in model predictions.

The partial dependence plots from evaluation 1 suggest that a variety of indicators are important in predicting undernourishment in the ensemble model (Figure 1). High previous undernourishment, a decrease in agricultural production, a high proportion of calories from fat and a negative food trade balance all led to high predictions of undernourishment. Highly variable humidity levels, low temperature variability, and warmer than average temperatures were also associated with high rates of undernourishment.

Without information on previous undernourishment, the ensemble model relies more heavily on GDP and latitude, with higher undernourishment predicted in low-GDP, low-latitude countries (Figure 2). Lower undernourishment is associated with a positive trade balance and high levels of food imports into the country. In terms of climatic variables, high undernourishment was predicted in countries with highly variable precipitation and soil moisture, a negative relative humidity anomaly for the 2000–2002 period, and low levels of temperature variability.

To further assess the predictive capacity of the ensemble model from evaluation 1, the ensemble and null models were used to generate predictions for the period 2006–2008 (the next period in which FAO data was available). The error (predicted undernourishment minus actual undernourishment) for each country is shown in Figure 3. Both the ensemble and null model tended to overestimate undernourishment in Africa and South America, although errors were smaller in the ensemble model. However, the ensemble model also tended to overestimate undernourishment in Southeast Asia compared to the null model.

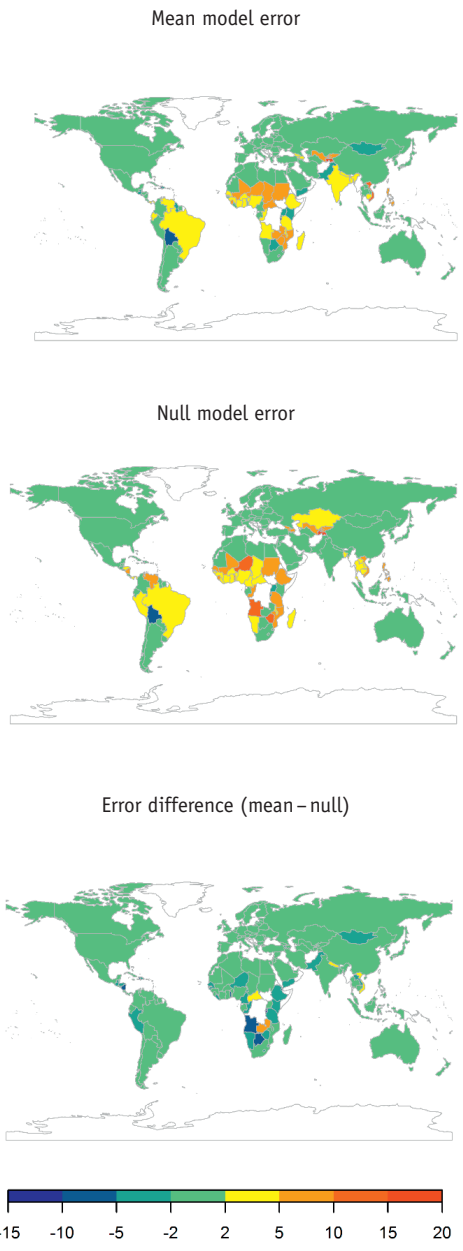


Figure 3. Model error in predicting undernourishment in 2006–2008 for the ensemble and null models

A useful tool?

While the predictions were not without error, the predictors used in the ensemble model are generally more easily estimated and more frequently available than FAO projections of undernourishment, providing a useful way to predict undernourishment when surveyed estimates of food insecurity are unavailable.

And while a cross-sectional evaluation such as this does not provide the statistical power needed to prove a causal association between predictors and undernourishment, evaluation of variable influence over model predictions can identify which indicators are the most valuable in estimating undernourishment (and thus inform data collection activities), illuminate correlations that may warrant further assessment, and inform the development of systems dynamics models.

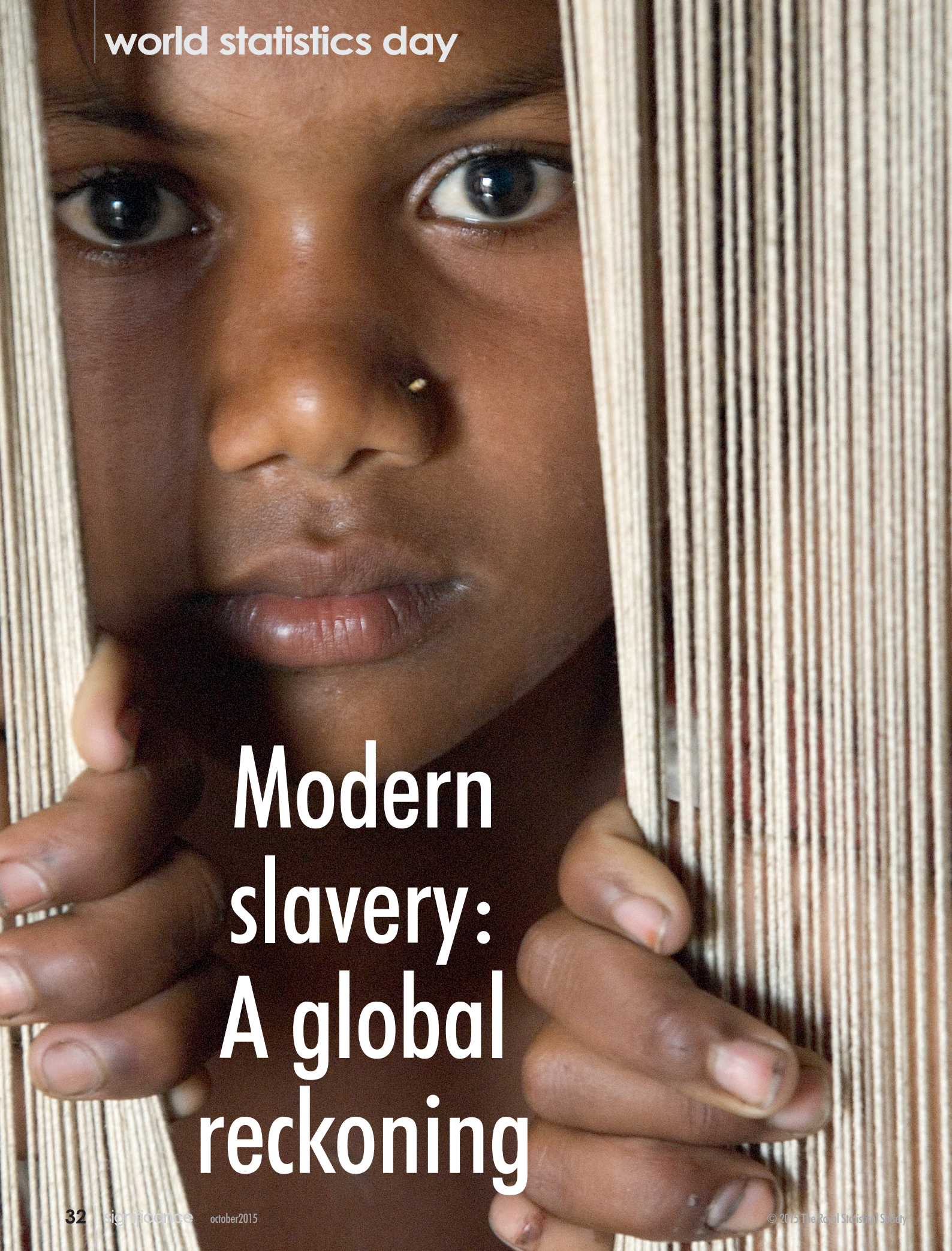
Moreover, it can identify overlooked variables that are often neglected from food security studies. For example, while a number of studies have assessed the influence that precipitation has on crop yields, our results indicate that other hydrologic variables, particularly relative humidity, evapotranspiration, and soil moisture, could be equally important in predicting food insecurity.

Indeed, we found that climatic conditions must be considered in conjunction with socio-economic influences on food availability and access to accurately predict food security outcomes. By providing insights into these relationships and a mechanism for predicting undernourishment using readily available data, statistical models like those described here could be a useful tool for organisations tasked with understanding and addressing food insecurity.

References

1. Food and Agriculture Organization (2011) *The State of Food Insecurity in the World*. Rome: FAO. bit.ly/XTiINx (accessed April 2012).
2. Barrett, C. B. (2011) Measuring food insecurity. *Science*, 327, 825–828.
3. Nordhaus, W. (2006) Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Science of the USA*, **103**, 3510–3517.

Julie Shortridge and **Stefanie Falconi** are PhD candidates in the Department of Geography and Environmental Engineering at Johns Hopkins University. **Ben Zaitchik** is an assistant professor of earth and planetary sciences at Johns Hopkins University, and **Seth Guikema** is an associate professor of industrial and operations engineering at University of Michigan



Modern slavery: A global reckoning

© 2015 The Royal Statistical Society

Slavery is an affront to the fundamental human right to liberty. In our June 2015 issue, we looked at estimates of the number of victims of slavery living in the UK. In this follow-up article, **Jacqueline Joudo Larsen**, **Monti Narayan Datta** and **Kevin Bales** discuss the crucial role extrapolation plays in assessing the number of global victims

Slavery, in its contemporary form, presents pressing challenges for the social scientist. The first is that it is ongoing, widespread, and destructive. The knowledge that any delay in illuminating and understanding this crime might have damaging or fatal results for those suffering enslavement requires a tenacious and timely response. As with public health statisticians confronting an unfolding epidemic, an unhurried research plan is not appropriate, however desirable.

The second pressing challenge is that slavery is a hidden crime. Illegal in all countries, slaveholders work hard to conceal their activities, offering a daunting task to anyone who would seek to measure the extent of slavery. This article presents both an extrapolation technique for estimating the prevalence of this crime at the national level, and a test of the accuracy of that technique.

Perhaps the last precise measurement of slavery occurred in 1860, when the United States Census enumerated those held in legal bondage. In that year, the total number of men, women, and children owned by slaveholders was 3 953 760, accounting for 13% of the US population.

From that time until the late twentieth century, however, there was no reliable measurement of the extent of slavery in any country, with the possible exception of the records of slave labourers kept by the Nazi regime during the Second World War. As the number of countries making slavery illegal increased, so did the difficulty of measuring this hidden criminal activity.

Advancing methods

With the end of the Cold War in 1989, criminal groups exploited newly opened international borders and human trafficking crime increased in the developed world. Responding to a growing concern about global trafficking and slavery in the 1990s, the first attempts were made to achieve a global estimate. In the short period since then, the methodology has advanced rapidly through three stages.

The first stage relied upon secondary sources, including governmental records, non-governmental organisations and service provider tallies, and reports in the media – in short, any source that might shed light on the extent of slavery. Even when sources were systematically assessed for reliability, these estimates[1,2] could only be seen as an approximation of the global situation.

One elaboration of this technique in the US was an attempt to triangulate secondary sources with surveys of service providers and government and law enforcement records. While the estimates derived in this first stage were not widely different from each other, there was no way to ensure their validity.

Perhaps the last precise measurement of slavery occurred in 1860, when the United States Census enumerated those held in legal bondage: 3.95m, or 13% of the US population

The second stage was set in motion by the pioneering work of Pennington *et al.*[3] in 2009. This team introduced a series of questions concerning human trafficking into a random sample health survey of five eastern European countries (Belarus, Ukraine, Moldova, Romania, and Bulgaria). Employing random sample surveys, they were able to build the first reliable estimate of the proportion of each country's population that had been enslaved, as many of those trafficked across borders often end up in hock to those who transported them and must work to pay off the costs of their journey.

The work of Pennington *et al.* was critical to the advancement of the methodology for two reasons. Firstly, it demonstrated that, at least in some countries and circumstances, enslavement could be measured through random sample surveys of the full population. Secondly, by fixing valid data points for these five countries, it became possible to begin the process of estimating the *range* of prevalence of modern slavery across countries, as well as using these and other emerging survey results to extrapolate the prevalence of slavery in other countries.[4]

In addressing the question of range, by 2009 it had become clear that slavery cases were being reported in virtually all countries with a population over 100 000. For that reason, the low end of the global range of prevalence of modern slavery for countries in which measurement was possible was assumed to be greater than zero. In the same year, a report from the United States Agency for International Development and the Pan American Development Foundation[5] provided a random sample survey of child *restavèk* slavery in Haiti. The proportion of the Haitian population (2.3%) found to be in this pervasive form of slavery was assumed to be in the upper range of the global distribution of slavery.

In 2014, a more robust extrapolation became possible with this sense of the range of prevalence across countries, supported by an increase in the amount of data available from random sample surveys. In addition to data from the Pennington *et al.* and Haiti surveys, random sample surveys of slavery were also identified in three more countries (Niger, Namibia, and the Democratic Republic of Congo), and an additional seven (Pakistan, Indonesia, Brazil, Nigeria, Ethiopia, Nepal, and Russia) were commissioned by the Walk Free Foundation using the Gallup International World Poll (bit.ly/1FfBCN8), which also provided data for three other countries (Saudi Arabia, Qatar, and Malaysia), resulting in data points from a total of 19 random sample surveys.

At the most basic level the process of extrapolation assumed that the 167 most populous countries were ranged along a continuum from those with the lowest

Image: © Kay Chernush for the US State Department

prevalence of slavery to those with the highest prevalence of slavery. Further, the data points derived from the 19 random sample surveys available were considered to be indicative of similar countries along the continuum.

In the 2014 Global Slavery Index, the process of extrapolating estimates of modern slavery followed five steps. In addition to the collection of data on prevalance, a large number of variables were collected for all 167 countries to measure vulnerability to enslavement across five dimensions: state stability, social discrimination, presence or absence of human rights protections, economic and social development, and presence of governmental slavery policy. The first step involved applying a *K*-means cluster analysis (see box) to group the 167 countries into seven distinct groups using these dimensions of vulnerability. The research team employed *K*-means as a clustering algorithm, deciding on the final cluster out of a dozen trials that had the highest pseudo-*F* score. Seven groups were chosen as there were seven prevalence surveys that used the same instruments and data collection methods, and ideally, each survey could be applied to a unique group.

The seven groupings were sufficiently distinct on overall mean values, although the minimum and maximum values did indicate some overlap among countries at the bottom of one list and top of the next.

The extrapolation process then followed these steps:

1. Countries within each group were ranked according to their mean vulnerability score, from low to high. This aggregated countries that are more alike as compared to those that are less alike.
2. Geography was also important to consider within and among clusters, given that some countries share many similar attributes that correlate highly with geography. Where available and relevant, the final calculation of estimated enslavement would incorporate survey data for a country within the same geographical region to reflect regional similarities.
3. Once this process was concluded, the research team examined countries on a case-by-case basis to determine if the extrapolation process corresponded to data collected from secondary source estimates. Country-level adjustments were also made in discussion with experts on specific countries.

K-means cluster analysis

K-means is a statistical method which groups similar items into clusters, ensuring that items not in the same cluster are as different as possible. This is achieved by allocating an item to the cluster with the nearest centroid, or the mean of the cluster. The cluster's mean is then recalculated and the process of allocating items to clusters begins again until no items change groups, or those changes do not make a substantial difference in the definition of a cluster. The Calinski–Harabasz pseudo-*F* score describes the ratio of between-cluster variance to within-cluster variance, whereby large values indicate distinct clustering.

Image: © Kay Chernush for the US State Department

4. A final, downward, adjustment was made for small island developing states in order to ensure that estimates for these nations erred on the conservative side in light of (in most cases) their remoteness, small but growing populations, limited resources, susceptibility to natural disasters, and limited potential for economic growth, among other development challenges.

This extrapolation process yielded a prevalence estimate for each country calculated as a proportion of the total population that was enslaved within that country. For all 167 countries this produced a total global estimate in 2014 of 35.8 million enslaved people.

A valid estimate?

Following the release of the Global Slavery Index in 2014, a further 19 surveys were commissioned by the Walk Free Foundation. In addition, multiple systems estimation (MSE) was used to compute a new prevalence estimate for the UK.[6] These new surveys allow a test of the extrapolation procedures, and can be added to previous surveys that were carried out after extrapolation estimates were made.

The survey data to which the extrapolation estimates can be compared is in three sets. Seven countries were surveyed in 2014, a further 19 in 2015, and surveys by other researchers were found for two other countries.[7,8] The addition of the UK's MSE based estimate brings the total number of countries to 29. Altogether these countries, and thus these surveys, represent 3.3 billion people, approximately 46% of the global population.

A comparison of the extrapolation estimates of slavery to the estimates derived from random sample representative surveys yields very promising results. As Table 1 shows, all but one of the extrapolation estimates of the prevalence of slavery within a national population fell within one percentage point of the estimates arrived through random sample surveys.

After the publication of extrapolation estimates in the 2013 and 2014 editions of the Global Slavery Index, there were concerns and reservations about the reliability of this technique. While we will continue to practise continuous and incremental testing and improvement of the methodology, the findings presented here provide a level of confidence in the use of extrapolation estimates in the absence of further random sample surveys.

The extrapolation technique presented here is not the only, or the best, means by which to estimate the prevalence of slavery. But at this time, and for many countries, it is the *only* method available, and will continue to be well into the future. This may be because ongoing conflict precludes on-the-ground survey work, or because the relevant government refuses to allow research to take place, or simply because of a lack of resources needed to carry out extensive random sample surveys. This test of the extrapolated estimates against the subsequent random sample survey results, however, supports the continued use of this technique until more precise methods are available.

Table 1. Difference in the percentage of the population estimated to be enslaved

	2013 extrapolation-based estimate, to 2014 survey-based estimate	*2014 extrapolation-based estimate, to 2015 survey-based estimate*
Pakistan	0.08	
Indonesia	0.20	
Brazil	0.03	
Nigeria	0.07	
Ethiopia	0.30	
Nepal	0.12	
Russia	0.37	
Namibia		0.21
Democratic Republic of Congo		0.43
India*		0.49
The Philippines		0.13
South Africa		0.26
Ghana		0.36
Thailand		0.69
Mauritania		1.59
Bangladesh		0.52
Vietnam		0.21
Cambodia		0.61
Myanmar		0.52
Sri Lanka		0.29
Tunisia		0.46
Guatemala		0.63
Chile		0.05
Dominican Republic		0.81
Poland		0.42
Hungary		0.03
Bolivia		0.21
Mexico		0.19

* In the case of India, this is a preliminary result based upon a nationally representative, random sample survey (n = 3000). The final estimate, however, will integrate data from a further 12 surveys, with samples representative at a state level. The combined sample size of 15 000 will provide a more robust analysis of India.

We have moved quickly to disseminate and publish these findings not simply because methodological testing and transparency are the absolute foundation of science, but also because of the immediate and pressing threat that slavery represents to human well-being. We are committed to ending the suffering and cost of slavery. We are certain that clear and objective measurement produced within the framework of globally accepted scientific process (open data, methodological transparency, and anonymous peer review) can provide the crucial metrics against which the progress of liberation – of human freedom – can be measured.

Acknowledgements

The research described in this paper was made possible through a significant financial investment in nationally representative, random sample surveys by the Walk Free Foundation, as part of the research programme that develops the Global Slavery Index. We also extend our appreciation to Gallup Inc., particularly, Pablo Diego-Rosell, Andrew Rzepa, and Mike Ilecki, for their methodological expertise, technical advice, and management of the Walk Free Foundation survey programme within the Gallup World Poll. With thanks, as well, for the support of the Global Slavery Index research team at the Wilberforce Institute for the Study of Slavery and Emancipation, University of Hull.

References

1. Bales, K. (1999) *Disposable People: New Slavery in the Global Economy*. Berkeley: University of California Press.

2. International Labour Office (2005) *A Global Alliance against Forced Labour: Global Report under the Follow-up to the ILO Declaration on Fundamental Principles and Rights at Work 2005*. Geneva: ILO. bit.ly/19WWQ8H

3. Pennington, J. R., Ball, W. A., Hampton, R. D. and Soulakova, J. N. (2009) The cross-national market in human beings. *Journal of Macromarketing*, **29**(2), 119–134.

4. Datta, M. N. and Bales, K. (2013) Slavery in Europe: Part 1, Estimating the dark figure. *Human Rights Quarterly*, **35**(4), 817–829.

5. Pierre, Y.-F., Smucker, G. R. and Tardieu, J.-F. (2009). *Lost Childhoods in Haiti: Quantifying Child Trafficking,* Restavèks *& Victims of Violence*. Washington, DC: Pan American Development Foundation. bit.ly/1N6TLVv

6. Bales, K., Hesketh, O. and Silverman, B. (2015) Modern slavery in the UK: How many victims? *Significance*, **12**(3), 16–21.

7. Johnson, K., Scott, J., Rughita, B., Kisielewski, M., Asher, J. and Ong, R. (2010) Association of sexual violence and human rights violations with physical and mental health in territories of the eastern Democratic Republic of the Congo. *Journal of the American Medical Association*, **304**(5) 553–562.

8. Directorate of Labour Market Services, Ministry of Labour and Social Welfare (2005) *Namibia Child Activities Survey 2005*. Windhoek, Namibia: Ministry of Labour and Social Welfare. bit.ly/1hulJNp

Jacqui Joudo Larsen is senior research manager, Walk Free Foundation. **Monti Narayan Datta** is associate professor of Political Science, University of Richmond. **Kevin Bales** is professor of Contemporary Slavery, Wilberforce Institute for the Study of Slavery and Emancipation, University of Hull

Image: © Kay Chernush for the US State Department

Monitoring travel behaviour in England

A 50-year journey

In 1965, Britain became the first country in the world to collect data on national patterns of travel behaviour through its National Travel Survey. As the survey reaches its 50th anniversary, it remains a key source of data on how and why we travel. **Delphine Robineau** and **Matthew Tranter** explore the trends

Transport matters. It provides access to employment and education opportunities, to shops and leisure. In 2014, residents of England made on average over 900 trips and spent the equivalent of an hour every day travelling. Collecting statistics to measure and understand patterns of travel behaviour to inform transport planning and policy is therefore of national importance.

The National Travel Survey (NTS) provides the only authoritative national source of data linking personal travel behaviour to the social demographic characteristics of the traveller. As the survey methodology (see box, page 41) has remained broadly unchanged since its introduction in 1965, trends can be monitored consistently. Together, these factors make the NTS uniquely important in exploring changes in patterns of travel as the background to policy formulation, modelling, and ultimately investment decisions.

The most striking trend of the past half century is that, at the individual level, we are travelling further but not more often. Indeed, the total number of trips people make on average in a year has remained fairly stable since 1965 (Figure 1). However, the distance we travel has almost doubled over the same period. And so the trips we make are getting longer on average.

But while the distance travelled has increased, the time we spend travelling has remained remarkably unchanged. So the trips we make have got faster. Perhaps even more remarkably, the distribution of trip times has also remained broadly unchanged (Figure 2).

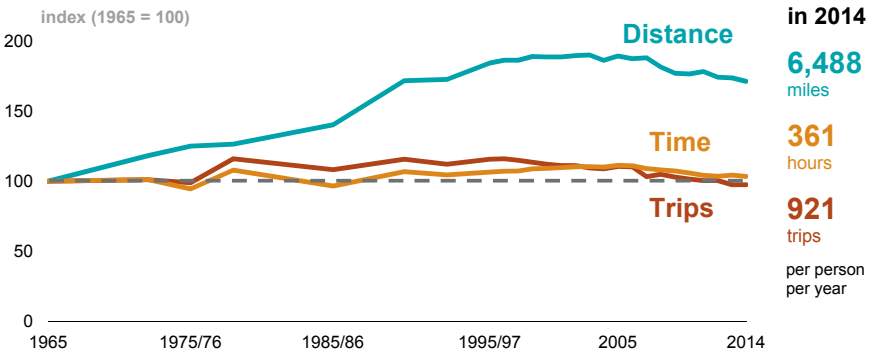


Figure 1. Trends in trips, distance and time travelled, 1965 to now

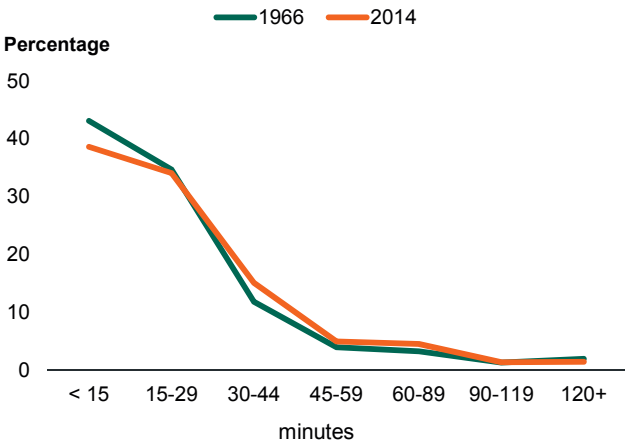


Figure 2. Distribution of trips by trip time, 1966 and now

© 2015 The Royal Statistical Society

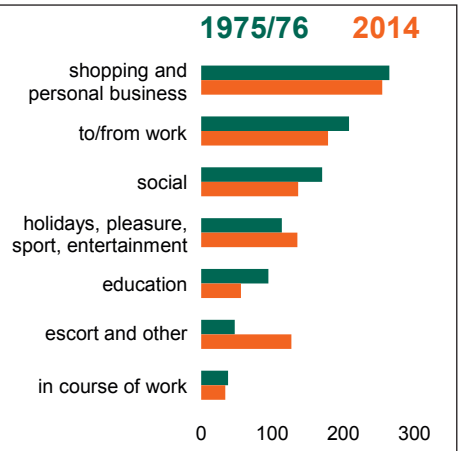
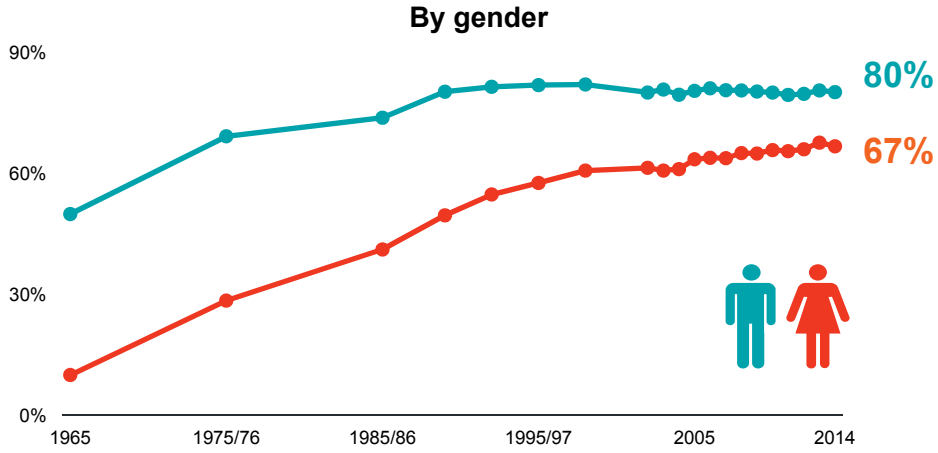
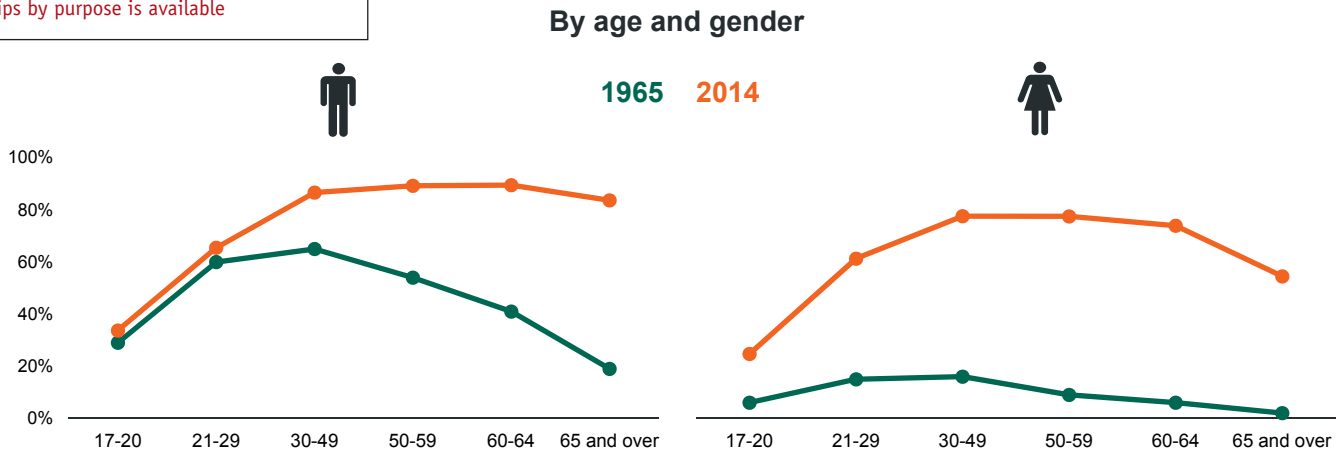


Figure 3. Trips by main purpose, 1975/76 and now. Figures presented here relate to 1975/76, the earliest date for which comparable and reliable data on trips by purpose is available





To what extent are these long-term trends driven by a change in reasons for travelling?

Why we travel

Transport is a derived demand; in most cases, people are travelling for a particular purpose, rather than solely for the pleasure of travelling. Figure 3 suggests little change in trip purpose over time. The major difference in trip purpose is the increase in escorting trips (where the main purpose is to accompany someone else), which could be partly because more people are being driven to places by car.

The rest of the difference observed in trip purpose is likely to be partly attributable to demographic changes. For example, the decrease in education trips may be explained to some extent by the decreasing share of school-age children in the population. Similarly, a larger retired population could be reducing the number of work-related trips per person averaged over the whole population. Overall, the stability of trip purposes over time is remarkable.

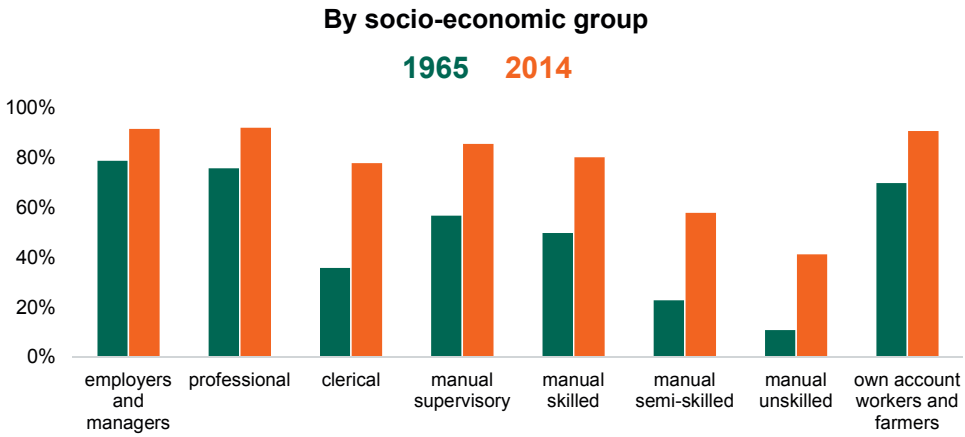


Figure 4. Change in driving licence holding, 1965 to now. (Note: The socio-economic grouping used here replicates the scale used in the publication of the 1965 results. The socio-economic position generally lowers when moving to the right of the graph)

The key factor affecting trends in travel over the past 50 years is the growth in access to cars. If we look at the share of the population that holds driving licences now compared to the past (Figure 4), we see that in 1965, just 10% of women aged 17 and above had a driving licence, compared to 50%

of men. Today, 67% of women hold a licence, compared to 80% of men. Driving licence holding has increased among women of all ages. For men, older age groups are now more likely to hold a licence than before, but for young males the numbers remain similar to 50 years ago. Driving licences have also spread more widely in lower socio-economic groups.

So there are now more people with access to cars, and household car ownership has increased. Indeed, it has become far more common for households to own several vehicles (Figure 5) – although, having a car, or several cars, is more frequent for higher socio-economic groups than it is for low socio-economic groups. The gap has narrowed over time, but disparities still exist.

How we travel

As car availability has grown, so car use has increased. Today, the car accounts for more than twice as much travel as it did in 1965. Use of other travel modes has decreased, particularly bus use, which has dropped by around two-thirds since the NTS began (Figure 6). Figures from later survey years show that walking trips have also declined, especially for short distances (under 1 mile).

All household types have seen an increase in car use, even those without direct access to a car (Figure 6). While the bus remains the most frequent mode of transport for people in no-car households (when walks are excluded), they are travelling more now as passengers in cars.

The increasing use of cars, along with the development of better road infrastructure, has meant that people can travel further and faster

The increasing use of cars, along with the development of better road infrastructure, has meant that people can travel further and faster, widening the range of opportunities available within a given travel time, leading to increasing trip distance. Again, this is particularly the case for women and older people, who travel much further on average than 50 years ago, whereas average mileage by young men was in fact lower in 2014 than in 1965 (Figure 7, page 40).

These trends are observed in other countries which have tracked travel behaviour over a similar period. For example, in France, the spread of car ownership, especially for lower socio-economic groups, has followed a similar trend in the last 30 years.[1] Similar

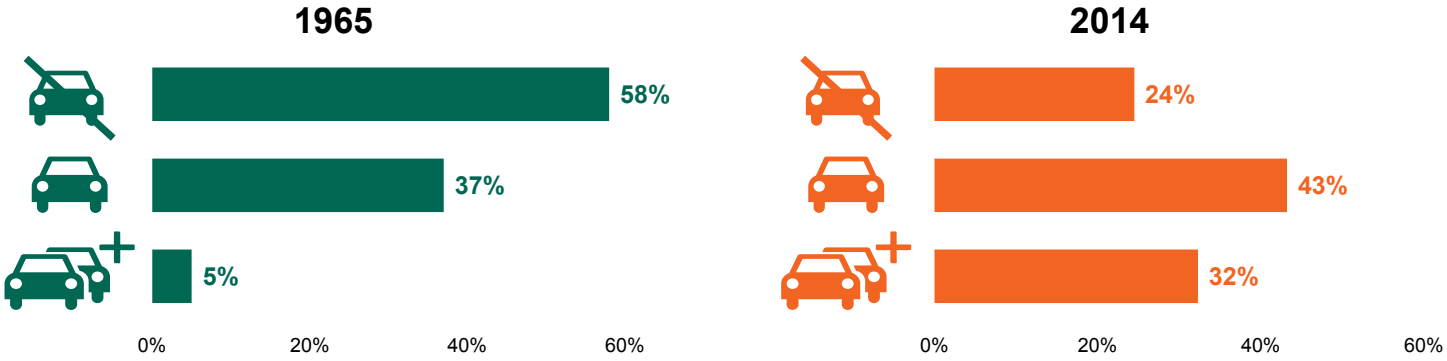


Figure 5. Change in car availability, 1965 to now

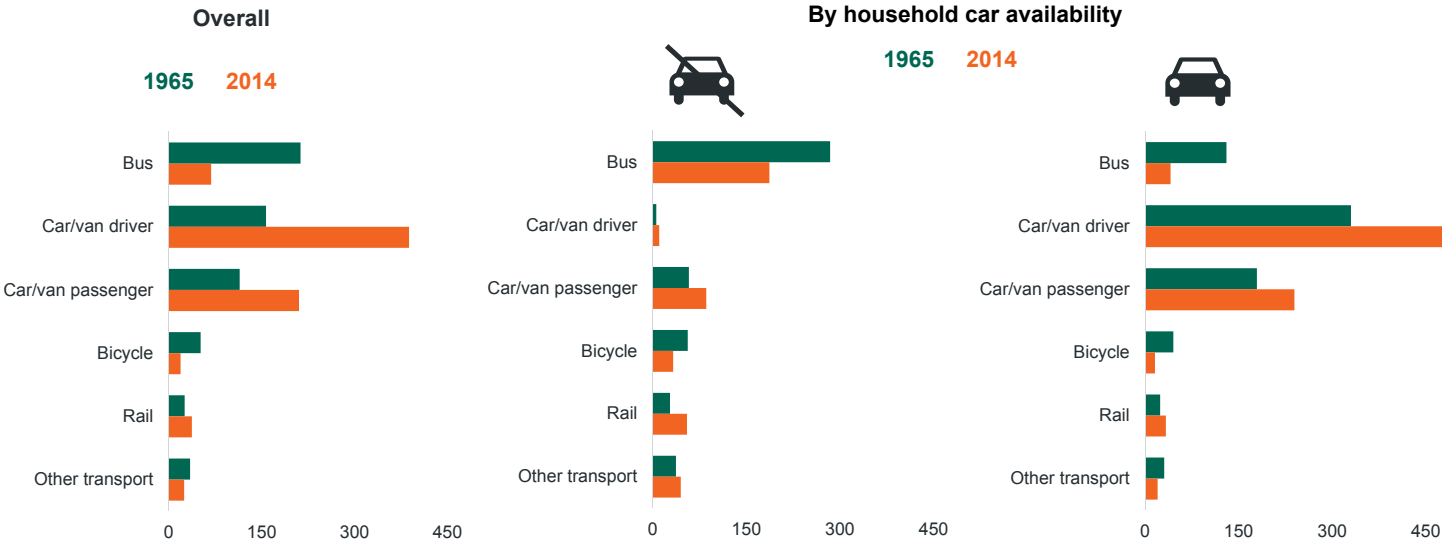


Figure 6. Stage mode of travel (excluding walks), 1965 and now. Figures for 1965 have been estimated from later survey data, but this was not possible for trip stages walked. Therefore, modal comparisons presented here exclude walks

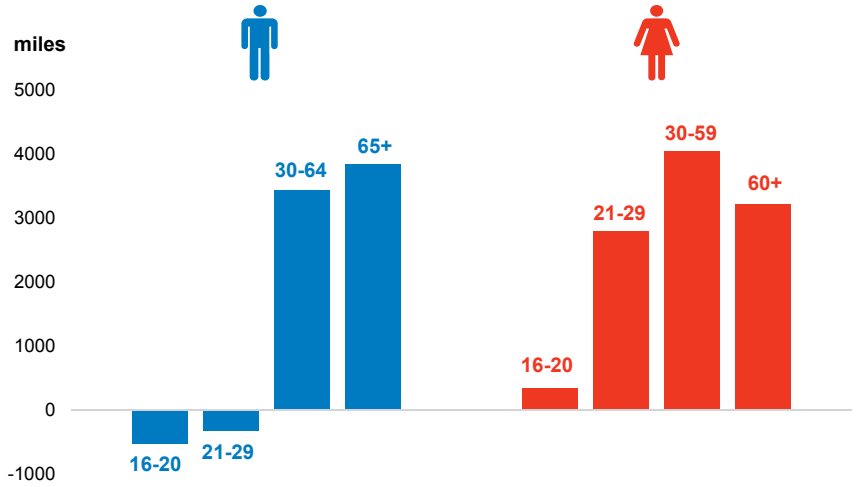


Figure 7. Change in distance travelled per person per year by age group and gender between 1965 and now

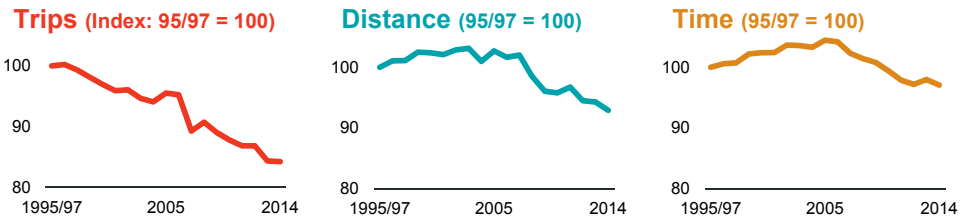


Figure 8. Recent trends in trips, distance and time travelled, 1995/97 to now

trends in distance travelled and increase of car availability have also been measured since the late 1960s by personal travel surveys in the United States.[2]

Changing trends?

While we might expect to see changes in travel patterns over the course of half a century, it is perhaps more important to note recent developments – particularly those of the last decade.[3] Overall trip rates appear to have peaked in the mid-1990s and to have fallen steadily since then, with a decline in walking and car driver trips in particular. The average distance travelled has stopped growing, which is influenced by trends in car travel (Figure 8). Conversely, public transport use – particularly rail – has increased, although it still accounts for only 3% of all trips made.

Understanding these trends matters because of their implications for the future. Trip rates are a fundamental part of the models and forecasts of the UK Department for Transport (DfT). In turn these models are crucial in the appraisal of transport schemes and investment, such as the recent £15 billion Road Investment Strategy. This is the reason why DfT continues to monitor travel trends and improve on the evidence base to better understand and attempt to explain the patterns observed, drawing on a range of sources including NTS data.[4]

For example, we see that the high costs of learning to drive and reduced employment rates have contributed to declining car use among the young, while changes in company car taxation have reduced mileage for higher-income, middle-aged groups. Meanwhile, mileage for women and older people continues to grow. Evidence in other areas is more limited – for example, growth in home-working appears to have reduced commuting trips, though the effect does not appear substantial, and the potential impact of growth in online shopping, use of social media (which could affect visits to friends) and other impacts through technology are not yet clear.

High costs of learning to drive and reduced employment rates have contributed to declining car use among the young

Urbanisation too is playing its part, with more people living in London and metropolitan areas where people are more likely to travel by public transport than drive. For example, 43% of households in London do not have access to a car today, compared to 24% of households living in urban areas and only 10% of households in rural areas.

Though drawing conclusions is not easy, the initial work suggests that, at the aggregate

Tom Shaw/Getty Images News/Thinkstock

NTS methodology: past, present and possible future

The National Travel Survey has collected data in essentially the same way since the first survey was carried out in 1965, through a face-to-face household interview and a 7-day travel diary completed by respondents of all ages. The survey was carried out infrequently until 1988, since when data collection has been continuous. Weighting for non-response was introduced from the mid-1990s, and from 2002 the annual sample size was increased substantially (from around 7000 to 16 000 individuals). The geographical scope of the survey changed in 2013 from residents of Great Britain to England only.

The purpose of the survey is to cover personal travel – where the reason for travelling is for the traveller to reach a destination. This excludes commercial travel (e.g. bus drivers), off-road journeys, and international travel, for example. The main unit of travel within the survey is a trip – a one-way course of travel with a single main purpose. Trips can consist of multiple stages, defined by a change in the form of transport. For instance, a journey to work could be counted as one trip (with a "commuting" purpose) but several stages (e.g. a cycling stage from home to a train station, and a train stage from the station to the workplace).

The NTS provides the longest and fullest time series of personal travel available internationally. However, like most data sources, the NTS has its limitations. For example, there have been some changes in definitions and scope over the years affecting comparability; most notably, the 1965 survey did not cover short walks (those under 1 mile), and these have been estimated from later data for the comparisons presented here.

There is a known drop-off in recording of trips by respondents throughout the travel week. A drop-off weight is used to adjust for this pattern. However, as short walks are only recorded on the last day of the diary (to minimise respondent burden), they are likely to be understated. Work is ongoing to explore the impact of changing the diary to record these on the first day, which is likely to result in a notable increase in the proportion of journeys made on foot, possibly necessitating some further weighting for historic data to maintain a comparable time series.

Of course, recent technological advances offer scope for reducing costs and respondent burden whilst improving data quality and modernising data collection methods. In 2011, a pilot was carried out to replace paper diaries with GPS devices, which 900 respondents carried for a week. Response rates for the GPS sample were acceptable (52%, compared with 59% for paper diaries) though younger respondents (12–24-year-olds) were found to be less likely to carry and return them.

The GPS devices included accelerometers – which detect relative motion in three planes – in an effort to improve the identification of mode of travel. An attempt was also made to code trip purpose from location of trip origin and destination. Ultimately comparison of results with diary data found that the processing algorithms for GPS data were not sufficiently well developed to produce robust data comparable with that collected via existing travel diaries, which have been retained for the current NTS. In particular, purpose was missing for around a quarter of trips, and there were issues with the accuracy of coding of mode, with too many trips classified as rail and too few walks.

As personal technology – such as smartphones and tablets – becomes more prevalent, exploring new data collection methods to ensure the survey remains relevant, and to maintain response rates for particular groups of respondents, is likely to be important. Ultimately, however, "big data" from these technologies – mobile phones in particular – may provide a better means of meeting the need for data to understand travel behaviour than the current survey-based approach.

level, growth in road traffic is likely to resume, as growth in incomes, declining driving costs (thanks to improved fuel efficiencies) and greater car accessibility push up car use. However, the DfT has – for the first time – used an extrapolation of the falling trip rate as a scenario in its most recent road traffic forecasts. It turns out that even in this case, the projected growth in population and commercial travel means traffic levels are expected to increase, although the timing and pattern of road demand do vary.[5]

The road ahead

Investment in transport and the growth of car use over the last 50 years have enabled us to travel further for the same purposes, widening employment and leisure opportunities, particularly for women and older people. But trends may be changing, and understanding the reasons and implications is crucial to ensure sound decisions regarding future investment. Good data and statistics are the cornerstone of transport decision-making.

The World Statistics Day is an occasion to celebrate the importance of data sources like the NTS, which provides an example of the value of maintaining a consistent, high-quality data source over time. It also reminds us of the need for constant improvement, especially in the context of new technologies changing both the way we travel and the way we collect data. Though it seems unlikely we will be collecting data in the same way in 2065 (see box), the need to understand travel behaviour is likely to remain crucial for transport planning and policy. Ultimately, it informs decision-making that has a real impact on the lives of millions of people.

References

1. Collet, R., Madre, J.-L. and Hivert, L. (2012) Diffusion de l'automobile en France: vers quels plafonds pour la motorisation et l'usage? *Économie et Statistique*, **457–458**, 123–139.
2. Santos, A., McGuckin, N., Nakamoto, H. Y., Gray, D. and Liss, S. (2011) *Summary of Travel Trends: 2009 National Household Travel Survey*. Washington, DC: US Department of Transportation.
3. Department for Transport (2015) *National Travel Survey: 2014*. bit.ly/1g5x1Ak
4. Department for Transport (2015) *Understanding the Drivers of Road Travel: Current Trends in and Factors behind Roads Use*. London: DfT. bit.ly/1AsqRYZ
5. Department for Transport (2015) *Road Traffic Forecasts 2015*. London: DfT. bit.ly/1JOaDxG

Delphine Robineau and **Matthew Tranter** are statisticians working in the National Travel Survey team of the Department for Transport in Great Britain. Any opinions expressed here are those of the authors and not necessarily those of the Department for Transport

Magical mirror

Improving foreign trade statistics in the Middle East

Trade between countries helps solidify relationships, but comparable statistics on exports and imports can be hard to come by. **Nili Karsai** and **Helen Brusilovsky** describe efforts to harmonise data in the Middle East using "mirror exercises"

Mickey Mouse stands in front of a magical mirror. Gazing at his reflection, he moves one way, but his mirror self does something entirely different. The poor cartoon mouse looks confused.

Foreign trade statisticians know how Mickey Mouse feels. In an ideal situation, country A's data on exports to country B should be equal to country B's data on imports from country A. That is, the person in front of the mirror and the figure reflected in the mirror should move in sync.

In real life, however, there are statistical discrepancies in official data collection which lead to confusing asymmetries. This was the situation statisticians in Israel, Egypt, Jordan and the Palestinian Authority (PA) faced.

Regional trade in the Middle East is important because common interests in trade help stabilise relationships between countries at the macro-economic level. This affects diplomacy, politics, and other fields, even if the trade volume is small.

The peace treaties between Israel and Egypt (1978) and Israel and Jordan (1994) led to the signing of trade agreements. Later, Qualifying Industrial Zone agreements (QIZs) were introduced – special free trade zones which allowed Egyptian and Jordanian goods to be exported directly to the US on the condition that Israeli inputs exceeded 10.5% of the product price. These trade agreements boosted the foreign trade of Egypt and Jordan, and increased trade between Israel and Egypt, and Israel and Jordan.

However, when comparing the import and export figures recorded by each country, discrepancies were evident. The statistical asymmetries found between Israel and its neighbours[1] – including the PA – are presented in Figure 1. In each graph dark columns represent data as reported by Israel. Lighter columns represent reflected data as reported by the partner country. In theory, each pair of columns should be equal; however, discrepancies – both large and small – can be seen in the data, with the greatest data asymmetry occurring between Israel and Egypt.

Cooperation needed

Discrepancies such as these should be reduced as much as possible, because accurate statistics facilitate actual trade, both for government and private sectors. Asymmetries in bilateral trade statistics can affect decision-making: for example, a business might get the wrong idea about the size of the market for its goods if statistics from country A and country B are different.

One way to detect asymmetries in bilateral trade statistics is to employ "mirror exercises", a methodological tool that is designed to reconcile trade data by identifying, explaining, and assessing the causes of discrepancies.

In mirror exercises, partner countries provide each other with very detailed statistical data based on the Harmonized System (HS), an internationally standardised system of names and numbers to classify traded products, which was developed by the World Customs Organization.

Next, the flow of exports from country A to country B is compared to the reflected flow in the imports received by country B, and vice versa. In theory the flows are supposed to be equal, but in reality discrepancies are found in either direction. In these cases, a data mining process starts, where statisticians attempt to identify the specific reasons for the gaps.

Mirror exercises were conducted between Israel, Egypt, Jordan, and the PA as part of the European Union's Mediterranean Statistical Cooperation Project, No. 3 (Medstat III). Group members – made up of statisticians from each of the four territories – met several times during the period 2009–2013. The main reasons for data discrepancies between Israel and its neighbours were as follows.[2]

© 2015 The Royal Statistical Society

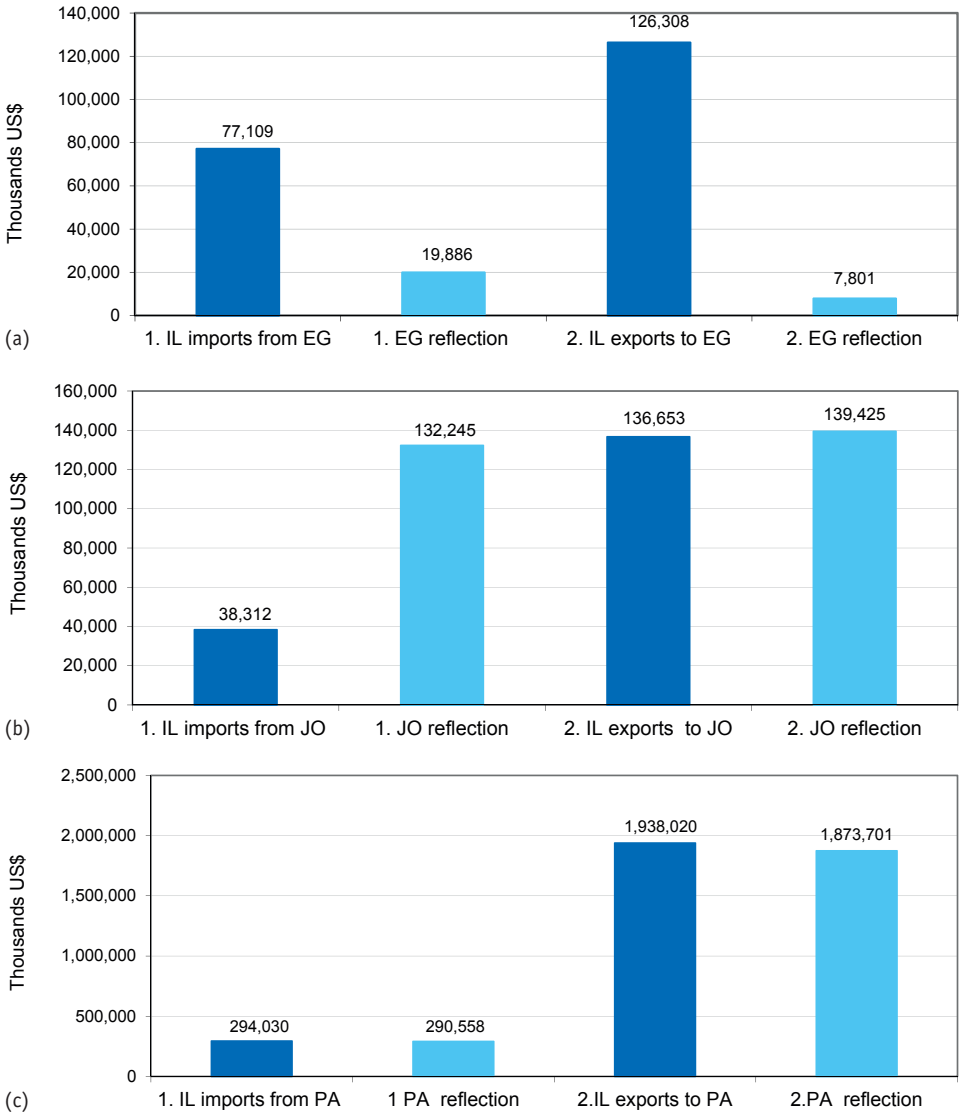


Figure 1. Data asymmetries between Israel and (a) Egypt, 2006; (b) Jordan, 2006; and (c) the Palestinian Authority, 2005

First, the four partners do not use the same trade system. For example, Israeli data on exports to Egypt were twice as high as the import records reflected in Egypt because the Israeli goods that were destined for the QIZ had not been included in the Egyptian statistics.

Second, the PA and Israel use VAT data to record their bilateral trade. In VAT vouchers, only a non-detailed description of the goods is available and it is difficult to obtain a detailed six-digit-level classification of goods. Therefore, Israel publishes the figures according to the economic activity classification of the importer/exporter firms, whereas the PA transforms the description in the voucher into HS codes.

Third, no records were registered in the data on Egyptian exports to match the data on Israeli imports of "natural gas in a gaseous state" and "petroleum coke".

Fourth, the records of Jordan's exports in "apparel and clothing" goods to Israel were much higher than the Israeli import records, because Israel registered only the value added of the goods in imports after processing, instead of registering the full value of the goods.

Finally, discrepancies between Israel and Jordan were due to the fact that Jordan was a transit country for Israeli gold and "waste of precious metals" goods, but Jordan was recorded as the final destination in Israeli records.

Reconciling the data

When comparing trade data between countries which allow for the use of different concepts and classifications,[3] there will be built-in discrepancies. But after months of collaborative work there were several improvements in the foreign trade statistics produced by Israel, Egypt, Jordan, and the PA. In particular, the mirror exercises led to better implementation of international standards, coverage, harmonisation and treatment of specific transactions.

Major asymmetries in Israeli and Egyptian data on fuel products were reconciled by analysing trade data from a new data source, while the classification gaps between Israel and the PA were narrowed by converting the PA's classification of goods into Israel's activity branch of the exporter/importer classification.

Discrepancies due to different concepts relating to goods in the QIZ zone were reconciled, and data on the final destination of exports were corrected.

The mirror exercises clearly showed that in order to produce quality data on bilateral trade, cooperation between statisticians on both sides is a necessity. And if better trade data do lead to deeper economic ties between the countries of the Middle East, the hope is that the end result is better, more prosperous lives for the people of this region.

References

1. Veronese, N. (2013) *Quality Report on External Trade Statistics in Southern & Eastern Mediterranean Countries*. Paris: EU Medstat III.
2. Tyrman, H. (2013) *Synthesis of the Asymmetry Studies on Trade between EU and the Southern and Eastern Mediterranean Countries*. Paris: EU Medstat III.
3. UN Department of Economic and Social Affairs (2011) *International Merchandise Trade Statistics: Concepts and Definitions* (Statistical Papers, Series M No. 52). New York: UN.

The Israel Central Bureau of Statistics (ICBS) web page on Foreign Trade Statistics can be found at bit.ly/1JOApP2

Nili Karsai works at the Israel Central Bureau of Statistics in the International Relations and Statistical Coordination Department as a senior supervisor. She has a degree in Business Administration and an MBA from the Hebrew University in Jerusalem

Helen Brusilovsky works at the Israel Central Bureau of Statistics in the Macro-Economics Department. She is the director of the Foreign Trade Sector. She has a degree in Statistics and Geography and an MA in Demography from the Hebrew University in Jerusalem