7. Setting scenarios: combining numbers and stories

Michaela Potančoková, Helga de Valk, Rafael Costa, Michaël Boissonneault and Jakub Bijak

7.1 QUANTITATIVE VERSUS QUALITATIVE SCENARIOS

Many attempts to create future migration scenarios have already been made, especially in the European context and with a spotlight on immigration (see e.g. Lutz et al. 2019; Acostamadiedo et al. 2020; Sohst et al. 2020). In a recent systematic review, we examined 107 migration scenario studies, nearly half of which focused on Europe (Boissonneault et al. 2020). The literature included in the review covered both *quantitative* studies (44), with migration flows described in terms of possible trajectories of future numbers, and *qualitative* studies (30), concentrated on narratives about the direction of change and interplay of migration with a wide range of drivers. Additionally, a sizeable number of studies (33) had tried to combine the advantages of both approaches within a *mixed* framework, of which Wiśniowski et al. (2021) is a more recent example. This was especially the case for studies relying on expert opinion both for setting the trajectories and providing the underlying narrative explanations.

In that review, once we looked at the focus and purpose of setting migration scenarios, an interesting picture emerged. By focus, we mean whether the scenarios are set specifically for migration or for other processes (e.g. population or the economy), with migration just one of the contributing variables. As for purposes, we distinguish *predictive* studies, with scenarios used as forecasting tools, *explorative* ones, both exploring and explaining possible futures, and *normative* studies, trying to answer the question about the desired levels of migration required to meet some social or economic objectives. In Boissonneault et al. (2020), we offered a simple typology of existing studies and approaches, across these two dimensions. One interesting finding was that among migration-focused scenarios, explorative studies ($n = 26$) clearly outnumbered predictive ones (8). When migration was just an input to scenarios for other variables, the studies were undertaken mainly for predictive (38) rather than explorative (29) or normative (6) purposes. As a possible explanation, this difference may indicate that the migration research community is rightly cautious around using migration scenarios as predictive tools, being aware of the associated difficulties, as discussed in Part I.

Existing studies, especially on the numerical side, have explored a wide range of methods for generating and quantifying scenario assumptions. Aside from time-series methods, as discussed separately in Chapter 6, these approaches range from Delphi methods (e.g. Drbohlav 1997; Acostamadiedo et al. 2020; Wiśniowski et al. 2021) to proper foresight studies (e.g. Foresight 2011; Vezzoli et al. 2017). In addition, some scenarios are conditional on established assumptions and narratives about drivers, such as the *shared socioeconomic pathways* (O'Neill et al. 2017), which are reflected in a number of migration studies (e.g. Abel 2018; Lutz et al. 2019), or gravity-type models (Rikani and Schewe 2021). Delphi-type surveys, in particular, attempt to combine the advantages of quantitative and qualitative methods by eliciting the numbers within a broader deliberative process. Still, despite covering a lot of ground in terms of the methodology and practice of migration scenario-setting, existing methods collectively leave a few important gaps.

The first gap is that the scenarios are very often set in terms of high-level aggregates (e.g. total immigration or emigration). Worse, the most popular variable in the review was *net migration*, which does not correspond to *any* actual migration process, being simply an arithmetical difference between the numbers of immigrants and emigrants (Rogers 1990; see also Chapter 10). The second, related, gap is that the high-level aggregates for which scenarios are set rarely reflect the diversity of flows, which have different characteristics and levels of uncertainty, as discussed in Chapter 4 (see also de Beer 2008; Bijak et al. 2019). The third gap is in the quantification of assumptions, which is fraught with many pitfalls, whether rooted in analysis of individual drivers or expert opinion. This particularly involves the ways in which scenario uncertainty is described or quantified, if at all (see discussion in Chapter 6). These gaps call for the adoption of a different approach; we present two attempts at this – a factorial experiment and microsimulations – in the remainder of this chapter.

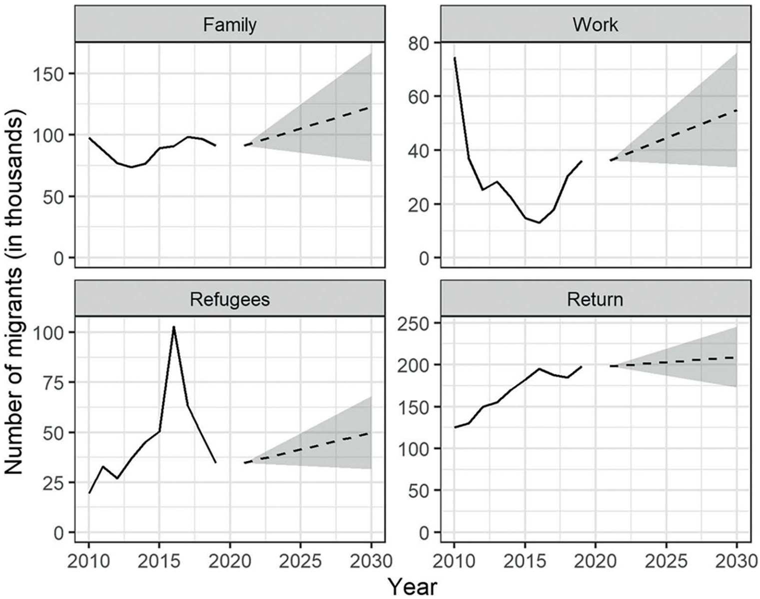
7.2 COMBINING SCENARIO-BUILDING APPROACHES: A FACTORIAL SURVEY EXPERIMENT

To combine the strengths of existing qualitative and quantitative approaches, we carried out a *factorial survey experiment* of migration researchers and other professionals, based on *vignettes* describing origin and destination regions (Boissonneault et al. 2022; Boissonneault and Costa 2023). The survey experiment involved asking invited migration experts to fill in an online survey on the expected future of selected migration flows by 2030, relative to benchmark data for the 2010s. The flows involved migration from Middle East and North Africa (MENA) to Europe and were disaggregated into four main types: family, work, asylum (refugees) and return migration. For each flow, we looked at a range of drivers, in both the origin and destination, following the work presented in Chapter 4. The choice of specific drivers was also inspired by broader theoretical views on social transformation as a mechanism driving migration (de Haas et al. 2020; de Haas 2023). The specific drivers included: (1) pace of population ageing; (2) shifts in attitudes; (3) policy changes in Europe; (4) changes in the size of the most mobile populations; (5) shifts in social stability in MENA countries; (6) shifts in political stability in MENA countries; and (7) convergence in economic conditions between the two regions.

In the next step, for each of these seven drivers, we assumed two *levels*: either increasing or decreasing values (e.g. decreasing pace of ageing in Europe, increasing political stability in MENA). Each of the seven drivers could thus be in one of the two possible states (increasing or decreasing), resulting in $2^7 = 128$ possible combinations. This broadly follows the argumentation about drivers given in Czaika and Reinprecht (2022) and reflects the fact that the possible changes in drivers can be described qualitatively. We referred to each combination as a *vignette*, describing one of the possible worlds in terms of the driver profiles.[1] That

[1] For a general description of the vignette methodology, see Atzmüller and Steiner (2010). An existing example of a migration application is the study by de Jong and Fonseca (2020).

our survey was *factorial* means that we used all possible combinations (vignettes) in our study, with each of the 138 expert participants being presented four of them.[2] Finally, the experts were asked to indicate, for each vignette, by how much they expected each of the four migration flows in 2030 (family, work, asylum and return) to change from the 2010s baseline. Aggregate results for all experts are summarized in Figure 7.1. On the whole, participants' answers varied between flows remaining broadly the same and doubling during the 2020s.



Note: Dashed lines: mean expert assessment of future trends; shaded areas: average range between the lowest and the highest migration scenarios.
Source: Boissonneault and Costa (2023, Figure 2), reproduced by kind permission of the authors. Past data from Eurostat.

Figure 7.1 Past trends (2010–19) and expert assessment of plausible future trajectories (2020–30) of main migration flows from MENA countries to Europe

2 For a discussion of the methodology of factorial experiments, see e.g. Auspurg and Hinz (2014). One related application of such a method, in the area of healthcare, can be found in Sheringham et al. (2021).

Our experiment aimed to create scenarios that were more in line with current theoretical thinking about the role of migration drivers and social change (de Haas et al. 2020; Czaika and Reinprecht 2022; de Haas 2023), while remaining within the broad confines of methodology of scenario-setting. For a single group of migration flows, from MENA to Europe, our experiment was successful, producing expert-based scenarios and engaging the participants in the process. We were also able to produce scenarios (with uncertainty assessment additionally offered by the experts) for flows disaggregated by main type (family, work, asylum and returns). If this approach could be implemented for other regions of the world, it would offer a very appealing alternative way of using expert opinion in migration scenarios. Such an approach would go beyond the state-of-the-art Delphi approaches (e.g. Acostamadiedo et al. 2020).

Unfortunately, when trying to scale up the experiment to different regions of migrant origin and for additional flows from other world regions, we encountered an important practical barrier. The experiment was well suited to a single flow but was very resource intensive and would require additional elicitation, or at least evaluation, expert time and knowledge, to derive similar parameters for additional migration flows. Because of the costs in terms of setting up and carrying out the experiment itself, we found that implementing this method for all possible combinations of flows was impractical. Further, reducing complexity of migration drivers into a few numerical parameters that can be scaled in scenarios to produce divergent migration outcomes was still deemed as too simplistic by some experts. For these reasons, we focused instead on addressing uncertainty arising from divergent migration behaviours. To that end, we decided to use microsimulation-based methods, which offer flexible tools for capturing heterogenous behaviours and modelling diverse populations. We present our solution in the remainder of this chapter.

7.3 SCENARIO-BUILDING THROUGH MODELLING: THE POWER OF MICROSIMULATIONS

As mentioned earlier, migration scenarios are most often set at the very aggregated level of overall immigration and overall emigration flows. However, individuals and groups have diverse migration aspirations and behaviours. Structural changes in populations, such as e.g. population ageing of entire populations but also of settled immigrants, and *cohort* *replacement*[3] of more homogeneous cohorts by diverse ones, result in increased population heterogeneity. This heterogeneity can manifest itself especially when we look at population groups in terms of birthplace (nativity) or origin country. The resulting diversity in population composition impacts migration dynamics, as different groups have different propensities to migrate. For scenarios this is particularly important, because migration has become the key driver of population change in Europe in the 21st century. Migration scenarios are an essential part of population projection models and should be linked to demographic change and consider multiple sources of heterogeneity.

The diversity in migration processes also affects uncertainty in future migration via changing population characteristics. Expert opinion may implicitly consider this impact by providing an informed guess as to altered future population composition, but multistate demographic modelling can project or simulate these changes more precisely. Dynamic microsimulation methods (e.g. Harding et al. 2009; Bélanger and Sabourin 2017) and agent-based models (e.g. Gilbert and Troitzsch 2005; Bijak et al. 2021b) are especially well suited to flexibly incorporating multiple socio-demographic characteristics of different actors. Especially microsimulation offers a way of handling multiple dimensions linked to the many sources of heterogeneity more flexibly and can provide better projection results than conventional population projection methods (see Chapter 10). Microsimulation is also better suited to modelling life courses and links between actors. This makes it easier to handle inter-generational transmission of some characteristics, such as education, or incorporate interaction between duration of stay of immigrants in the country and their childbearing or labour force outcomes.

Microsimulations are based on data on the frequency of people's *transitions* between the different *states* they can find themselves in.[4]

---

3 This is also known as *population renewal* or *demographic metabolism* (Lutz 2013). In essence, as time goes by, older cohorts of people (here, groups born in the same year) are replaced by younger ones, as a result of an interplay of three demographic processes: *births* create new cohorts, *deaths* affect the survival of older cohorts throughout their lives, and these dynamics are influenced by *migration*. This framework can be extended to different *states* (e.g. country of residence, labour market activity status, educational level).

4 For example, consider a 44-year-old woman, highly educated and economically active, residing in country A (state 1 = higher education, active,

Such models are particularly well tailored to simulating future population changes, which are behind the observed macroscopic, population-level demographic process of cohort replacement. Such models allow assessment of the demographic and labour force impacts of migration (e.g. if the models are set to model future labour supply). Including different migration behaviours and intensities, for different population groups and at different stages of the life course, is a feasible way of reducing epistemic uncertainty (see Chapter 4).

Migration theory and empirical analyses of migration patterns provide guidance to modelling this heterogeneity. Selectivity in migration arises though differentiated aspirations and capabilities to migrate. First, age is the most important predictor, as migration is linked to life-course changes and peaks in the 'rush hour of life' – young adulthood – when many important life transitions happen (Rogers and Castro 1981; Courgeau 1985; Bernard et al. 2014). Empirical data on migration by age, including our estimates presented in Chapter 5, clearly show that migration peaks around age 25 for most flows concerning Europe. Second, native-born individuals usually have a lower propensity to migrate internationally than the foreign-born, who may remigrate or return to their country of origin (see Chapter 4). Among immigrants, emigration (including returns) peaks in the first years after arrival. Duration of stay in a country is therefore a very important predictor of migration, alongside age and nativity status (van Hook and Zhang 2011).

Educational aspirations are also an important migration driver (Chapter 4), as is labour market activity. It is well established that people with above-average educational attainment are more likely to migrate internationally (Dustmann and Glitz 2011), and for labour migration, the higher propensity to move for those who are economically active is almost tautological. Including at least some of these differentials – not only age, but also education level and economic activity – is an important step towards causal modelling of the migration process. To that end, considering multiple sources of population diversity indirectly supports the integration of causal migration modelling into population projections, by approximating the underlying mechanisms driving migration flows

---

country A). If she moves to country B to take up a new job, she is making a transition to another state (state 2 = higher education, active, country B). The intensity of such transitions is described by migration rates from A to B for 44-year-old, highly educated and economically active women.

(Willekens 2018). Specifically, it can be argued that mechanism-based migration forecasting must account for the heterogeneity of cohorts and changes in personal attributes over the life course (ibid.).

In our microsimulation model, called QuantMig-Mic, we simultaneously model populations of the EU+ countries (excluding Liechtenstein, due to problems with data availability), with emigration rates varying by nativity (region of birth) availability and age. Changes in the age structure and the composition of each cohort by region of birth are subsequently translated into simulated emigration flows. The destination is assigned based on rules driven by time series of past migration flows. In this way we capture the fact that European-born migrants are more mobile within Europe than people born outside the EU+. For the last group, for example, migration involves moving to other parts of the world more often than for the European born, and it often includes return migration.

Differences in emigration rates by birthplace need not be fixed: we can project assumptions on the future dynamics of these differences according to evidence from past data and insight from relevant migration theories and spatial regularities. In the microsimulation, immigrants enter the model as new actors, with immigration linked to age structures in origin countries and changes in educational compositions of immigrant populations (Marois et al. 2023). This results in changes in the composition of projected immigration flows in terms of region of birth, as immigrants are attracted to destinations with established migration links and networks. In this way, by following past trends in various transition rates, including migration rates, we can obtain the reference scenario – the *baseline* – which we then use for the analysis presented next.

Improved data on emigration in origin countries, availability of origin–destination migration data by socio-economic status, and data organized by immigration cohorts rather than by calendar years can further reduce the epistemic uncertainty in future migration scenarios. Granular emigration data by birthplace and duration of stay would allow us to better model the diversity and life courses of foreign-born populations. This could help equip migration models and scenarios with some causal elements. In the context of microsimulations and agent-based models especially, this would enable the links between people (in families or households) and the ways in which they make migration decisions to be explored (see e.g. Klabunde et al. 2017; Willekens 2018). More granular migration data could therefore help with utilizing the power of simulation tools more fully.

7.4 ADDRESSING THE ALEATORY: INCLUDING UNCERTAINTY IN SCENARIOS

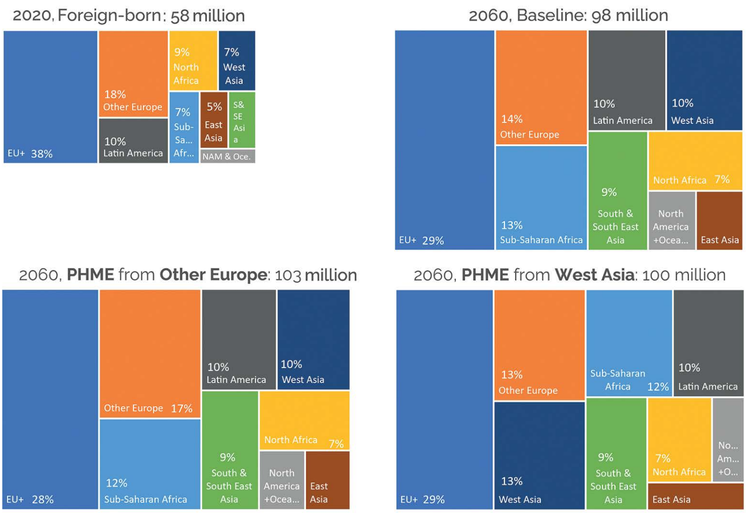
In population projections, future migration uncertainty can be accounted for by using either probabilistic methods (Bijak 2010; Azose and Raftery 2015) or a range of alternative scenarios. Data impose important limitations on how much epistemic uncertainty can be reduced by incorporating nuanced understanding of migration through causal modelling. As illustrated in Chapter 6, there are limits to mitigating epistemic uncertainty over long-term horizons. Driver-based scenarios, commonly used in migration forecasting, are by definition epistemic, as they try to make sense of macroscopic regularities and correlations between migration and its drivers. But the longer the time horizon, the greater the role of aleatory uncertainty (Chapter 2). In other words, whatever we think migration will be like decades from now is most likely incorrect. This is due to the unpredictability of micro-level migration decisions and aspirations and even more so the unpredictability of macro-level events that may impact migration (including wars and natural catastrophes but also large economic or geopolitical shifts). These factors and their causes cannot easily be distilled to a set of drivers.

To address the aleatory randomness of future migration, we propose a novel framework based on the modelling of rare migration events for long-term time horizons, introduced in Chapter 6. We start with the estimated magnitudes of rare migration events from different world regions presented in Table 6.1, for once-in-a-decade and twice-in-a-century high-migration events related to immigration into Europe. In the next step, based on these assumptions, for each immigrant origin we build four sets of scenarios for migration events that combine the magnitude of the event and the duration (persistence) of the effect. Of course, as mentioned earlier, we chose immigration for illustrative purposes and as a current policy priority in the EU, but the method would work equally well for any other migration flow of interest.

Our proposed solution approximates the inherent volatility of migration due to unpredictable events (e.g. wars, natural disasters). Such events – sometimes referred to as ‘shocks’, although we prefer to refer to *high-migration events*, as a more neutral term – cause disturbances to the migration system and shift the trends. Alternative scenarios can then simulate such hypothetical high-impact migration events to better understand their demographic and labour market impacts (Potančoková et al. 2023). In our case, we simulated two contrasting scenarios. The first is a *transient high-migration event*, where after one year migration returns to the projected trend shown in the baseline scenario. This could reflect an adequate policy response that mitigates the impact or simply the driving forces ceasing to exist. Transient high-migration events even of twice-in-a-century magnitude would not be expected to leave a lasting imprint on the projected population sizes, labour force or share of foreign-born population. To put the simulated numbers into perspective, the estimated twice-in-a-century high-migration event from West Asia into EU+ would roughly correspond to the 2015–16 migration wave from Syria and Iraq (see Table 6.1 in Chapter 6).

The second situation reflects a prolonged crisis and chain migration: after the initial high-migration event, immigration from a given region remains elevated and returns only gradually to the volumes envisaged in the baseline scenario. We refer to this as a *persistent high-migration event* (PHME). This persistence in migration after the initial event lasts for a decade. Such a situation is not unusual, as family reunifications and migrant networks incentivize the perpetuation of migration from the same origin (de Haas 2010). In fact, migrant networks and past migration corridors are used in the simulation to approximate the choice of future destinations. Therefore, only PHMEs from the origins previously linked to a particular European country are assumed to impact population size and composition in that country, as illustrated by Figure 7.2. We model the alternative scenarios sequentially, with high-migration events introduced at the same time in all scenarios (from mid-2025 to mid-2030). In other words, we are asking a question: What would be the population and labour market impacts of high-migration events in the late 2020s?

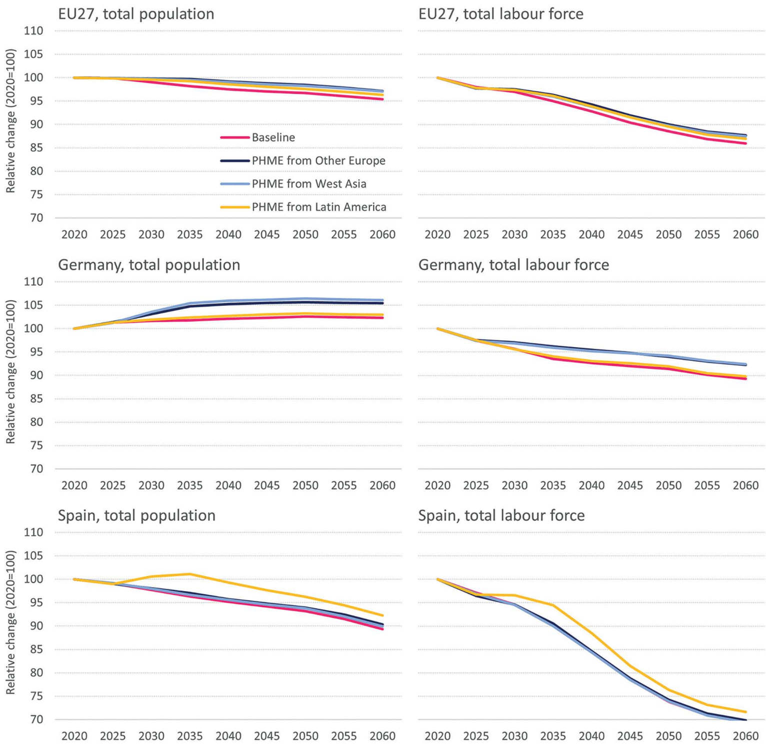
Among the two situations, PHMEs can clearly be expected to leave a larger imprint on the make-up of European populations than transient migration events. If migration into Europe continued at the same rate as in the past, the share of foreign-born population in the EU27 would increase from 13 per cent to a projected 24 per cent in 2060. The share of migrants born in EU+ countries would be smaller (Figure 7.2, upper boxes), while the foreign-born population would, of course, be larger under alternative migration scenarios (Figure 7.2, lower boxes). A PHME from Other Europe would increase the total foreign-born population in the EU27 by 5 million by 2060 in comparison with the baseline scenario, while the share of people born in Other Europe would be similar to their share in the 2020 stock. Over the same horizon, a PHME from West Asia would result in a foreign-born population larger by 2 million compared with the baseline, with the share of foreign born nearly doubling. Still, given the importance of natural demographic dynamics (births and deaths), the long-term impacts even of PHMEs on population size and labour force are relatively limited. This confirms earlier findings (UN 2000; Bijak et al. 2008) that in spite of some diversity in trends, even seemingly large migration flows only temporarily offset the projected impact of population change, particularly ageing, as shown for the EU27, Germany and Spain in Figure 7.3. All results are available from the QuantMig Migration Scenarios Explorer, at https://bit.ly/quantmig-scenarios.



Note: Baseline and two selected persistent high-migration event (PHME) scenarios shown. The box areas are proportional to population sizes.
Source: Eurostat (for 2020); own elaboration (for 2060).

Figure 7.2 Estimated (2020) and projected (2060) distributions of the origin of the foreign-born population in the EU27

From the point of view of causal migration modelling, the proposed framework is even better suited to emigration than immigration (where PHMEs can be attributed to the countries of origin) and can also be adapted to bilateral flow forecasts. The latter can be achieved by combining the estimated volume of high-magnitude emigration – number of migrants from a country or region of origin – with decision rules



Note: PHME = persistent high-migration event related to immigration to Europe.
Source: Own elaboration.

Figure 7.3 Projected total population and labour force for selected countries and scenarios, 2020–60

regarding the absorption of the flows to different destinations. A particularly appealing possibility would be to apply this approach to climate change and environment-related migration. Linking its scenarios to the relative frequency of events could explicitly take into account caveats about the uncertain character of such mobility, avoiding unnecessary alarmism (see e.g. Foresight 2011, de Haas 2023, and discussion in Chapter 4). This would necessitate not only more nuanced emigration data but a better understanding of different situations and strategies (e.g. involuntary immobility and adaptation capacity; short-term and short-distance mobility vis-à-vis long-term and long-distance international emigration).

Our scenarios have been prepared for Europe, but with improving data availability the same process could be applied to any region of the world. To that end, improved migration data for countries of origin and advances in modelling can also reduce some of the epistemic uncertainty for environment- and conflict-related drivers of migration. However, for climate change and conflict, especially over longer time horizons, much of the uncertainty will remain aleatory and therefore irreducible. This shifts the policy focus back to ensuring adequate preparedness. We continue this discussion in Part V, but first, in Part IV, we make a foray into discussing effective science communication and use by policymakers: a necessary prerequisite of any attempts to prepare for the unexpected.