International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report IR-09-037

Toward a Systematic, Argument-Based Approach to Defining Assumptions for Population Projections

Wolfgang Lutz (lutz@iiasa.ac.at)

Approved by

Detlof von Winterfeldt (detlof@iiasa.ac.at)
Director

October 13, 2009

International Institute for Applied Systems Analysis
Registration number: ZVR 524808900

Contents

Abstract

This paper proposes a new approach for defining the assumptions in population projections. It is based on a broad discussion and argumentation exercise which critically assesses the science basis of alternative arguments that relate to the forces which jointly shape the future trends in fertility, mortality and migration. It starts by discussing the results of a recent Eurostat survey among all EU national statistical offices in which there appears a clear consensus that the current practice of defining assumptions for population projections should be improved through stronger interactions with the scientific community. As an instrument for such interactions, the paper goes on to propose an interactive spreadsheet in which more than 100 arguments relating to the forces shaping future fertility, mortality and migration have been formulated. Invited experts then evaluate these arguments with respect to their validity and their potential impacts on the force under consideration. This interactive questionnaire has already been used to define the assumptions in the official UK population projections and with an international group of 17 mortality experts. With further fine tuning, this approach can be used to collect and synthesize the expertise of large numbers of experts for defining median assumptions and the associated ranges of uncertainty. It has the potential to become a standard tool for the future production of national and international population projections.

Acknowledgments

Many thanks to Chris Wilson, Isolde Prommer, Chris Shaw and Pertti Saariluoma for their valuable input.

Funding for this work was made possible by the European Commission 6th Framework Project "Bridging the Micro-Macro Gap in Population Forecasting" (MicMac), Contract No. SP23-CT-2005-006637. This paper will eventually become a chapter in the final book for this project.

About the Author

Wolfgang Lutz is Leader of the World Population Program at IIASA, Director of the Vienna Institute of Demography of the Austrian Academy of Sciences, Professor of Social and Economic Statistics at the Vienna University of Economics and Business, and Professorial Research Fellow at Oxford University, UK.

Toward a Systematic, Argument-Based Approach to Defining Assumptions for Population Projections

Wolfgang Lutz

1 Introduction

In times of accelerating social, economic and environmental changes, science-based information about the future becomes ever more important. The demand for such authoritative information often exceeds the supply of credible projections. There are many dimensions along which projections could be doing better than is the current practice. In the field of population projections, such improvements could be essentially along three lines:

(a) **Offering more detail** about the stratification of the population along dimensions other than only age and sex. The European Commission's 6th Framework Project "Bridging the Micro-Macro Gap in Population Forecasting" (MicMac) has pioneered approaches which explicitly consider educational status, health status and household status in addition to the traditional dimensions of age and sex. These additional dimensions are assumed to cover important sources of heterogeneity for demographic behaviour which implies that their explicit consideration in projections will both improve the accuracy of the projections and provide more detailed information that is of direct use for planning and other purposes.

(b) **Providing more information about uncertainty**. While most users of projections are primarily interested in a best guess forecast, there is increasing demand for explicit consideration of the full range of uncertainty. Traditionally, this demand has been met by producing alternative "variants" or "scenarios" which are supposed to cover some "plausible" range of future trends. More recently, however, planning agencies explicitly demand probabilistic population projections in order to have a "demographic risk function" to be matched to their cost function. In terms of future pension entitlements, for instance, a very minor deviation from the base line projection of the proportion of the population above age 65 in 2030 implies additional government expenses or savings in the billions of Euros. This is also the reason why the UK Ministry of Finance recently explicitly requested a probabilistic population projection from the UK Office of National Statistics in order to enter it into its main economic model.

(c) **Provide a better substantive justification of the assumptions made**. The results of projections crucially depend on the specific assumptions made. As the survey of statistical agencies discussed below will describe, these assumptions are typically defined within those agencies after consultation with the experts. This is usually done in a rather informal way and the resulting choices tend to reflect "expert opinion" rather than transparent science-based reasoning which some of the users might expect. All of the statistical agencies state that this particular aspect of producing projections is the one most in need of further improvements.

This paper will primarily address the third issue concerned with the process of defining assumptions. But inevitably such discussions also touch upon the issues of uncertainty and that of population heterogeneity.

The paper begins with a summary of the results of a survey among all EU national statistical agencies about the current practice in population projections and their views of desirable further improvements. We will then step back and discuss some more theoretical meta-scientific points about the common fallacies associated with expert opinion and frequently stated arguments, and the possible science-based solutions. We continue by presenting a new argument-based questionnaire that was used in the most recent population projections of the UK and summarize the results of a more extensive exercise along these lines carried out in the context of a MicMac workshop on future mortality trends. The paper will conclude with a discussion of what we have learned and an outlook to the future of population projections.

2 Survey of Current Practices in National Statistical Agencies in the EU

Virtually all national statistical agencies in the world as well as inter-governmental agencies such as the United Nations and Eurostat have been producing regular population projections by age and sex following the so-called cohort-component projection method. The arithmetic of making cohort-component population projections has essentially remained unchanged since Edwin Cannan proposed it in 1895. However, the way in which demographic forecasters make their assumptions concerning future trends in fertility, mortality and migration is always potentially changing. This section reviews the current practice of statistical agencies in Europe (EU-25) for defining the fertility, mortality and migration assumptions in population projections, and draws preliminary conclusions from the review. Perhaps the most significant finding is that all national statistical offices that replied to a questionnaire on this topic agreed that there is a need for improvement in the methods used to make assumptions. In particular, the offices charged with making population projections would welcome more structured interactions with the demographic research community.

One of the first steps of the European Commission's MicMac project was to collect information on the current use of external experts in defining fertility, mortality and migration assumptions. The national statistical offices (NSOs) of the European Union countries were asked to provide information on what had been done during the production of the *most recent* population projections. Each office received a questionnaire from Eurostat – designed by the International Institute for Applied Systems Analysis (IIASA) – and 21 out of 25 national statistical offices returned the completed questionnaire by late November 2005. In addition to pre-set answers to the 12 questions, the questionnaire provided space for open-ended comments on each question. A complete account of these answers and comments has been provided by Prommer and Wilson (2006).

The aim of the questionnaire was to assess the current status of expert involvement and methodology in making population forecasts by the national offices. The second aim of the questionnaire was to evaluate what future improvements could be made in the process by which experts contribute to the definition of assumptions in population projections. The results of the survey are presented in Appendix Table 1.

The majority of the offices indicated that they generally use three future pathways for fertility, mortality and migration; some used only one or two. Only one office used stochastic methods to project future population. The most common approach is to create scenarios that cover a “plausible” range. The involvement of external experts and meetings are clearly important, but there is a marked gap between “old-15” who tend to have more elaborate ways of involving experts and “new-10” member states. If there are problems in finding a consensus on values, most offices make in-house decisions after consulting the experts. Three offices commissioned scientific studies from outside experts for the explicit purpose of helping with the definition of assumptions (though one of those does not publish its own forecasts).

Error analysis of past assumptions is also important. The decision to carry out either a systematic or a more qualitative analysis of past errors splits the respondents into two groups. However, no statistical office provided a description on the methodology they use for error analysis. Half of the respondents define storylines (either combined for the three components of change or for each component separately) behind the assumptions. The other offices do not discuss storylines.

All national statistical offices agreed that there is a need for improvement in the methods used to make assumptions. Generally speaking, improvements in networking and in advancing the conventional methodology of scenario-based forecasts seem to have priority. The introduction of stochastic/probabilistic forecasting methods is not a high priority for most offices. It is not clear, however, if the slow pace of adoption of stochastic/probabilistic forecasts has to do with scientific criteria, or is simply due to the non-availability of human resources with the appropriate knowledge of the methodology. But there is evidence of lack of human recourses as three offices explicitly stated in question 12.

In the following we will have a closer look at the answers to some of the key questions. Anonymity of the specific quotes given is maintained throughout the text.

2.1 Number of alternative future paths defined for fertility, mortality, and migration

The majority of the statistical offices use the “scenario” design for the definition of the assumptions of the future paths of fertility, mortality, and migration. That means that the forecasters project the future population by defining at least one development path of future fertility, mortality and migration. This is commonly known as the “best-guess” or “most-likely” development. Roughly half of the central statistical offices prepare three different paths for each variable: fertility (52 percent), mortality (52 percent) and migration (43 percent). The second most common situation is that the NSOs use only one future path of each indicator; this is the case for five statistical offices. Two NSOs generated two different future developments for fertility, two NSOs for mortality, and four for migration. Generally speaking, the 3-3-3 version is most common, but there are some NSOs that use the 3-2-2, or the 3-1-1 composition. Only one statistical office identifies more than three possible future developments of mortality and migration, while two offices do so for fertility. Using stochastic population projections methods is still rare among the central statistical offices; only one institute stated that they use stochastic forecasting. Finally, one national statistical office replied that it did not officially publish projections and therefore uses the projections published by Eurostat.

The open-ended comments give us more insight how these assumptions are used in the projections. One statistical office that answered with a 1-1-1 assumption “produces demographic trend-calculation projection by using the so-called demographic component model, and does this for all municipalities”. One office made national and regional projections for the “low-central and high hypothesis” in their latest projections. For another, the latest projections "were centred on the ventilation of those perspectives at the NUTS-3 level and calibrated them only on the former central hypothesis". Another office uses the 1-1-1 assumption hypothesis for short-term projections, but included two migration scenarios for long-term projections. Another NSO published the Eurostat baseline projection as the national variant, therefore, the answers in this questionnaire refer only to the procedures that took place during the discussion and harmonization of the assumptions for the three variants of population projections regarding fertility, mortality and migration for 2004 to 2050. One NSO that answered that they use the "standard" three variants (high, principal, low) also produce projections of a few special case scenarios, e.g., replacement fertility, no mortality improvement and zero migration.

2.2 Interpretation of the range defined by alternative assumptions

Thirteen national offices stated that the assumptions described cover a "plausible range". The office that uses the stochastic forecast methodology specifies intervals (usually 95 percent intervals) together with the type of distribution, etc. Two NSOs did not make any specification in the meaning of the alternative variants.

Some offices described the alternative variants, or how the "plausible range" was defined by filling in the open-ended comments: For instance, in one case the variants are considered as scenarios and extreme uncertainty and future international migration development is explained. Another NSO states that "… the high level of TFR (2.1) is explained by the fact that it is the symbolic value of the replacement level of fertility. As the baseline level was 1.8, the low level retained is 1.5 because it is 0.3 lower than the baseline (2.1 is 0.3 higher). It is as being the mean EU-15 level. For mortality, the various assumptions are explained by the expected trend of the future decrease in sex and age mortality rates (future trend is the same as past three decade trends / slow down in the trend at all ages / faster decrease at old ages). For migration balance, the level retained as baseline is explained as the mean level observed over the past two decades. A scenario 'without migration' is investigated but this assumption is only combined with the baseline assumptions on mortality and fertility and not with all assumptions (…). The high assumption on migration was 100,000 (+50,000 in comparison to the baseline), which was considered as a reasonable value by experts. The extra migration balance is supposed to consist in immigration only."

One office defined the standard variants as follows: "These are intended as plausible alternative scenarios and not to represent upper or lower limits for future demographic behaviour. For the special case scenarios we say: 'It is also sometimes useful to prepare special case scenarios, or 'what if' projections, to illustrate the consequences of a particular, but not necessarily realistic, set of assumptions.' "

There are variations of the definition when we consider the definition of another office. "For fertility the alternative assumptions are considered as covering a certain quantitatively specified uncertainty interval; for mortality alternative assumptions are considered as covering a 'plausible' range; and, for migration, behind the null scenario, the other assumption is based on values considered possible."

2.3 Who was involved in defining the assumptions

The baseline projection, or the "most-likely" or "best-guess" forecast, generally, is first discussed within the national statistical offices, and then discussed with outside experts (11 of 21 offices, or 52 percent of the respondents). The second most common approach is that the baseline forecast is initially proposed by experts and then discussed within the office (24 percent). Four NSOs (19 percent) discussed and defined it only within the office, and for one NSO the baseline projection was defined fully by outside experts. The one national statistical office that uses the Eurostat projections answered that both answers c) and d) applied. Despite appearances, this answer is not contradictory, as Eurostat produced the population projections in cooperation with the scientists of the country and the Central Statistical Bureau.

Typical comments on the procedure were: "The assumptions were drawn up by an expert group chaired by a member of the NSO and attended by outside experts as well as NSO experts." Or, the answer b) "is closest to the … situation. However, our initial meeting with six … academic experts was largely based on the assumptions used for the previous projections and their views were part of the evidence we took into account in preparing new assumptions."

The absolute majority of the respondents (86 percent) answered that when they defined more projection variants for fertility, mortality and migration, they used the same mechanism to define the alternative scenarios. However, one bureau states that "the mechanism is not applied for each new forecast round for each component; this depends on whether new evidence is available or striking developments have taken place for a specific component."

2.4 Extent of involvement of external experts

This question refers to the number of experts involved in the decision process, in the number of consultations, and if separate meetings took place with different experts for fertility, mortality and migration. In total, 76 percent of the national offices answered that they involved external experts to define the future paths of the model determinants. Of those 16 institutes, 11 told us that they involved 10 or more outside experts; two institutes consulted between five and nine experts; and two offices one to four external experts. One bureau responded that "there is no fixed number of experts; mainly for migration".

The number of consultation meetings with such experts as a group is below 10 meetings in general. Some answers are vague, such as 1-10 meetings. Nine offices recorded fewer than five meetings. Eight of the 21 NSOs stated that there were separate meetings with different experts for fertility, mortality and migration, possibly explaining the relatively high number of meetings. Higher numbers of experts involved and a larger number of meetings were mostly seen in the statistical offices of the "old-15" countries of the European Union than in the "new-10" member states.

2.5 How did you deal with situations when experts disagreed

Three statistical offices told us that they were able to agree on common values, and hence they did not need any specific way to handle discrepancies. In cases where the experts could not reach a consensus for the model input, there is no clear policy. Around 50 percent of the national offices listen first to the experts but then take the final decisions in-house. Only three national statistical offices chose the way of majority vote, and two offices applied a variant of the "Delphi" method. To our understanding there is a clear need in the methods used to reach an agreement in case there are some problems among the external and/or internal experts. A vote is a fast procedure and is a good tool for generating quick, democratic decisions, but may not be applicable for scientific decision-making procedures.

One of the three bureaus that did not feel the need to deal with a "consensus mechanism" reacted as follows: "This was a common discussion to understand everyone's argument." Another NSO that answered with a) and d) answered that, "We base our assumption on facts and reasoning. We try to present as much [sic] facts as possible and we also try to specify where we are uncertain and show how and why we have decided in a certain way. We also try to describe the reasoning behind the assumptions thoroughly in the publication about the population projection." One office that answered that there was a majority vote, describes it in this way: "Assumptions are defined initially within the office. Papers are prepared to justify these assumptions. The experts usually agree with the basic assumptions." The in-house decision choice could be interpreted as the following possibility described here. "The discussion on assumptions should finish by selection [of] the most likely variant for each component or by suggestion of verification. Conclusions from the meeting are implemented by NSO in calculation of official projection."

2.6 Definition of storylines, i.e., plausible visions of the future conditions that would result in certain trends

Over half the institutes (12 or 58 percent) do not use storylines to describe the reasoning behind the assumptions of the future paths of fertility, mortality and migration; seven institutes do use them. Two NSOs did not indicate any of the listed answers, and another office was unsure what was meant by "storylines". Five out of the eight offices that answered with yes define storylines for all three components; three of them defined consistent storylines to bind together at least two components. One of the three noted that they use both variants with more focus on the definition of the storylines for each component separately. The second didn't specify, and the third defined the storylines for low fertility and mortality jointly.

One NSO defines in detail the "binding together" of the storyline and thus provides insights into what probably happens in many of the NSOs. "The main variant is based on what are considered to be the most probable, and therefore reliable, trends for the future: a further fall in mortality, a slight increase in period fertility, inter-regional migration showing a constant probability, international migration at around levels experienced in the '90s. In addition to the main variant, two alternative scenarios have been considered regarding the development for each demographic component. The two alternative assumptions are intended to define the range of variation within which the future population will develop. The scenario imagined in the low variant is marked by minimal economic growth and limited attention paid to social problems. Given such a context, improvements in life expectancy would slow down and there would be no recovery in the fertility rate. Regarding migration, inter-regional and international flows would exhibit modest levels, resulting in a kind of stagnation owing to the low level of attraction exerted by the destinations concerned. Such a scenario would give rise to the lowest projected population level, characterised by the most unbalanced age structure. In the high variant the scenario assumes lively economic growth, providing the opportunity to increase investments also in the social and health fields. This would lead to a higher life expectancy than in the main variant and a considerable recovery in fertility rate. Furthermore, this scenario is also marked by a more intense population movement among regions and an increase in the attractiveness of … as a destination for immigrants from abroad. All of the foregoing factors would lead to the highest projected population level together with a more balanced age structure."

One office describes the need to define separate storylines for all three variables: "On some aspects the storylines are consistent, example: Many young [immigrants] … in the country look for a partner in their country of origin. The rising number of young … will have an upwards effect on the number of marriage migrants. As marriage migrants are rather traditional in behaviour these migrants will have an upwards effect on the fertility level of the … in the country, and of course also on overall fertility."

2.7 Expressed need for further improvements in the procedure for defining the assumptions and directions for improvement

There is a clear common agreement among all respondents: There is a need for improvement to define the assumptions on future fertility, mortality and migration. All 21 NSOs answered that they think further improvements in the procedure are required.

Only one office appeared to be somewhat hesitant about this choice in the questionnaire, but then stated: "We are satisfied with our procedures, but of course improvements can always be made, so we are continuously looking for improvements." Suggested improvements are: "In the sense that the analysis done on the data and the justification of the options should be more released in the publication." And "Everything can be improved. We are always open to discuss improvements. However we investigate new procedures carefully before we accept them as an improvement."

In what direction do the NSOs think the improvements should go? Here the tendency is fairly clear: The most widely chosen improvement options were to "have a more systematic review of all the substantive arguments behind the assumptions" (13 NSOs) and "have some structured interactions with the European demographic research community about the state of the art in our knowledge about future demographic trends (13 NSOs). This was followed by the desire to "involve more experts" (11 NSOs). Further down in the ranking came the hope for more exchange with other NSOs (9 NSOs) and with Eurostat (8 NSOs).

In a nutshell, this extensive and very informative enquiry among EU national statistical offices makes it very clear that all national offices think that the current practice for defining the assumptions for population projections is sub-optimal and needs further improvement. The results clearly indicate the direction into which to move on from the current practice: **Have a more systematic review of the substantive arguments behind the assumptions in the form of a structured interaction with the demographic research community which also facilitates the involvement of more experts**.

In the rest of this paper we will propose a new procedure for defining the assumptions that would try to achieve exactly this goal in all its dimensions. Furthermore, it would not only address the three top ranked directions of improvement but also, if conducted as a Europe-wide exercise in the future, have the potential to significantly strengthen the substantive collaboration among NSOs and with Eurostat.

3 Problems and Fallacies Associated with Expert Opinion and the Road Toward Argument-Based Forecasting

Given the prominence and importance of the assessment of likely future demographic trends in Europe, it is surprising to see how little systematic attention the scientific community has been given to the evaluation of arguments underlying the assumptions of future fertility, mortality and migration trends. While the above-described survey shows that the National Statistical Offices put great hope in input from the demographic research community, this topic has largely been confined to the described processes within statistical agencies. Since these offices are in a way forced to make choices on assumptions in order to fulfill their mandate of producing population projections, they cannot escape this challenging task as easily as academics seem to be able to.

The users of population projections – which are by far the most visible and most relevant products that the demographic research community provides to the rest of society – would rightly expect a broadly-based discussion of likely future trends to be the main topic of international population meetings. They would also expect governments and intergovernmental bodies, which greatly rely on the accuracy of population projections in their policy formulations, to commission major studies to make sure that they get the best possible information about likely future demographic trends. But in stark contrast, the reality shows that at scientific meetings in the field of demography as well as in government-sponsored activities around Europe, the discussion of assumptions used in projections is largely absent or at best a marginal topic.

It is important to point out here that the absence of such structured and prominent discussions is not due to the fact that we know all about the future. Quite the contrary, the sense of uncertainty about whether fertility in Europe will recover or continue to decline, or whether we are already close to a maximum life expectancy or will see continued increases, seems to be even higher than in the past. Moreover, studies on the accuracy of past population projections produced since the 1960s have shown that significant errors were made particularly with respect to anticipating the speed of population ageing in Europe. Generally, for most European countries, the national statistical agencies as well as the United Nations Population Division have assumed far too high fertility levels and far too low gains in life expectancy. While these two independent errors tend to cancel out when one is only interested in population size (fewer than expected deaths compensate the effect of fewer than expected births), they strongly reinforce each other when it comes to population ageing (ageing is enhanced by higher life expectancy and by lower fertility). One could even argue that these significant errors of past projections which failed to anticipate the actual speed of ageing have contributed to the fact that today's societies are not as well prepared for ageing (e.g., in terms of pension systems) as one could have hoped.

Given this situation, the approach presented in this paper will try to show the way for a better inclusion of available scientific knowledge into the process of defining the range of assumptions on future fertility, mortality and migration levels. In other words, it will attempt to facilitate the translation of the vast body of relevant research that exists in the demographic community as well as other related research communities into a definition of specific sets of science-based assumptions for projections. This follows up on earlier work by Lutz, Saariluoma, Sanderson and Scherbov (Lutz et al. 2000), on which the following section partly draws.

As shown in the previous section, up to now this process has mostly happened through the collection of expert opinions. Such procedures typically follow the tradition of Delphi methods that have been well developed and extensively documented elsewhere (Linstone and Turoff 2002). But the problems with expert opinions is precisely that they tend to be opinionated; this can result in all sorts of biases and distortions that are not desirable and do not necessarily reflect the best state of the art in the field. There is abundant evidence that experts tend to hold strong beliefs about the future, which are at the level of emotions and intuitions. Hence, the approach proposed in this paper is nothing short of trying to go beyond opinion-based Delphi and suggest a more objective science-based way. Of course, whenever one has to rely on the views of people in one way or another, this cannot be fully objective, but one can move into this direction by making it inter-subjective and applying the standard scientific tools of peer review and critical evaluation. But in order to make progress in this direction there needs to be something on the table to be evaluated and analytically reviewed. Hence, the argument-based approach will put specific arguments on the table that are directly relevant for the future course of the demographic force under consideration and which can then be critically assessed.

The following considerations are the product of an interdisciplinary collaboration between demographers involved in population projections and an experimental psychologist working in the field of cognitive science, in a way an expert on experts.

One important contribution that meta-science can make to any scientific approach is to investigate the problems in the way arguments are built in specific scientific fields. Such work can help the applied scientists find a more analytical way of thinking in their own fields. One may ask why is it important to critically inspect the argumentative basis of a science. The logic of the answer is very straightforward: All scientific argumentation ends somewhere and from that point on, the area of intuitive assumptions begins. Infinite chains of arguments are impossible, but we need to be aware of this and reflect on the point when we choose to end the chain of argumentation. This point can be right next to the object of observation, in which case there is no argumentative foundation at all. It can also be too far away from the object in which case the arguments considered and the objects are hardly linked any more. The choice of this cut-off point needs to be based on expert judgment. But this is judgment at the meta-level rather than at the level of the object itself. Such judgment must be based on some sense of plausibility or intuition as it is typically called in cognitive science and foundational analysis.

Intuitions in the foundations of scientific ways of thinking are unavoidable. We cannot get around them; we have to learn how to live with them. The first step in this direction is to understand them in the right manner. The problem with intuitive foundations of science is not that all our intuitions would immediately and necessarily be false, but that we do not know whether they are true and to which degree they are true. This means that we have to adopt a dynamic stance toward them. We have to turn our attention to them and carefully consider the possible strengths and weaknesses in them. When we understand the intuitive foundations better, we are able to use this new understanding for the advancement of science. We can open new perspectives to knowledge and justify the search for new types of knowledge. Indeed, the ultimate goal of such foundational work is to deepen our understanding of what we are doing. This is a way to speed up the progress in science in general and in the field of making necessary assumptions for population projections in particular.

In the following, we will critically review some of the most common problems with expert judgment and the reference to empirical findings and discuss how arguments should be framed in order to avoid such problems. The goal is not to prove such arguments incorrect or empty in content, but rather to make sure that the specified arguments actually refer to possible causal mechanisms and are specific enough to be falsifiable. Only the evaluation of such arguments will add to our science-based knowledge about likely ranges of future demographic trends.

An important prerequisite for valid argumentation is the clarification of what is the assumed cause and what is the effect. The *explanans*, i.e., the explanatory premise, refers to statements that explain the *explanandum*, i.e., the phenomenon which should be explained on the grounds of *explanans*. In argumentation analysis, it is always central to consider carefully the form and explanatory power of the *explanans*.

A typical example for the confusion between *explanans* and *explanandum* is the assumption that something will not happen because it has not yet been observed, something the literature calls “curve illusion.” In this kind of false argument, one views the shape of an observed curve (the phenomenon to be explained) as the driver that produces a pattern. An example from the field of fertility assumptions is the frequently held view that there is some "rock bottom" fertility level below which fertility will not fall. This is simply justified by the fact that fertility has never fallen below such a level in any country. There may well be good arguments to assume that fertility will not approach zero in the future, but they cannot be based merely on the description of the "curve" observed so far. Interestingly, in the field of projecting life expectancy, although human history has never experienced a national life expectancy of above the current maximum of some 86 years for Japanese women, few people think that future increases will be impossible. But while a historically unprecedented level is no longer used as an argument in making mortality assumptions, the current practice of (blind) trend extrapolation is not much better from a meta-scientific perspective, if it does not provide any plausible reasons for why life expectancy is assumed to continue to grow at the same speed as in the past or at a decelerating speed, as some agencies assume. More generally, in order to avoid such circularity of taking the *explanandum* for the *explanans*, one would have to anchor the argument in the world outside of the curve itself (the observed trends). If such an anchoring is not explicated, the argument cannot be valid.

The circulatory problem also exists when we refer to two different measurements which may be affected by the same cause but do not influence each other. When we measure a fever in a child, we do not think that the high temperature in itself is the illness or that the temperature in the mouth is caused by the temperature in the armpit, even though the correlation would be substantial. Instead, we look for the illness in the body, which explains the high fever measured at both points. We know that the body defends itself from many different types of illnesses by producing a fever and therefore, we look for further symptoms to cancel out incorrect diagnoses and to find the true explanation.

Another problem in this context of defining valid arguments is the confusion of differentials with causes. Much of the social sciences have been inspired by the observation of differentials. Individuals and their behavior differ from place to place, over time and among individuals. These differentials typically give rise to the formulation of explanations as to why the observed patterns of behavior differ. These explanations point the way to the more general causes of behavior. In many cases, however, the analysts stop short of providing real explanations for the observed differentials and suffice by describing only the differentials. An example for such inference from differentials to causes is when people point at the fact that urban women typically have lower fertility than rural women and conclude from this that increasing urbanization will lead to lower fertility. But this conclusion is only correct if it is assured that there is indeed a causal relationship from the kind of living environment on the number of children. There probably are such real causes, but in order to make it a valid argument, the possible causal mechanism has to be identified and discussed. This does not necessarily mean that they have to be proven in the sense of strong causality which may be very difficult. But at least the identified mechanisms should result in a plausible storyline. Such an argument can then be properly evaluated both with respect to its validity and its relevance.

The same problem affects the currently popular notion of a "second demographic transition" (SDT). It is a name given to a bundle of observed trends in certain values related to sex and partnership and is by its very nature an *explanandum* (or a "curve" in the above terminology). Although SDT is sometimes referred to as a theory, it does not potentially have predictive power (such as stating that a country that will move into the direction of a more liberal attitude toward sex will have fewer children in the future) and hence must not be mistaken as a testable *explanans*. In this respect the concept of SDT also suffers from another frequent problem that makes many proposed patterns of explanation inappropriate as valid arguments, namely, the lack of specificity or, in other words, the fact that they are too general. If an argument is too vaguely formulated or too broad and general in its content so that there is no way to potentially reject it, the argument is not helpful for broadening our science-based understanding of the future.

On the opposite end of the spectrum of problems lie those with arguments that are too specific and too narrow. While such arguments may well be falsifiable in the sense that they have specific information content that can be evaluated in the light of empirical evidence and theoretical cohesion, they may not add much to our overall understanding of likely future trends because they only address a very narrow aspect of all factors that jointly determine the future trend of the demographic force under consideration. A good example for such an argument in the field of fertility determinants is a focus on declining human sperm counts. While there seems to be convincing evidence that in some countries there have been significant declines in the quantity and quality of sperm counts, and a sufficient number of healthy sperm clearly are a prerequisite for natural conception, some commentators have taken this as an explanation for the declining birth rate. But this relationship is far from straightforward. As discussed in a recent special issue of the *International Journal of Andrology* (Jørgensen et al. 2006), declining sperm quality and counts may well affect the waiting time to conception (and only in rare cases lead to infertility), but this interacts in a complex way with characteristics of the partner as well as the nature of the partnership.

But the problem of partial explanations is much broader than the example above. Essentially all arguments about future trends in fertility, mortality and migration focus on certain partial aspects, while leaving others out. Hence, it is one of the most challenging tasks for the development of a new model for argument-based assumption making to bring these different aspects together in a comprehensive way in which the relative importance of the different arguments in determining the future course of the force under consideration are assessed. This will be done in the form of weights to be attached to the different factors that should resemble reality as closely as possible. In other words, we will distinguish between assessing the validity of certain arguments and their relevance in terms of influencing the overall trend of the demographic force.

Based on the above-described considerations and the identification of possible traps and pitfalls in the specification of arguments, in the following section we will present a scheme of core substantive arguments that try to avoid (as far as possible) the above problems. These could become the basis of a systematic future scheme for defining argument-based assumptions for population projections for essentially all countries in the world. While the specific formulation of forces and arguments in the following section is geared toward population projections in industrialized countries, it will be relatively easy to adapt the framework for use in developing countries as well.

4 Specification of an Argument-Based Questionnaire and Application to the Future of Mortality

Based on the above-described principles and in consultation with fertility, mortality and migration experts from around the world, as part of the MicMac project, IIASA's World Population Program has produced a prototype questionnaire in an attempt to further operationalize the general approach of argument-based definition of assumptions for population projections. This questionnaire presents extensive sets of pre-defined arguments around 5-6 forces for each of the three components fertility, mortality and migration. For lack of space in this paper, we will only consider the mortality part in more detail.

The choice of moving to a predefined set of arguments that have been derived by the so-called meta-experts from the literature presents an important further step in the evolution of expert- and argument-based population projections at IIASA. The book entitled *Future Demographic Trends in Europe and North America: What can we assume today?* (Lutz 1991) was the first systematic attempt in this direction. It summarizes the results of an IIASA conference on Future Changes in Population Age Structures held in Sopron, Hungary, October 18-21, 1988, in which prominent demographers from Europe and North America were invited to summarize the state of the art in terms of substantive knowledge about the likely future trends of fertility, mortality and migration. In doing so the authors were explicitly instructed to discuss competing hypotheses and arguments. In a final roundtable, the authors were then asked to quantify their views about the future in terms of alternative numerical assumptions about the future trends in fertility, mortality and migration, and argue the chosen values with the substantive arguments discussed in their papers. The IIASA team then translated these alternative assumptions into alternative population scenarios for Eastern Europe, Western Europe and North America. Although this book went much further than typical edited volumes in pushing the authors to nail down their science-based views in terms of specific numerical assumptions and relate these to the text, the whole exercise was still largely the expression of informed views of prominent demographers.

A similar and even more clearly structured exercise was then carried out at the global level. *The Future Population of the World: What can we assume today?* (Lutz 1994) contains only invited and clearly targeted papers on the three components in different world regions summarizing the empirical evidence, the state of the art in terms of analysis and the reasoning behind the definition of alternative scenarios for the future. Moreover, the volume includes authors on the same topics (such as the future of longevity in industrialized countries) who were known for holding opposing views on these topics. In a rather intensive "scenario definition session," all authors had to agree on a common set of alternative scenarios definitions for the three components for the different world regions. It was a highly interesting session particularly with respect to the authors of opposing views who had to agree about their disagreement in quantitative terms. Again, the IIASA team then translated these alternative assumptions into global population scenarios for 13 world regions, the same regions that are being used by the Intergovernmental Panel on Climate Change (IPCC) in their scenarios. In the 1996 revised edition of the book (Lutz 1996), some updates were made on the basis of new empirical data and a section was added which converted the high-low ranges defined by the experts into the first probabilistic world population projection. They were also published in *Nature* under the title "Doubling of world population unlikely" (Lutz et al. 1997). Some of these IIASA population scenarios were also used by the IPCC in their SRES scenarios (Nakićenović et al. 2000). All further revisions of the IIASA world population projections (Lutz et al. 2001, 2004, 2008) have essentially been based on updates of this original assumption definition exercise.

The new argument-based projection approach, developed as part of the MicMac project, will be used for a new set of IIASA world population projections which for the first time will be carried out at the level of individual countries. A serious, argument-based exercise carried out at the level of large numbers of countries can no longer be based on a group of experts coming together in room. Moreover, under the past approach the choice of arguments and reasoning supporting the specific assumptions made was still largely up to the invited expert and hence contained strong elements of subjectivity. In addition to that, even two experts holding identical views about the future would likely present their arguments in different forms, which would make it hard to compare and evaluate the arguments systematically. This is even more so when larger numbers of experts are involved, such as the

17 in the mortality exercise described below, or hundreds as will be the case in the new set of global projections. For these reasons it was decided to move to a large set of predefined arguments that are derived from the scientific literature. This clearly puts the workload more on the "meta-experts", i.e., the team leading the exercise, but comes with the great benefit of having a better structured, managed, more transparent and therefore more science-based and credible process. Still, individual experts have the opportunity to add further arguments to the list that they think the meta-experts have forgotten or not specified properly.

When it comes to defining specific arguments about, e.g., the future of fertility, it soon becomes clear that they can operate at very different levels ranging from bio-medical factors to individual preferences or the effect of economic conditions or government policies, and that these levels require differential treatment. Hence as a first step in the attempt to come up with arguments, the three components of demographic change were each decomposed into 5-6 presumably independent forces. The future trends of these have to be argued separately.

Since this questionnaire has already been operationalized by the UK National Office of Statistics and used in two recent rounds of population projections, all reference to a specific country below will be for the example of the UK. These specified forces are:

Major forces on which future fertility (F) will depend:

F1. The trend in ideal family size and the strength of individual desires for children as compared to other joys in life

F2. The trend in the patterns of education and work, including the proportion of time to be dedicated to the professional side of life (life – work balance)

F3. Changing macro-level conditions (government policies, child care facilities, housing, etc.) that influence the cost of children in a broader sense

F4. Changes in the nature and stability of partnerships

F5. Changing bio-medical conditions (sperm quality and counts, female fecundability, new methods for assisted conception)

F6. Changes in population composition and differential trends in population subgroups

Major forces on which the future of life expectancy (L) will depend:

L1. Changes in biomedical technology

L2. Effectiveness of health care systems

L3. Behavioral changes related to health

L4. Possible new infectious diseases

L5. Environmental change, disasters and wars

L6. Changes in population composition and differential trends in population subgroups

Major forces influencing net migration gains (M):

M1. Trends in the main motives for international migration

M2. Trend in migration pressure resulting from changes in countries of origin

M3. Trend in the attractiveness of the UK as a country of destination

M4. Costs of migration in the broader sense

M5. Effectiveness of barriers to unregulated migration flows

Once these major forces have been defined, the next step is to define a set of specific arguments that would have a potential influence on the future course of these forces that in turn would contribute to future changes in fertility, mortality and migration in one or the other direction. Depending on the force, between five and nine different arguments were predefined for each force. This makes a total of over 100 arguments which cannot be presented or discussed here because of space limitations. The arguments referring to mortality are listed in Appendix Table 2. It should be stressed that the specification of the individual arguments is still open to further modification. Hence, for each new application, the specifications need to be reviewed. Such further modification will be necessary for the application to developing contries.

For each of the specified arguments, experts are asked to make a judgment about its validity based on the scientific evidence to the best of their knowledge. They are given five predefined choices about the validity of the argument ("very likely to be right", "more right than wrong", "do not know/ambivalent", "more wrong than right", "very likely to be wrong"). In addition experts are given space for free format comments.

But it is not only the validity of an argument that matters; it is also the relevance of the argument for the future course of the force under consideration. There may be arguments that are very likely to be true, but which are completely irrelevant for the question under consideration. Hence experts have to be asked a second question concerning the likely impact of each argument on the demographic component (see Table 1).

How should one make sense of and combine expert judgment about the validity and relevance of more than 100 arguments? This requires a weighting of arguments according to their assigned validity and impact. To calculate the scores for the overall assessment of all arguments per force, and over all forces, the following weights were applied (Table 1):

Table 1. Weighting factors for the validity and impact of arguments.

A Validity of argument		Weighting factor
1.	Very likely to be right	1.00
2.	More right than wrong	0.75
3.	Do not know / ambivalent	0.50
4.	More wrong than right	0.25
5.	Very likely to be wrong	0.00
B Impact of argument		**Weighting factor**
1.	A large upward influence on life expectancy	1.0
2.	A small upward influence on life expectancy	0.5
3.	Little or no influence on life expectancy	0.0
4.	A small downward influence on life expectancy	-0.5
5.	A large downward influence on life expectancy	-1.0
6.	Don't know	0.0

Assigning weights is only the first step toward a comprehensive evaluation of the different forces that are likely to shape the future path of life expectancy. The further processing of these scores is best illustrated through the exercise in which 17 international experts participated at the MicMac Meeting on Assumptions on Future Mortality and Morbidity Trends in Europe, held at IIASA September 10-11, 2007. It used an Excel-based interactive version of the questionnaire that was developed in collaboration with the British National Statistics Office (ONS) and was sent to the selected experts via email before the meeting.

The questionnaire begins with a self-assessment of the competences of the experts in different fields of demography. It is followed by three broad sectors on fertility, mortality and migration. We focus in this paper on the mortality sector.

The mortality sector (and the others) is structured along the major forces (L1-L6) described above. For each of these forces, several arguments are listed that imply either an upward or downward pressure on that driver. The experts are asked to evaluate individual arguments and to assess the validity and impacts of these arguments on future trends in the driving force and also the relative importance of the driving forces on the future course of overall mortality. These assessments are stored in a database. Before the user is shown the overall implications of the sum of individual assessments made on the future course of fertility, mortality and migration, he/she is asked to make an independent assessment of what is considered the most likely future trend in these components. In this context the experts are also asked for probabilistic assessments about the future trends of overall fertility, mortality and migration. This information can be used later to study the consistency of argumentative and numerical statements given by the expert and possibly give those experts who show significant inconsistencies a lower weight when calculating aggregate results.

The questionnaires on the mortality sector were completed by 17 experts who returned them to IIASA prior to the meeting. The resulting data base was then analyzed in the following way:

Step 1:

For each argument and for each respondent we applied the proposed weighting factors for the impacts and validity of arguments. The numbers in Tables 2 and 3 give the distributions of the responses of the 17 experts over all the arguments.

Table 2. Answers on the validity of the arguments listed in Appendix Table 2.

L1: Changes in bio-medical technology will have:

	Argument L1.1	Argument L1.2	Argument L1.3	Argument L1.4	Argument L1.5
Number of respondents					
1	6	6	6	8	0
0.75	8	8	7	6	1
0.5	1	2	4	2	5
0.25	1	1	0	1	6
0	1	0	0	0	5

L2: Effectiveness of health care systems will have:

	Argument L2.1	Argument L2.2	Argument L2.3	Argument L2.4	Argument L2.5	Argument L2.6
1	4	1	2	0	10	3
0.75	6	11	7	5	5	10
0.5	2	3	6	3	0	1
0.25	5	1	2	7	1	2
0	0	1	0	2	1	1

L3: Behavioral changes related to health will have:

	Argument L3.1	Argument L3.2	Argument L3.3	Argument L3.4	Argument L3.5	Argument L3.6
1	5	4	2	2	3	1
0.75	9	3	7	8	7	14
0.5	1	7	2	1	5	1
0.25	2	1	2	3	1	1
0	0	2	4	3	1	0

L4: Possible new infectious diseases and resurgence of old diseases will have:

	Argument L4.1	Argument L4.2	Argument L4.3	Argument L4.4
1	2	2	4	3
0.75	4	4	9	8
0.5	2	4	2	4
0.25	4	3	1	2
0	5	4	1	0

L5: Environmental change, disasters and wars will have:

	Argument L5.1	Argument L5.2	Argument L5.3	Argument L5.4	Argument L5.5	Argument L5.6	Argument L5.7
1	3	1	4	1	0	3	0
0.75	2	3	9	8	3	8	5
0.5	4	7	4	6	7	4	4
0.25	4	4	0	1	3	0	6
0	4	2	0	1	4	2	2

L6: Changes in population composition and differential trends in population subgroups will have:

	Argument L6.1	Argument L6.2	Argument L6.3	Argument L6.4
1	2	1	7	3
0.75	1	12	8	4
0.5	5	1	1	8
0.25	8	1	1	0
0	1	2	0	2

Table 3. Answers on the impact of arguments.

L1: Changes in bio-medical technology will have:

	Argument L1.1	Argument L1.2	Argument L1.3	Argument L1.4	Argument L1.5
1.0	3	6	4	3	0
0.5	13	11	11	9	1
0.0	1	0	2	5	12
-0.5	0	0	0	0	2
-1.0	0	0	0	0	1
0	0	0	0	0	1

L2: Effectiveness of health care systems will have:

	Argument L2.1	Argument L2.2	Argument L2.3	Argument L2.4	Argument L2.5	Argument L2.6
1.0	1	4	0	2	6	3
0.5	2	12	1	11	10	10
0.0	8	0	9	4	1	4
-0.5	4	0	6	0	0	0
-1.0	2	0	0	0	0	0
0	0	1	1	0	0	0

L3: Behavioral changes related to health will have:

	Argument L3.1	Argument L3.2	Argument L3.3	Argument L3.4	Argument L3.5	Argument L3.6
1.0	7	0	1	3	0	1
0.5	9	0	10	8	0	12
0.0	0	3	6	6	8	4
-0.5	1	13	0	0	7	0
-1.0	0	1	0	0	2	0
0	0	0	0	0	0	0

L4: Possible new infectious diseases and resurgence of old diseases will have:

	Argument L4.1	Argument L4.2	Argument L3.3	Argument L4.4
1.0	0	0	1	1
0.5	0	0	6	1
0.0	7	6	10	3
-0.5	6	7	0	8
-1.0	4	3	0	3
0	0	1	0	1

L5: Environmental change, disasters and wars will have:

	Argument L5.1	Argument L5.2	Argument L5.3	Argument L5.4	Argument L5.5	Argument L5.6	Argument L5.7
1.0	0	1	0	0	0	1	0
0.5	0	0	0	2	1	4	0
0.0	11	9	10	10	8	9	7
-0.5	6	5	6	2	6	0	6
-1.0	0	1	0	0	0	0	1
0	0	1	1	3	2	3	3

L6: Changes in population composition and differential trends in population subgroups will have:

	Argument L6.1	Argument L6.2	Argument L6.3	Argument L6.4
1.0	0	0	0	0
0.5	6	6	2	0
0.0	7	7	11	16
-0.5	4	4	4	0
-1.0	0	0	0	0
0	0	0	0	1

Step 2:

The scores for validity and for impact are multiplied for each expert and each argument. Hence only the arguments that are considered "very likely to be right" and at the same time have "a large upward influence on life expectancy" get the maximum score of 1.0. Answers that are either considered to be wrong or irrelevant for life expectancy get a score of zero. The lowest score of -1.0 is given to arguments that are considered to be right and have a large downward effect.

In order to summarize the results across all arguments for one force, L, the average of all non-zero scores for the arguments referring to that force was taken. The rationale behind this calculation is that a score of zero implies that the argument is either considered totally wrong or irrelevant and should not be treated on an equal basis with arguments that imply either a relevant shift upward or downward. Those arguments are given equal weight in the calculations presented here, but this could easily be changed should there be a good reason for doing so. Table 4 gives the averages of the non-zero scores for each force by individual experts.

Table 4. Weighted average score over all arguments for each force.

	Average (excluding zeros) over all arguments					
Expert #	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**
1	0.33	0.38	0.69	-0.38	0.38	-0.04
2	0.88	0.58	0.38	0.38	0.31	0.00
3	0.50	0.35	0.23	-0.69	-0.38	0.13
4	0.42	0.25	0.38	-0.38	0.00	-0.13
5	0.63	0.25	0.25	-0.13	-0.38	0.00
6	0.50	0.50	0.50	-0.83	-0.50	0.50
7	0.38	0.35	0.06	0.13	-0.09	-0.25
8	0.38	0.13	-0.13	-0.22	-0.33	0.00
9	0.38	0.34	0.38	0.13	-0.19	0.31
10	0.21	0.13	0.04	0.13	0.00	-0.13
11	0.47	0.25	0.20	-0.38	-0.25	0.38
12	0.34	0.63	0.16	-0.29	-0.13	0.00
13	1.00	0.81	-0.04	-0.38	-0.38	0.00
14	0.34	0.30	0.16	-0.25	-0.13	-0.06
15	0.50	0.13	0.19	-0.25	-0.38	0.13
16	0.22	0.23	0.38	0.06	0.13	-0.13
17	0.65	0.23	-0.22	0.58	0.19	0.44

Step 3:

In order to combine the effects of the different forces in jointly shaping the future course of life expectancy, the expert had to provide an additional piece of judgment. They were asked to distribute 100 points over the six forces indicating the relative weight of each force in shaping the overall trend in life expectancy. The software was programmed in such a way that the expert could not make mistakes. They were only allowed to continue to the next step of the questionnaire if the sum of the points given to individual forces actually equaled 100. This information is presented in Table 5. Table 6 gives the weighted sum over all forces. The columns in Table 6 result from a multiplication of the respective columns in Tables 4 and 5.

Table 5. Relative importance of forces as stated by experts (sum up to 100).

Expert #	L1	L2	L3	L4	L5	L6
1	5	60	20	0	0	15
2	30	20	20	10	10	10
3	40	10	5	20	20	5
4	35	25	25	5	5	5
5	20	20	40	5	10	5
6	50	10	30	5	5	0
7	25	15	35	5	10	10
8	20	20	10	20	10	20
9	25	25	15	5	5	25
10	25	15	40	5	3	12
11	25	20	25	10	10	10
12	20	20	20	5	5	30
13	30	30	30	3	5	2
14	5	30	35	5	10	15
15	15	20	30	10	10	15
16	25	35	25	5	5	5
17	23	23	20	8	20	8
Average	**25**	**23**	**25**	**7**	**8**	**11**

Table 6. Overall assessment: Relative importance (Table 5) x score for each force (Table 4).

Expert #	L1	L2	L3	L4	L5	L6	SUM
1	0.02	0.23	0.14	0.00	0.00	-0.01	**0.37**
2	0.26	0.12	0.08	0.04	0.03	0.00	**0.52**
3	0.20	0.04	0.01	-0.14	-0.08	0.01	**0.04**
4	0.15	0.06	0.09	-0.02	0.00	-0.01	**0.28**
5	0.13	0.05	0.10	-0.01	-0.04	0.00	**0.23**
6	0.25	0.05	0.15	-0.04	-0.03	0.00	**0.38**
7	0.09	0.05	0.02	0.01	-0.01	-0.03	**0.14**
8	0.08	0.03	-0.01	-0.04	-0.03	0.00	**0.01**
9	0.09	0.09	0.06	0.01	-0.01	0.08	**0.31**
10	0.05	0.02	0.02	0.01	0.00	-0.02	**0.08**
11	0.12	0.05	0.05	-0.04	-0.03	0.04	**0.19**
12	0.07	0.13	0.03	-0.01	-0.01	0.00	**0.20**
13	0.30	0.24	-0.01	-0.01	-0.02	0.00	**0.50**
14	0.02	0.09	0.05	-0.01	-0.01	-0.01	**0.13**
15	0.08	0.03	0.06	-0.03	-0.04	0.02	**0.11**
16	0.03	0.05	0.11	0.01	0.01	-0.02	**0.23**
17	0.15	0.05	-0.04	0.04	0.04	0.03	**0.27**
Average	**0.12**	**0.08**	**0.05**	**-0.01**	**-0.01**	**0.01**	**0.24**

The extreme right column in Table 6 presents what can be called the final quantitative result of the argumentation exercise for each expert. Up to this point the experts had not yet been asked to give their views in terms of actual number of likely increases in life expectancy per decade. This was consciously done in order to avoid the expression of a premature and argumentatively unfounded statement about the end result of the process which would only reflect a personal opinion. In contrast the described exercise resulted from a pure bottom-up assessment of individual arguments on the six forces combined.

The resulting overall scores have to be interpreted in the context of the possible range of scores which theoretically can go from an extreme high of +1.0 to an extreme low of -1.0. Table 6 shows that the individual overall scores range +0.49 (expert #13) to +0.02 (expert #8). In the first instance, this result implies that when viewing all possible drivers together, all experts expect a continued increase in life expectancy. A negative overall score would have implied a likely decrease in life expectancy.

If all experts are given equal weight, the average score across experts is +0.21. This implies the expectation of moderate further increases in life expectancy. A comparison across the columns in Table 6 also shows that the experts attribute the biggest share of this expected increase to expected future changes in bio-medical technology (force L1) followed by the future effectiveness of health care systems (force L2). Of the six forces considered, only environmental change, disasters and war (L5) and the possible emergence of new or the re-emergence of old infectious diseases (L4) are seen as rather minor downward risks on life expectancy which are outweighed by the other forces of change.

The great variance among experts also indicates quite a degree of uncertainty in the future trend. While in these illustrative calculations all experts have been given equal weight, one can easily choose to weigh them differently based, for instance, on their self-stated expertise in the field, on some indicators of self-contradiction with the questionnaire itself, or through external assessment of status and credibility (although this can enter difficult territory).

At the end of the questionnaire, after going through all of the arguments, the experts were asked to state their numerical suggestions for future ranges in male and female life expectancy in the UK that would cover roughly two-thirds (67 percent) of the possible distribution in 2030. The resulting distributions are shown in Figure 1 for men and in Figure 2 for women. They clearly indicate the big difference in assessments among experts, with some of them being much more pessimistic than others, and some of them much more certain about the likely future trend than others as reflected in very narrow intervals. The overall averages show a life expectancy of 82 for men and 86 for women in 2030. This corresponds roughly to an increase of two years per decade.

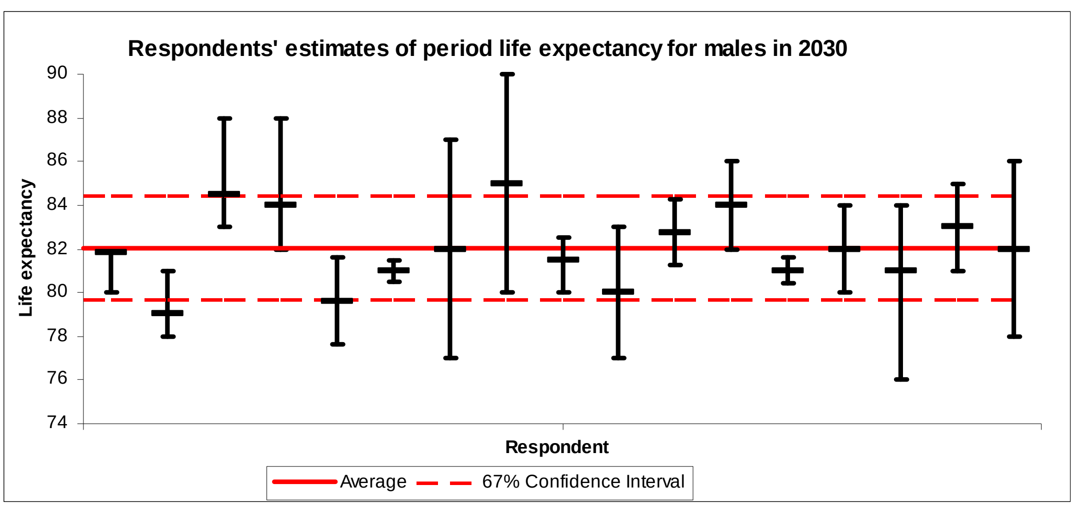

Figure 1. Best guesses and 67 percent uncertainty distributions about male life expectancy in 2030 by expert.

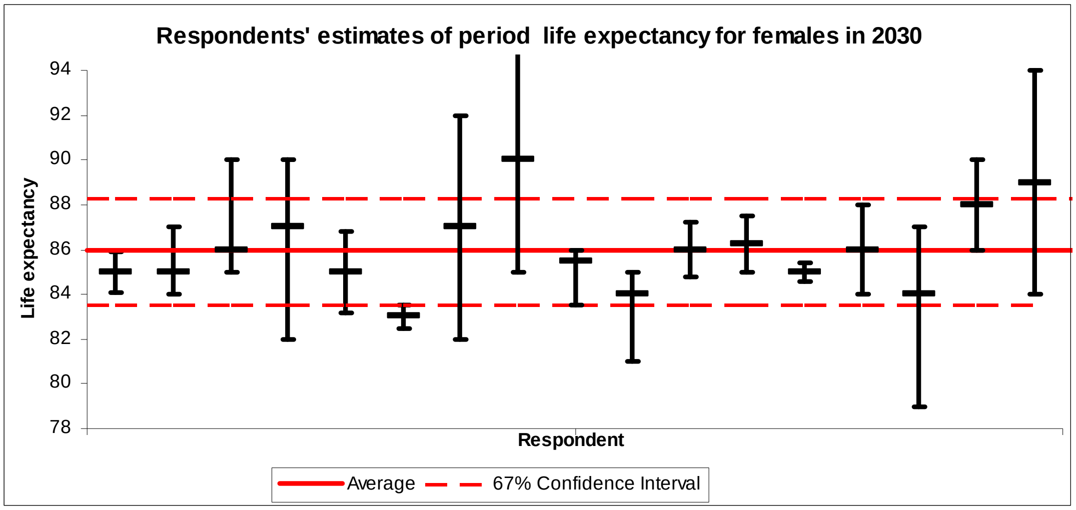

Figure 2. Best guesses and 67 percent uncertainty distributions about female life expectancy in 2030 by expert.

5 Discussion and Outlook

This new argument-based approach for providing a more scientific basis for defining assumptions for population projections is directly responsive to the needs and priorities as expressed in the above-described survey of EU National Statistical Offices. It attempts to facilitate the desired, more systematic review of the substantive arguments behind the assumptions in the form of a structured interaction with the demographic research community, which also allows for the involvement of many more experts. It offers an option for the future for all national statistical agencies in Europe and, if conducted in a concerted effort, it can facilitate the interaction among different agencies and with Eurostat. The UK Office of National Statistics has already used this questionnaire for two rounds of projections and further improved it. There are currently plans to use a further modified version of the questionnaire for a major new exercise to produce new population projections for most countries in the world to be carried out by IIASA in collaboration with Oxford University.

The application of this argument-based forecasting approach is not limited to applications for fertility, mortality and migration. An isomorphic procedure can easily be applied to other demographic dimensions considered in the MicMac model such as education, health status or household status. It is an appropriate way to go beyond rather simple ad hoc scenarios in the definition of future demographic transition rates and exploit the entire body of existing literature and knowledge about the substantive forces that are likely to determine these transitions in the future. One may choose to further decompose the trends under consideration. In the case of mortality, for instance, one may consider it desirable to separately study the trends in different causes of death. This would be feasible in the context of the described approach although one would have to be careful not to increase the complexity beyond a manageable point.

Another potentially consequential innovation of this approach is that the standardized computer-based format of the questionnaire allows for the inclusion of almost unlimited numbers of experts. The questionnaire can be widely distributed electronically and certain quality criteria can help to make sure that only the information provided by competent persons is taken into account. For this the first criterion would clearly be the ranking of competence on the subject matter and the statement about the years already working in the field as provided by the respondent. It would also be advisable not to invite every chosen expert to answer all three sections on fertility, mortality and migration, but to choose only the field in which he/she feels most expert. It would be easy to "test" the competence of the respondent in the chosen field by including a few factual questions or questions relating to the key literature in the field to filter out those who do not know the field well enough. Since the processing of the responses will be anonymous, the "failing" respondents would not know about this and hence there would be no hurt feelings. As already indicated above, certain criteria applied to the internal consistency of the answers given could be used to assign different weights to different experts when aggregating the answers across experts.

The inclusion of a larger number of experts who are not only drawn from academic demography, but also from a broader range of other relevant disciplines as well as from other government agencies and civil society organizations, might serve two additional important goals: (1) it could bring up some more unconventional but relevant information and views that the traditional experts do not see in a possibly too narrow perspective, and (2) it could create a broader public interest and therefore ownership of population projections in society. Once population projections are not seen as something cooked up behind closed doors by a few people, but rather convey the sense that a large number of persons from different segments of society have contributed to them, more people will feel ownership of them and possibly take the projections more serious in their own work. And last but not least, more people will get a better sense of what is known and what is unknown about the future demographic trends once they have to work through all the argumentation process themselves. It may also sensitize them as to what factors can be changed by policies and what has to be taken as given.

This points to another possible use of the results of such an argumentation exercise that goes beyond forecasting and has policy relevance. It helps to prioritize the forces that lead to better health and lower mortality. While the synthesis of arguments as described above only results in a score that indicates into what direction and how strongly the demographic component is likely to change, it does provide specific quantitative information about the relative importance of different forces on the future level of that component. In terms of the above-described mortality-related exercise, one can, for instance, equate the overall score of +0.21 to the average decadal increase of two years of life expectancy given by the experts and then decompose the gain in overall life expectancy into the elements which are due to the specific forces. In this sense, the score of force L1 (changes in bio-medical technology) can account for more than a year of the expected future decadal increase of two years. It also shows that without the risk of new infectious diseases (L4), life expectancy would increase by 2.1 years, and without the risk of environmental change, disasters and wars (L5) by 2.4 as compared to the actual 2.0 assessed by the experts. While such external factors may be harder to influence, factors associated with the effects of public health (a score of +0.08 indicating about two-thirds of the effect of bio-medical technologies) are better candidates for policy action. In this sense the argument-based approach presented here not only offers a better scientific basis for projections than current practice, it also offers otherwise unavailable insights into the relative importance of the different factors that influence the different forces which jointly shape our demographic future.

6 References

Jørgensen, N., C. Asklund, E. Carlsen, and N.E. Skakkebæk. 2006. Coordinated European investigations of semen quality: Results from studies of Scandinavian young men is a matter of concern. *International Journal of Andrology* 29: 54-61.

Linstone, H.A. and M. Turoff, Eds. 2002. *The Delphi Method: Techniques and Applications*. Newark, NJ: New Jersey Institute of Technology. Digital reproduction of the original 1975 version: http://is.njit.edu/pubs/delphibook/

Lutz, W., Ed. 1991. *Future Demographic Trends in Europe and North America. What Can We Assume Today?* London: Academic Press.

Lutz, W., Ed. 1994. *The Future Population of the World. What Can We Assume Today?* London: Earthscan.

Lutz, W., Ed. 1996. *The Future Population of the World. What Can We Assume Today?* Revised edition. London: Earthscan.

Lutz, W., P. Saariluoma, W.C. Sanderson, and S. Scherbov. 2000. New Developments in the Methodology of Expert- and Argument-Based Probabilistic Population Forecasting. Interim Report IR-00-020. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lutz, W., W. Sanderson, and S. Scherbov. 1997. Doubling of world population unlikely. *Nature* 387: 803-805.

Lutz, W., W. Sanderson, and S. Scherbov. 2001. The end of world population growth. *Nature* 412: 543–545.

Lutz, W., W.C. Sanderson, and S. Scherbov, Eds. 2004. *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Sustainable Development*. London: Earthscan.

Lutz, W., W. Sanderson, and S. Scherbov. 2008. The coming acceleration of global population ageing. *Nature* 451: 716-719.

Nakićenović, N. et al. 2000. *Emissions Scenarios*. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge, UK: Cambridge University Press.

Prommer, I. and C. Wilson. 2006. An Analytical Summary of the Current Practices of Definition of Assumption Making in Population Projections. Deliverable D14. Report prepared for the European Commission 6th Framework Project "Bridging the Micro-Macro Gap in Population Forecasting" (MicMac), Contract No. SP23-CT-2005-006637. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Appendix Table 1. Summary of Answers to the Questionnaire

ST = stochastic forecasts are made; M = more

	QUESTION / NSO id.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	**In your officially published population projections, how many different assumptions do you have for the future paths of ...**																					
a	Fertility	3		3	3	3	1	2	1	3	1	3	M	3	3	ST	1	3	1	M	3	3
b	Mortality	3		3	3	1	1	2	1	3	1	3	M	3	3	ST	1	3	3	2	3	3
c	Migration	3		3	3	2	1	2	1	3	1	3	M	3	1	ST	1	2	3	2	3	3
2	**If there is more than one (most likely) assumption for fertility, mortality or migration, has there been an explicit statement in your discussions about what these alternative assumptions should stand for? Are they considered as …**																					
a	the most extreme cases that one can consider possible?																					
b	covering a “plausible” range?	X		X	X	X		X	X	X		X	X	X	X				X			X
c	covering a certain quantitatively specified uncertainty interval (such as 67% or 80% of all possible future paths)?															X						
d	There was no explicit specification of the meaning of alternative assumptions.																X				X	
e	Other (please explain).																	X		X		
3	**Have the assumptions for the baseline (most likely) projection primarily been ...**																					
a	defined and discussed only within the statistical office?							X			X								X	X		
b	defined initially within the office and then discussed with outside experts?			X	X	X	X			X			X	X	X	X	X					X
c	generated fully by outside experts?		X																			
d	initially proposed by experts, then defined within the office?	X	X									X						X			X	
4	**Have the assumptions for the alternative variants or scenarios been defined by the same mechanism?**																					
a	Yes.	X	X	X	X	X		X		X		X	X	X	X	X	X	X	X	X	X	X
b	No. (By what mechanism: a, b, c, d from																					

	above list for Question 3?)																					
5	**In case that external experts were involved in this process:**																					
a	How many external experts were involved all together?	11	10		5	14	11		12	5		14	3	20	X	20	4	12			12	40
b	How many consultations (meetings) with such experts were there all together?	7			10	4	2		1	1		5	5	4	1	1	1	?			8	4
c	Were there separate meetings with different experts for fertility, mortality and migration?	X		X	X							X		X		X		X			X	
6	**What was the background of the external experts involved? (more than one answer possible). They were …**																					
a	Scientists in the field of demography.	X	X	X	X	X	X		X	X		X	X	X	X	X	X	X			X	X
b	Other social scientists, including economists.		X			X	X					X	X	X	X	X		X			X	
c	Medical and public health researchers.						X					X		X		X		X			X	
d	Representatives of government agencies (ministries).		X		X	X	X		X	X		X		X	X	X		X			X	X
e	Representatives of social partner organisations (e.g., trade unions).													X				X			X	
f	Representatives of independent NGOs.						X		X													
7	**How did you deal with situations in which experts had different views about the values to be assumed?**																					
a	This did not happen.					X						X		X								
b	The dissenting experts were convinced with arguments to change their mind.																					
c	There was a (formal or informal) majority vote.		X														X				X	
d	We listened to the experts and their arguments, but then made the decisions in-house.	X		X	X		X		X	X			X	X		X		X				X
e	We applied a variant of the "Delphi" method.														X						X	

	Please specify how:																					
8	**Has your office commissioned any scientific studies from outside experts for the explicit purpose of helping with the definition of assumptions?**																					
a	Yes.	X		X											X							X
b	No.		X		X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	
9	**Has there been a systematic analysis of the errors of past official projections in your country?**																					
a	No.						X													X		
b	We looked at the past assumptions and studied which were too high or too low.	X	X	X		X			X	X	X	X			X			X			X	X
c	We not only studied the errors in past assumptions but also tried to analyze the causes of the errors.		X		X			X	X				X	X		X	X		X			
10	**Has there been an attempt to define storylines (plausible visions of the future conditions that would result in certain trends) behind the assumptions?**																					
a	No specific storylines were discussed.						X		X	X	X	X	X				X	X	X	X	X	X
b	We defined storylines for each component (fertility, mortality, migration) separately.	X		X		X									X	X						
c	We defined consistent storylines that bind together the future trends of at least two components. (Please provide more information about what you assumed, if possible.)				X			X								(X)						
11	**Do you think you have a near to optimal procedure for defining the assumptions on future fertility, mortality and migration, or is there a need for improvement in the process?**																					
a	No improvement necessary.															(X)						
b	Improvement required.	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12	**If you think there is the need for further improvement, in what direction do you think it should go? (You can choose more than one answer.)**																					
a	Involve more experts.			X		X	X	X		X			X	X			X		X	X	X	
b	Have a more systematic review of all the substantive arguments behind the assumptions.	X	X		X		X			X	X	X				X	X	X	X		X	X

c	Have more exchange with other national statistical offices about their assumptions.				X	X	X				X	X		X		X					X	X
d	Have more interactions with Eurostat in the process of defining national assumptions.		X				X				X			X			X			X	X	X
e	Have some structured interactions with the European demographic research community about the state of the art in our knowledge about future demographic trends.	X				X	X	X		X	X		X	X	X		X			X	X	X
f	Move toward stochastic/probabilistic projections.	X			X		X								X					X		X
g	Other directions (please explain).	X							X					X			X					

Appendix Table 2. List of Individual Arguments Defined for the Future of Life Expectancy

Force L1	**Changes in bio-medical technology**
Argument L1.1	Increased understanding of bio-medical ageing processes will allow us to develop effective anti-ageing strategies.
Argument L1.2	Breakthroughs in the understanding of carcinogenic processes will lead to substantial reductions in mortality from cancers.
Argument L1.3	Innovative medication will make hitherto life threatening diseases containable.
Argument L1.4	Improvements in surgery including transplants and implants will enhance longevity.
Argument L1.5	Unintended adverse consequences of new bio-medical technologies will outweigh their benefits.
Force L2	**Effectiveness of health care systems**
Argument L2.1	The cost of new treatments will be prohibitive to large segments of the population.
Argument L2.2	There will be some very effective and easily affordable new technologies.
Argument L2.3	Because of the growing elderly population there will be limited access and increased waiting times for treatment.
Argument L2.4	Society will be able and willing to afford expensive new treatments.
Argument L2.5	Progress in preventive medicine (screening, genetic testing) will lead to lower death rates.
Argument L2.6	Better and faster medical and health information dissemination will increase longevity.
Force L3	**Behavioral changes related to health**
Argument L3.1	Smoking prevalence will continue to decline.
Argument L3.2	Substance abuse (alcohol and drugs) will lead to more premature mortality and accidents.
Argument L3.3	Increased awareness of the importance of physical activity will lead people to exercise more.
Argument L3.4	Increased awareness of the importance of nutrition will lead people to adopt healthier diets.
Argument L3.5	Increased stress levels will impact negatively on health.
Argument L3.6	Increasing mental and social activities at old age will lead to greater longevity.

Force L4 Possible new infectious diseases and resurgence of old diseases

Argument L4.1 There will be a growth in infectious diseases leading to increases in overall mortality.

Argument L4.2 Increasing drug resistance to known infectious diseases will lead to higher mortality.

Argument L4.3 Increased capability of early detection and control will help to contain the spread and impact of new infectious diseases.

Argument L4.4 A major flu epidemic (avian or other) is likely to occur over the next 25 years.

Force L5 Environmental change, disasters and wars

Argument L5.1 Increased frequency and intensity of natural disasters (such as flooding and strong storms) will lead to increasing mortality in the UK.

Argument L5.2 Global warming will lead to the spread of malaria in Europe and result in higher mortality.

Argument L5.3 More intensive heat waves during summer will lead to higher mortality among the elderly.

Argument L5.4 Less extreme cold spells during winter will lead to lower mortality among the elderly.

Argument L5.5 Global climate change will lead to a decline in food production in certain parts of the world and, as a result, uncontrolled mass migration and conflicts will increase mortality in this country.

Argument L5.6 Because of the European Union, we will not experience wars in our country in the future.

Argument L5.7 A “clash of civilizations” will lead to major conflicts that result in lower life expectancy.

Force L2 Changes in population composition and differential trends in population subgroups

Argument L6.1 The UK “golden cohorts” born between 1925 and 1945 have experienced relatively high rates of mortality improvement throughout their lifetimes. The rate of improvement in overall population life expectancy will slow down as these cohorts reach advanced age.

Argument L6.2 For ethnic minority groups already resident in the UK and their descendants, mortality rates will converge to those for the indigenous population.

Argument L6.3 The majority of new immigrants will come from countries where mortality rates are higher than in the UK.

Argument L6.4 In the future, people who emigrate from the UK are likely to have lower mortality rates than the UK average.