Why business and socio-economic models often do not live up to expectations

Dr. Heinz Schild
CaseLab Bonn
Hopmannstrasse 6, D-53177 Bonn
Tel. 0049 228 323702
Heinz.Schild@CaseLab.Info

Summary

Looking back critically at about 30 business and socio-economic dynamic simulation models developed over the last 40 years, I notice that many of them did not quite live up to what was expected of them. Very often this was caused by organizational shortcomings during the implementation process. Other reasons were discrepancies between thinking and reality ("known known"), insufficient treatment of uncertainties ("known unknown") and/or the occurrence of events neither expected nor anticipated ("unknown unknown"). Based on the analysis of these shortcomings and on references in relevant publications, this paper presents some suggestions on how to avoid them.

Introduction

This study draws on almost 40 years' experience working as a consultant in the development and application of business and socio-economic dynamic simulation models ("models" in short), often based on the System Dynamics concept. It tries to answer the question as to why in a fair number of cases the expectations pinned on these models remained partially or even totally unfulfilled and to come up with suggestions as to how this situation could be improved.

The primary purpose of the models considered here was to support executives and their aides in forming an opinion during decision-making and planning activities with uncertain outcomes.

It is assumed that the expectations in connection with a model are not fulfilled when

- the client and his aides get the impression that working with the model and the delivered results does not substantially deepen their understanding of the relevant part of reality or does not contribute substantially to a better assessment of the consequences of alternative courses of action;
- the real data for a considerable part of the planning period deviate from the data computed by the model in a way that suggests that decisions or alternatives other than those based on the model would have better fulfilled the goals of the client.

Somewhat simplified, the first case can be ascribed to organizational shortcomings and the second to deficits of the model. „

Organizational shortcomings

As a rule organizational shortcomings are already visible while the model is being developed and applied. They prevent the modelling process from unfolding to full effect. In addition they may negatively influence the opinion of executives and their aides about the usefulness of models in general for some time to come.

The following table shows the main participants to be found in the model development process and the shortcomings which can result from their interplay. The table distinguishes 5 types of participants:

- The power promoter is the executive who issues the order for the development of the model and who in the end has to carry responsibility for the use to which consumed resources have been put. On the basis of his authority or his position in the hierarchy he can either convince or force the other participants to follow his intentions. The subject promoter reports to him about the progress of the model development process and its results, usually conjointly with the model developer.

  The function of the power promoter may be assumed by a steering committee.

- The subject promoter is either an executive in the line below the power promoter or a member of a staff. Often it will originally have been his idea to support a planning process or a decision-making task by commissioning the development of a model. He acts as interface to the model developer in all questions relating to the subject and should be well informed about the intentions of the power promoter as regards the model and about details of the relevant micro-world. He answers the model developer's questions and makes the necessary data available, to which end he may call on specialists to provide information. If data or information from external sources are needed it is his responsibility to obtain them. He controls the project in cooperation with the model developer, overseen by the power promoter.

  Occasionally the subject promoter acts as model developer (if the enduser concept is being applied). After final delivery of the model he may be in charge of its maintenance.

- The specialist is somebody in the organization who is knowledgeable about aspects of the relevant micro-world and able to provide the data needed for developing and running the model. If he is not positioned in the line reporting to the power promoter, it is the responsibility of the power promoter to reach an agreement with the specialist's superior to ensure that he will answer the model developer's questions and provide the required data.

  Often there is a demand for more than one specialist and for different specialists in different development and application phases.

- The user is the person in the organization who has to implement the decisions based on the model results. He may also have to take over the documented model from the developer and, after sufficient briefing, run the model to support future planning and controlling tasks.

  Sometimes more than one user is involved.

- The model developer is either an internal or external professional. An internal model developer is generally an adequately trained and experienced member of an in-house consulting group or a person hired specially for this purpose from a university, a consultancy or the supplier of the modelling software. An external model developer may be an independent consultant or an employee of a consulting company, the modelling software supplier or an academic institution.

  An internal developer normally requires less induction and meets with less resistance but may become a victim of inter-organizational jealousy or scheming. An external developer is dependent on the protection of a dedicated power and subject promoter. He will sometimes be tempted to play down inadequacies in the model or to make the model overly complicated so as to lock in the client.

The fields of the table "Organizational shortcomings" list frequently observed organizational shortcomings which may lessen or negate the benefits of the model. They are self-explanatory and therefore do not need further discussion. Hidden agendas which deviate from those that are openly declared or expected of their holders can prove extraordinarily harmful. This is especially the case when they are nursed by the power or subject promoter or the model developer. A serious problem can arise if specialists refuse to bring in their specialist knowledge and experience. This happens when they secretly reject the whole modelling exercise because they want to defend their "guru status" or believe they can deliver better results using their own spreadsheets. They justify their behaviour by referring to allegedly possible time and resource savings.

The power promoter and to some extent the subject promoter should be in a position to overcome organizational problems. The developer on the other hand is in a much weaker situation. He can threaten to terminate the co-operation but then risks loss of income or contractual penalties. A cynical but pragmatic course of action for him to adopt is to soldier on as long as the client pays for the model development and to learn from negative experiences how to minimize the chances of organizational shortcomings in the next project.

Sometimes an other more positive reason for the premature end of the model development is that the power and/or subject promoter get the impression they have learned enough already from the exercise. The model becomes too complex and therefore too confusing for them, and too many assumptions are applied to cope with the uncertainties of future developments. Therefore they prefer to rely on their gut feeling for the decision in hand. Whenever uncertainties are high and hard facts scarce, this can by all means be considered reasonable behaviour [1].

⬇	Power Promoter	Subject Promoter	Specialist	User	Developer
Power Promoter	no real interest in the model; hidden agenda; person replaced during the development of the model	doubts about the usefulness of the model and the intentions of the power promoter; insufficient access to the power promoter; lacking in loyalty	doubts about the competence of the power promoter	doubts about the commitment and the will to enforce; lacking in loyalty	insufficient access to the power promoter; unconvincing ap-pearance and presentations
Subject Promoter	unclear directives; not given enough resources or authority; inadequate gov-ernance	insufficient compe-tence and interest; burdened with other tasks; hidden agenda; person replaced during the devel-opment of the model	doubts about the competence of the subject promoter	doubts about the competence the subject promoter	doubts about the competence and commitment; insufficient participation in the validation process; mediocre or bad personal relation-ship
Specialist	insufficient co-ordination with the specialist´s superior	no power to enforce the collaboration; imprecise and in-complete briefing; insufficient quality checking of the contributed infor-mation	”spreadsheet-jockey“; superior´s negative attitudes toward the management of the model development; hidden agenda		doubts about the competence and loyalty; personal tensions
User	unclear orders; not given enough resources or authority; insufficient checks and reviews	insufficient and unconvincing briefing; bad feelings		dissonant problem view; feels insufficiently trained, equipped and authorized; disgruntled	insufficient expla-nations and train-ing; incomplete and incomprehensible documentation
Developer	intentions not communicated clearly; little interest for the model development and the results	unclear requests; insufficient partici-pation; disturbed personal relationship	doubts about validity and useful-ness of the approach; incomplete and misleading infor-mation; seen as competitor	no confidence in the model results; doubts about put-ting them into practice	not experienced or competent; unsure or arrogant behaviour; hidden agenda; person replaced during the development

[1] The results of research conducted at the Max-Planck-Institut für Bildungsforschung support this. See Gigerenzer, Gerd: “Gut Feelings“, Viking, New York, 2007

Modelling deficits

Damages resulting from model deficits may be substantial but can only be detected at a later point in time. In real life they are hardly every noticed, let alone publicized by the original power or subject promoter, who anyway might in the meantime have moved on to other positions and possibly does not consciously associate such damages with the model. If anyone becomes aware of them, it will most likely be the model developer; however, for contractual or other understandable reasons, he will not be interested in making them public. Therefore the following observations and considerations are largely based on the personal experiences of the author as developer and consultant, with references from the literature to underpin them.

The cognitive, methodical and procedural aspects that are preconditions for successful models are discussed in relevant publications [2]. The essential ones are:

- gain an adequate world view and clearly state the boundaries of the micro-world to be modelled;
- select the most effective and most efficient modelling method and programming language;
- choose appropriate filters and granularity and apply these to identify the elements of the model, the relationships between them, and their positive and negative feedback loops;
- program correct formulae for the tautological (definition) relations;
- identify, calibrate (whenever possible with professional statistical tools) and program the behaviour relations;
- determine values for the exogenous variables which map reality;
- wrap up the program code in a way which allows for easy handling and documentation of alternative model versions and runs;
- test the model program meticulously and undertake sufficient attempts to falsify it.

In the experience of the author, crass violations of these preconditions for successful model development are rare, with the exception of the selection of the modelling method and programming language. Since the advent of powerful spreadsheet software for personal computers, this tool is used too frequently for modelling in cases where specialized systems with graphical development interfaces and elaborated support of experiments [3] would be more suitable. The appliers of spreadsheet software normally do not have much professional know-how in the development of dynamic simulation models. Because they have little knowledge of pertinent micro-economic concepts they often

---

[2] See Sterman, John D.: "Business Dynamics: Systems Thinking and Modelling for a Complex World", Irwin McGraw-Hill, 2000

[3] For instance ITHINK from ISEE Systems in Lebanon NH, USA; VENSIM from Ventana Systems, Inc in Harvard MA, USA; POWERSIM from Powersim Software in Bergen, Norway; COSIDEO from Consideo GmbH in Lübeck, Germany

restrict their world view to elements for which they have (financial) data at hand. In addition they prefer simple tautological relations without feedback loops. Often they use exogenous variables where combinations of endogenous elements and behaviour relations with feedback loops would have been more appropriate [4]. For the exogenous variables they then insert values from other sources (which are normally based on different micro-worlds) or naively extrapolate them from available time-series or even simply guess them.

There seem to be three main causes for the majority of model deficits: overconfidence in information delivered by specialists or gathered from official sources, called the "known known", sketchy or plainly incorrect treatment of the risks, or the "known unknown", and insufficient precautions against the impact of the "unknown unknown".

- "Known known" characterizes situations where there is a serious discrepancy between what the model developer thinks some information might mean and what it stands for in reality. Sometimes this discrepancy is resolved in the development process, sometimes during the presentation of the model and its results to people who are not members of the core team, and sometimes not at all or too late to prevent flawed decisions or plans.
- "Known unknown" reflects the fact that data about the real world are always unknown to some extent; historical data are spoiled by measurement or observation errors, data about the future are per se uncertain. Therefore the developer cannot be sure about the actual start values of the endogenous variables, the values of parameters in behaviour equations and the development of exogenous variables in the future.
- "Unknown unknown" [5] describes events which have never been observed in the past and which the promoters, specialists, users and developers of the model are unaware might happen within the period up to the planning horizon. Only after the occurrence of such events can precursory signs be perceived and correctly interpreted.

Sometimes there is no obvious distinction between the three causes of model deficits. A wrong opinion may have the same negative effect as an unknown event. Or the probability that a known variable will have an effect is rated so low that this variable will be neglected in the model.

The following narrative serves to demonstrate the effects of "known unknown" and "unknown unknown". This particular case was chosen because it can be grasped intuitively and dates back far enough for all information about the model deficits to be available.

[4] Prof. Jay Forrester highlighted this problem in his guest lecture at the 9th International Symposium on Forecasting in Vancouver / Canada 1989.

[5] Taleb, Nassim Nicholas: "The Black Swan: The Impact of the Highly Improbable", Random House, 2007

In 1966, as part of an electric power study, a model which used the iconography of "Industrial Dynamics" [6] was developed and programmed in FORTRAN for an IBM 1130 computer. It simulated the combined deployment of nuclear, conventional thermal and hydro electric power stations in Austria with a planning horizon of 40 years. The aim was to estimate what capacity of nuclear power and pump storage stations would be required for the long-term optimum-cost generation of electric power in that country.

Looking back after 15 years, i.e. after less than half of the original planning horizon, one must concede that some of the assumed data and relationships deviated widely from real developments and that decisions based on the model results would have been wrong. What had happened?

The model assumed an oil price of about 3 US$ per barrel and only small increases were expected. In 1973, as a reaction to the Yom Kippur war, OPEC increased the price per barrel from 2.89 US$ to 11.65 US$.

Following the practice at that time, it was assumed in the simulation model that burned-out nuclear fuel would be bought back by the US Atomic Agency for their civil and military programs. This changed at the beginning of the 1970s. Burned-out nuclear fuel was not bought back any longer but had to be expensively re-enriched and stocked.

These two developments changed the costs of nuclear and oil-fired power stations and their optimal operating regimes significantly and reduced the demand for pump storage capacity which had been intended to exploit excess electric power. The estimation of the unknown values for the two known variables "price of oil" and "life-cycle costs of nuclear fuel" had been much too optimistic in 1966. They caused a model deficit of the type "known unknown".

Another event which was not considered at all during the model development had an even more dramatic effect. Despite the first oil shock the public mood turned massively anti-nuclear in the 1970s. This culminated in a referendum in Austria in 1978 which ended with a tiny majority for the opponents of nuclear power. This forced the government to ban the start of operations at the brand new nuclear power station in Zwentendorf. As a substitute a conventional thermal power plant burning imported coal had to be built. Within 12 years an "unknown unknown" event had made the model from 1966 and all decisions based on it totally obsolete.

Deficits in modelling "unknown" variables and the occurrence of unknown events are not always as crass as in the case outlined above. But experience shows that, more often than not, reality ends up significantly different from what had been assumed in the model, due to not handling the "known unknown" and the "unknown unknown" properly. By making small adjustments to plans and future

[6] Forrester, Jay W.: „Industrial Dynamics“, The M.I.T.Press, 1961

decisions whenever the "known unknown" and "unknown unknown" happens, it is frequently possible to reduce the damage caused by model deficits. Nevertheless all efforts should be undertaken during model development to supply the planners and decision makers with a view of the future which will differ as little as possible from actual future events.

The following sections try to give advice on how to deal with the "known known", "known unknown" and "unknown unknown" when developing dynamic business and socio-economic models.

"Known known"

The most common form of this problem occurs when the developer is not aware that the values of a variable which have been extracted from a document or delivered by a specialist mean something different from what he thinks or naively assumes. The difference might be due to an imprecise wording in the name of the variable, or an incomplete definition, or ignorance of disturbances in the data, especially if they are time-series. This happened for instance when a developer used a public statistic about the number of phone calls in the network of a national telecom carrier but was unaware that the numbers did not include the calls of the telecom and postal service employees and their family members.

A less common situation, but one that can still be observed, is when the developer sees a relationship or feed-back loop between elements which in reality does not exist or is only the proxy for a hidden or neglected part of the micro-world. This might be due to over-confidence on the part of the developer or specialist about his know-how and experience, or to over-interpreting the results of regression calculations.

Although these shortcomings look trivial they are hard to avoid. The standard recommendation is to set up a dictionary of meta-data for every exogenous and endogenous variable, but also for any relationship between variables used in the model, and have it double-checked by the subject promoter and the specialists involved. As a result of high external and/or peer pressure the world view of a project team's members may become lopsided, shutting out all information which does not fit in and developing a so-called "bunker mentality". This bunker mentality can run deep or be only a facade. In the latter case at least one project member is aware of the lopsidedness but keeps quiet for opportunistic reasons.

Unfortunately for the credibility of the model, these shortcomings are frequently not discovered until the final presentation and then possibly by some not so well-meaning superior. Occasionally they remain undiscovered and lead to decisions which are sub-optimal but not so obviously wrong that they draw attention. If a development team gets locked into in a bunker mentality, totally wrong decisions might be ensue.

"Known unknown"

In the case of "known unknown" the model includes variables whose values are stochastic. These variables may be start values of endogenous variables, parameters of behaviour relations or forecast values of exogenous variables. Often the users of the model results are not sufficiently aware of their stochastic character. For planning or decision-making purposes the developer contents himself with model results calculated by using only the mean values of the variables. He does not consider the latent uncertainties hidden in the results. Sometimes these uncertainties are deliberately played down by the power and subject promoters or the developer.

To avoid this more emphasis should be placed on the whole bandwidth of the values of the stochastic variables. However this requires that all participating parties search for appropriate forms of the behaviour equations and density functions of the stochastic variables and that the developer codes them properly. Based on these density functions, simulation runs should be executed which demonstrate clearly the bad risks implicitly hidden in the model due to the "unknown" character of the "known" variables.

In recent publications it has frequently been pointed out [7] that stochastic variables in business and socio-economic models cannot as a rule be properly represented by ”mild“ density functions like the Gaussian normal distribution. Instead, a wildly asymmetric distribution with "fat tails" should be applied [8]. Asymmetry is essential because in business and socio-economic models optimistic and pessimistic values do not deviate equally from the mean. Anticipated optimistic turnover values, for instance, will differ much less from today´s actual values than pessimistic ones. For asymmetric distributions the most probable or modal value is intuitively easier to grasp than the mean or median and should therefore be preferred in presentations.

Power functions are gaining in popularity for modelling the right-hand branch of right-skewed ”wild“ density functions [9]. For modelling the left-hand branch of right-skewed distribution functions a beta or gamma distribution is suitable. Often it is sufficient to approximate them using the left branch of a simple triangular distribution.

Sometimes variables can be realistically modelled using fractal generators [10]. Fractal generators also produce fat tails due to their characteristic self-similarity. Unfortunately, in some cases it does not make sense to apply them in a period-oriented simulation because the integration of all events in

---

[7] See for instance Bookstaber, Richard: "Demon of Our Own Design: Markets, Hedge Funds, and the Perils of Financial Innovation", Wiley & Sons, 2007

[8] Ball, Philip: "Critical Mass: How One Thing Leads to Another", William Heinemamm, 2005

[9] Sornette, Didier: "Critical Phenomena in Natural Sciences; Chaos, Fractals, Selforganization and Disorder: Concepts and Tools", Springer-Verlag, 2006

[10] Mandelbrot, Benoit, Richard L. Hudson: "The (Mis) Behaviour of Markets: A Fractal View of Risk, Ruin and Reward", Basic Books, 2004

a period smoothes out the extreme values and draws a too optimistic picture of the degree of uncertainty.

When analysing historical data with a goal to identifying and calibrating suitable density functions for behaviour relations, relatively long time-series or extensive data are required, or „wild“ functions must be enforced if this is recommended in the literature. Short series often do not include extreme values, or these are rejected as outliers by the calibration software. The computerization of ever more business and socio-economic fields permits the cost-efficient collection and storage of vast quantities of data, [11] which facilitates identifying and calibrating wild density functions. This should reduce model deficits of the "known unknown" type and will make it easier to detect extreme values. On the other hand one can expect the frequency of extreme values to increase due to the growing complexity and networking of the business and socio-economic world.

Analogous considerations must be borne in mind when forecasting the values of exogenous variables to be used as model inputs. As a rule the semi-automatic time-series analysis and forecasting methods on offer today assume errors with a Gaussian distribution and so produce distribution functions that are too ”mild“. This is one of the explanations why sophisticated forecasting methods fare no better than simple ones [12] in the M-Competitions.

A combination of "known unknown" and "unknown unknown" is found in variables which display bifurcations [13], i.e. their values may lie on different levels and can suddenly jump from one level to another. Well-known examples can be observed in the field of finance where a rumour can trigger a "jump" in the behaviour of a crowd of people. Common distribution functions are not suitable for modelling them and one must fall back on intervention variables.

As already mentioned, running the model with mean values only is not sufficient. To gain an impression of the degree of uncertainty surrounding the model results and thereby the risks attached to decisions based on them, the calculation must

- either produce the so-called behaviour space, based on runs with a combination of modal, acceptably extreme and in-between values of the stochastic variables [14],
- or be prepared as a Monte-Carlo simulation.

Extreme values are rated as just acceptable when the probability of their happening is between 10 and 5 per cent. This probability should be in line with the subjective risk preferences of the planner or decision maker. To elicit a number for their subjective risk preferences it can help to take bets on

[11] Ayres, Ian: "Super Crunchers: Why Thinking-by-Numbers Is the New Way to be Smart", Bantam Books, 2007

[12] See www.neural-forecasting.com/m-competitions.htm

[13] Poston, Tim, Ian Stuart: "Catastrophe Theory and its Applications", Pitman, 1978

[14] The "BehaviourSpace" feature of the software system "NetLog" shows how this can be handled in a comfortable manner.

events in their personal environment [15]. Runs with in-between values are only recommended when bifurcations are expected.

Model runs with different combinations of values for the stochastic variables or parameters communicate a "tangible" impression of the uncertainties inherent in the simulated development of the micro-world. One can present the findings in the form of naive implications:

> If everything works fine, expect result A. But if the extreme value occurs for variable X, expect an x% deterioration in the result. If in addition variable Y takes on its extreme value, the result will deteriorate further by xy%.

One should restrict the presentation of the combinations to those extreme values for which one believes that they will have a really strong negative influence on the result. Otherwise one may get entangled in too many influences and loose the ability to analyse and communicate the results convincingly. The analysis of the effects of extreme values often leads to successive adaptations of the original plan until an alternative is found which produces a spectrum of results for a reasonable combination of extreme values that is acceptable to the decision makers.

It is more demanding to calculate and communicate the risks involved in decisions based on a Monte-Carlo simulation. Even when wild distribution functions are applied, which is not at all easy using the relevant standard software, the risks resulting from extreme values may be underestimated. This happens because the density functions of the stochastic variables are treated as mutually independent, whereas in real life the expectation is that if one thing goes wrong, other things may get worse. One can try to overcome this by including conditioned control loops in the model. This often requires working with negative feed-back loops in the first few simulation periods and later converting these into positive feed-back loops. Doing this introduces an additional and hard-to-communicate level of uncertainty or arbitrariness. The standard graphical presentation of the results of a Monte-Carlo simulation showing bands of increasing width is hard to interpret and does not highlight the effects of extreme values.

To avoid model deficits due to "known unknown" effects the evaluations should not be restricted to mean values or medians but should be performed with combinations of modal and extreme values and by applying wild density functions.

"Unknown unknown"

In conventional model development one rarely watches out for events which are not obvious or evident but which would nevertheless strongly influence simulation results if they occurred before the planning horizon was reached. On the contrary: for pragmatic reasons the boundaries of the micro-world are often drawn in such a way that data for the included elements are readily available or can

[15] See the procedures of the descriptive preference theory, especially the theory of Kahneman-Tversky as outlined in Eisenführer, Franz, Martin Weber: "Rationales Entscheiden", Springer-Verlag, 2003

be delivered by the specialist assigned to the project. This allows the subject promoter and the developer to avoid some of the organizational problems outlined above and to finish the work within the proposed and agreed timeframe. The downside of this tactic is that the model might produce results which lead to wrong decisions.

To avoid model deficits caused by "unknown unknown" effects, it is highly recommended that sessions be held regularly in all phases of the model development to spot any new trends, tipping points and outliers which might crop up in the periods up to the planning horizon. Members of the model development team should use these sessions as brainstorming events to scan and discuss the external and internal news for hints of hitherto unexpected, but relevant, future events. Sources of unexpected events are often changes of ownership (for instance following a hostile take-over), national or international politics (for instance the cancellation of subsidies after a change in government, or the results of a public referendum), scientific or technological break-throughs or setbacks, or natural disasters.

Sometimes these events are indicated by micro-trends [16]. It is worthwhile looking out for these specifically and trying to project or extrapolate them and estimate their influence on the development of the relevant micro-world. Macro-trends about which professional futurologists have been making noisy pronouncements are normally already on the radar screens of subject promoters and developers and will have been automatically built into the model. Unusual information from unorthodox sources should be included in the active search for hints about as yet unknown, but possible, future events. Unorthodox sources can be popular science and literary [17] publications.

Information about developments in other specialisms which may be of importance for the micro-world of the model can often be garnered from popular science publications, obviating a steep learning curve. Bold or even cranky speculations about future possible developments are especially valuable. The authors of these publications often have outstanding didactic skills, making it easy to digest the knowledge offered.

Literary publications, including science fiction stories, try to depict consistent (”stimmig“ in German) sets of situations and developments. To this end their authors may introduce features and events other than those known to exist or familiar to the reader. Science fiction authors from the first half of the last century, for instance invented radiation phobia to demonize the consequences of the physical and chemical discoveries of their time and initiated public action groups to militate against such discoveries. Suitably interpreted, these stories could have pointed towards the emer-

---

[16] Penn, Mark with E. Kinney Zelesne:“Microtrends: The Small Forces Behind Tomorrow´s Big Changes“, Allen Lane, 2007

[17] Hörisch, Jochem: "Das Wissen der Literatur", Wilhelm Frank Verlag, 2007

gence of the anti-nuclear movement [18] in the 1970s and avoided the "unknown unknown" deficit of the electric power case outlined earlier.

One cannot treat the analysis of trends or unorthodox literature as serious sources for the construction of density functions for future events. A better option is to implement these events as intervention variables which generate discontinuities in behaviour functions or forecasts at guessed periods. If such a conjectured event could have a substantial influence on decision making, it should be treated as a dimension of its own in the behaviour space in combination with the other variables and their modal and extreme values.

Although there is always a high possibility that these conjectured events might be "red herrings", one should not be too critical. The primary purpose of accounting for them in models is to delineate the risks which may be hidden in the simulation results and not to increase their numeric precision. However intensively and systematically the searches in standard and unorthodox sources are carried out, one must bear in mind that unexpected events will still occur in the period up to the planning horizon. The probability of this happening increases more than linearly the further you look into the future. The developer must stress this point when presenting the model results. A somewhat risk-averse but pragmatic approach is

- to run the simulation up to the intended planning horizon to get an impression of what might happen in the long term as well as to check the over-all plausibility of the model;
- but to base any decision on the results of only the first third of all periods.

It seems that the extent of "unknown unknown" events and the erroneous decisions and plans in connection with them is growing. Reasons for this may be the permanent increase in the speed of business and social processes due to greater networking and to higher automation based on ever more sophisticated algorithms [19]. Therefore it is recommended that model teams design and implement coarser, mathematically simpler, but more robust models [20], include micro-trends and information from unorthodox sources and opinions more frequently in their considerations, and put more emphasis on the presentation of the risks attached to the model results.

Retrospect

A subjective analysis of the 30 models mentioned at the beginning of this paper yields following findings:

---

[18] Weart, Spencer W.: "Nuclear Fear: A History of Images", Harvard University Press, 1989

[19] See "Business by numbers", The Economist, September 15th, 2007

[20] In doubt prefer event- and agent-based simulation over period- and macro-based simulation. Event-based simulations normally do not improve the accuracy of predictions but add to the understanding of the micro-world and thereby improve the mental models of the participants. See Grimm, Volker, Steven F. Railsback: „Individual-based Modeling and Ecology“, Princton University Press, 2005

- About 23% of the models can be rated as a full success; no serious organizational problems occurred during the development process and no model deficits became known.
- About 77% had organizational problems and/or model deficits. This does not mean that all of these models were totally useless or that they totally misled the decision makers, but only that real developments differed significantly from simulated ones. In some cases the model development was prematurely aborted because the power and/or subject promoter got the impression that their gut feelings could cope with the problem in question at least as well as the model exercise.
- Referring to all models
    - 43 % had at least organizational problems,
    - 3 % had at least deficits of the "known known" type, [21]
    - 43 % had at least deficits of the "known unknown" type,
    - 23 % had at least deficits of the "unknown unknown" type,
    - 23 % had exclusively organizational problems,
    - 17 % had exclusively deficits of the "known unknown" type,
    - 10 % had exclusively deficits of the "unknown unknown" type.
- Older models more frequently showed problems or deficits than more recent ones.

Of course these numbers are only based on a relatively small sample and on subjective ratings. Although these findings should not be over-generalized, they may help to explain to what extent and for what reasons business and socio-economic models often do not live up to expectations.

[21] When ”known known“ errors happened they were normally taken care of in the course of the model development or else never became known. Only in one case was a model with ”known known“ deficit developed which became known. It was due to the bunker mentality of the project team caused by external pressure.