Modelling tools for policy Analysis: dimensions of reliability

Wolfgang Britz, Thomas Heckelei
133rd EAAE-Seminar in Chania, June 14-15, 2013
Why reflect on reliability of modeling tools?
- Reliability – scientific perspective
- Reliability – client view
- Institutional aspects
- Conclusions
Why reflect on reliability?

Podhora et al. 2013

In FP6 and FP7
99 projects claimed to develop tools for policy analysis in the context of agriculture + environment

Uptake of quantitative tools in EU policy IAs very limited despite desire on the client side to increase the use

⇒ Tool use by clients might depend on criteria not in the focus of scientists ... reliability as a key one?

Britz/Heckelei: Reliable models, 133rd EAAE-Seminar Chania 14-15 June 2013
In scientific evaluation, reliable means

- "Sound": appropriate theoretical underpinning (behavioral model) state-of-the-art methodology, data choice and treatment, transparent and reproducible...
- "Fit and significant": proven relevance to explain observed phenomena

⇒ Checked by model of "impartial double blind" peer review: de-personalized, (seemingly) objective criteria for reliability
Reliability from a client-perspective different: reliability here more likely goes in the direction of

“The ability of an apparatus, machine, or system to consistently perform its intended or required function or mission, on demand and without degradation or failure” (business directory)

Why these demands: because they have to rely on it as in

to depend on (synonym)
Clients do not communicate with models but with people.

Our methodologies and especially our models are grey to black boxes.

Do we meet our deadlines? Keep confidentiality?

Are we self-critical, but able to convincingly defend key results?

Deliver the expected results?

Knowledgeable in domains which matter for clients, for example specific policy implementation or current market developments?

Clearly express our findings?

Clients often thus judge about our tools by judging about us:

Britz/Heckelei: Reliable models, 133rd EAAE-Seminar Chania 14-15 June 2013
Trustworthiness hinges on

1. External evaluation
   ⇔ peer reviewed papers
   ⇔ scientific standing

2. Repeated positive experiences

=> Clients favor stable relations
=> Models/tools in this context hard to „de-personalize“
⇔ different from „objective“ double blind review
Key for us for scientists:
- Competition for "brilliant" ideas
- Concepts matter more than results and far more than operational implementation

Clients tend to be more conservative, rather use
- what has proven to work
- what was repeatedly peer reviewed, especially in applications

=> Policy relevant tools are often not at the methodological forefront
Often key for client: how fast are usable results available

Interlinked challenges:
- Model up-to-date and operational
- Efforts needed to perform requested analysis
- Quality of results
Up-to-date historical data:

- key for **solid ex-post analysis** and **model specification**
- can involve large and repeated efforts (outlier removal, internal consistency, plausibility, dealing with data gaps)
- data e.g. on policy instruments might not be readily available from providers
- needed resolution changes with policy question

=> Obvious field for co-operation across tools and models, take GTAP as example
Up-to-date future reference ("baseline"):

- typically for partial equilibrium models (AGLINK-COSIMO, FAPRI, ESIM, CAPRI, AG-MEMOD ...)
- based on Delphi process
  - often involving the client
  - reflects expectations about macro-economy, market developments
- and factors in already decided future policy changes

=> Co-operation already institutionalized at coarser spatial resolution for key markets (FAO, OECD, EU)
Tools operational for task?

- Response time:
  - Licenses up-to-date
  - Parameters up-to-date
  - No compilation and run-time errors

- People available who can successfully apply model

- Policy targeted:
  - Current model structure suitable to deal with targeted new policy question

Britz/Heckelei: Reliable models, 133rd EAAE-Seminar Chania 14-15 June 2013
New policy questions

Policy targeted

Response time

high resolution (regions, markets, policy instruments)

versus savings in time and monetary costs to change data + model structure

explicit representation of policy instruments: requires structural changes for new instruments

versus faster, more generic representation of policies
Quality management

Not always on our agenda, but important for clients:

- Model code documented and systematically tested
- Sensitivity experiments, ex-post evaluations
- Code accessible for external reviews

⇒ Expensive and neither scientifically rewarding nor entertaining

⇒ Requires expertise typically not available with academic economic groups

IT implementation of model/tool:

- State-of-art IT management: software versioning, stable releases, ticket system...
- GUIs can help to avoid steering errors and to efficiently check results

Britz/Heckelei: Reliable models, 133rd EAAE-Seminar Chania 14-15 June 2013
Clients are increasingly pushed to use quantitative tools, e.g. by guidelines for Impact Assessment and RD evaluation...

Complex, policy relevant tools cannot be built quickly from scratch, clients partly „locked in“, for example

- GTAP as the sole available global data base for CGEs
- Only a few Europe-wide models of CAP relevance (MAGNET, ESIM, AG-MEMOD, AROPAJ, CAPRI ...)

High market entry costs, especially if trust building phase is considered
Critical transition phase:

- Some **successful and relevant applications needed** to build up trust: „proof of concept“
- Shift in cost structure **from development to maintenance and quality management**
- Shift in key competences **from scientific expertise to successful consulting**

How many researchers really want to take the necessary steps?
Institutional solutions for tool use

**Open tenders**

**Pros**
- competition
- tools targeted to policy

**Cons**
- high transaction cost
- asymmetric information
- trust?
- opt out: perhaps no suitable tenderer

**In-house capacity**

**Pros**
- strong control
- short response time
- procedural knowledge

**Cons**
- right people to attract?
- cost occur all the time
- credibility of tools?
- limited scientific environment
In-house capacity plus external support

- Keeps original developers on board
  - Training of and support to in-house staff if more complex changes are needed
  - Helps to avoid credibility loss:
    - Tool also applied for other clients and questions
    - Developers will strive for peer reviewed publications

- Combined with advantages of in-house capacity

- Example: IMAP Modelling platform for agricultural policy analysis (production structure, markets and trade) of the JRC/IPTS
Different aspects matter for scientific and policy relevant performance of tools.

Strategy for people, maintenance + up-dating, and quality management are key to move tools into application phase.

Institutional solutions matter.

... observed tool under use might also reflect preferences and job context of scientists (“publish or perish”, teaching, low incentives for what we saw is necessary...)

Britz/Heckelei: Reliable models, 133rd EAAE-Seminar Chania 14-15 June 2013