# Risk in the modelling process

Prepared by: Michael Clarke, PwC
Email: Michael.t.clarke@nz.pwc.com

Date: October 2016

# 1 Executive Summary

Models of some form are used and relied on heavily in almost every organisation. This could be as simple as a model to determine how much to invoice a customer, or as complicated as a prophet model that estimates policyholder liabilities. Based on PwC research from the US, executives underestimate how many key models their organisation relies on. For this reason it is incredibly important to understand the risks involved in the modelling process and be able to clearly communicate them to users.

There is an old saying from statistician George Box that "all models are wrong, but some are useful". Wrong because you cannot approximate reality perfectly. Useful because despite not approximating reality perfectly, models provide valuable insight and help inform decisions.

It is the usefulness of a model that is important. If there are errors in the model, then it could lead to poor decisions. There is a significant amount of research that looks into real world modelling errors that have had disastrous impacts. Most organisations have a mix of policies, procedures and controls to reduce this risk. However, the reality is there will always be errors that are not corrected.

The following table summarises the areas of risk that are featured in this paper, as well as how some of these risks can be mitigated:

| Part of process | Risk area | Potential mitigation |
|---|---|---|
| **Model build** | • Technical mistake<br>• Inefficient model<br>• Complex model | • Specification phase<br>• Test and review<br>• Training and guidance |
| **Model use** | • Incorrect inputs<br>• Interpretation of outputs<br>• Manual updates<br>• Transitioning<br>• Links to other models | • Reasonableness checks<br>• Good modelling practices<br>• Documentation<br>• Controls around model links<br>• Using the appropriate software |
| **Model adjustment** | • Implementing adjustment<br>• Adjusting working part of model<br>• Layout poor for adjustment<br>• Version control | • Documentation/understanding<br>• Specification phase<br>• Re-performing test/review<br>• Model rebuild |

The most important mitigation for reducing modelling errors is allowing more time. Time is needed to specify a model, time is needed to build a model properly, time is needed to review a model, time is needed for a new person to learn a model, and finally time is needed to adjust a model.

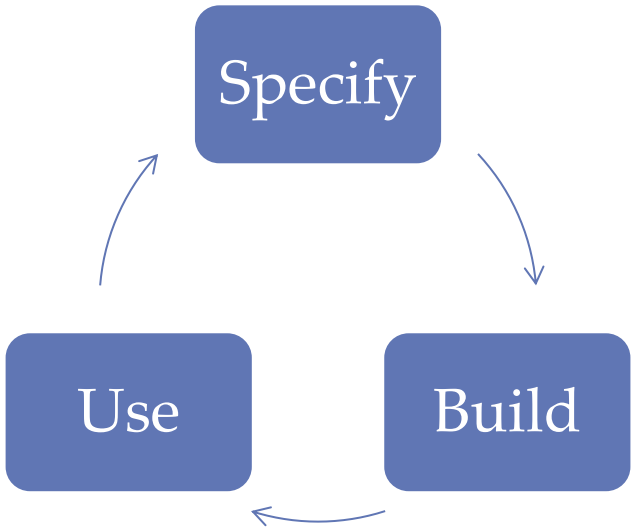
# 2 The Modelling Process

## Creating a model

The modelling process begins when you first start thinking about the model and finishes only when you no longer wish to use it. Possibly with many iterations of the model in between. The process can be separated into three distinct phases:

1. Specifying the model – how will it work, what will the calculations be, what software will be used, what data is used. The importance of this phase should not be underestimated. Without a clearly specified model, the following steps are much more prone to error.
2. Building the model – implementing what has been specified. If well specified, then with sufficient expertise this stage should be relatively straightforward.
3. Using the model – getting outputs as well as operating and maintaining the model.

Creating and using a model should be a circular process as follows:




The circular nature is important, as if an adjustment is needed to the model, then this adjustment goes through the specifying stage. In particular, an important consideration will be whether the existing model should be adjusted, or whether a new model should be built.

The phases of the modelling process can be compared to a construction process. The diagram below compares the modelling process to the construction and use of a house. This example will be used throughout this paper.

| | |
|---|---|
| Specifying | Drawing up plans for the house |
| Building | Constructing the house |
| Using | Living in the house and maintenance |

## Types of model risk

Emanuel Derman (Goldman Sachs) in his publication on model risk identifies a number of types of model risk. These are:

- Inapplicability of modelling – not applicable to the situation being modelled
- Incorrect model – for example incorrect assumptions and approximations
- Correct model, incorrect solution – such as a technical mistake
- Correct model, inappropriate use – used in ways not intended
- Software and hardware bugs
- Unstable data

Alternatively, you can consider model risk by mapping the risk types to the different phases of the modelling process. This can be useful because risks can be pinpointed easily while following the three phases outlined earlier. Given the circular nature of the modelling process, risks can also occur when adjusting a model. These risks would be:

1. Model being specified incorrectly
2. Error in building the model as specified
3. Model being used incorrectly
4. Model adjustment errors

Notice that this list is equivalent to Derman's:

| | |
|---|---|
| Model being specified incorrectly | • Inapplicability of modelling<br>• Incorrect model |
| Error in building the model as specified | • Correct model, incorrect solution |
| Model being used incorrectly | • Correct model, inappropriate use<br>• Software and hardware bugs<br>• Unstable data |
| Model adjustment errors | • All of the above |

This paper will not focus on a model not being useful due to it being a materially incorrect representation of reality; this is a much more complicated issue. Rather, it will focus on the other ways to ensure a model is correct and hence useful.

# 3 Model Build

The model build phase involves taking the model as specified and building it to these specifications. The more detailed and comprehensive the model specification is, the easier the model is to build. The major decisions, such as which software to use and the layout of the model, should already be made by the time the model build begins. It is costly to change these decisions later, so time should be taken to consider them carefully.

## Areas of risk

### Technical mistake

The major risk in building the model is that a technical mistake occurs. If this technical mistake goes unnoticed, an organisation may be unaware it is using materially incorrect outputs. This risk could manifest in many ways. In Excel for instance, it may be a cell that is referenced incorrectly or a calculation that is not implemented as specified. In programming languages, it may be due to using the wrong variable. These errors are often easy to fix.

> An example is building a home that has one gap where it is not weathertight, which causes leaking.

### Inefficient model

An inefficiently built model takes a long time to run. While this does not necessarily lead to materially incorrect outputs, it can cause problems due to the time required to run the model. Further adjustments to the model will decrease efficiency further. Poor choices in setup or programming often result in an inefficient model, and once built this way, it is hard to make the model more efficient.

> An example is building a home that is not energy efficient. It is not going to cause your house to fall down, but it will result in higher costs later on.

### Complex model

Models will often be complex if they represent a complex process or idea. Sometimes this complexity cannot be avoided. However, in a number of situations, unnecessary complexity is added during the build phase. With an Excel model, this could involve using long formulas, when a simple one will do the same job and be easier to understand. Needless complexity can cause inefficient models and makes transitioning a model to someone new more difficult.

> An example is building an unnecessarily complicated frame inside the walls.

## Reducing risk

### Specification phase

The more detail that goes into the specification phase, the easier a model is to build. Detailed specification allows the model builder to focus on implementing the model as specified, rather than trying to design the model at the same time. A model that is well specified is also likely to be less complex and more efficient than one which is designed and built at the same time. Consideration in the specification phase of the best way to design the model to accommodate further adjustments can also reduce potential inefficiency and complexity.

This aligns with the planning and designing phase of constructing a house.

### Test and review

A key part of the model build phase is testing and reviewing the model. There can be two levels of review – a technical review and a high-level review. The technical review is a detailed review often undertaken by an experienced analyst. In an Excel model, this would consist of performing a cell-by-cell review of the calculations.

A high-level review is generally undertaken by someone more senior, and consists of checking that the model is implemented in line with the specification and that the outputs will be useful.

Involving two more people in the test and review of the model has the additional benefit of ensuring additional people understand how the model works and can work with it if necessary. Testing can also include stress testing and backtesting.

The technical review is similar to the role of regular building inspections in constructing a house, checking the house has been constructed correctly. The high-level review is similar to tests the house owner will perform, such as considering whether it feels right to them.

### Training and guidance

The person building the model should have adequate experience and training in the area. There are a number of external training programmes offered by providers. These are often general, as they cater to a range of audiences rather than to a specific organisation. More specific training can be provided by senior analysts in the organisation. These staff should be familiar with the organisation's models and can provide more targeted training. The materials for these trainings can often be reused as new staff join the organisation.

In addition to being well trained, analysts need others around to ask questions if they are unsure about the best way to build something and to have a sounding board for ideas. Leaving an analyst to create a model independently can often result in a model that is poorly built.

Builders typically undertake reasonably long apprenticeship programmes before they start to build a house. There are also buddy/mentor systems in place if any additional guidance is needed.

# 4 Model use

Using a model is often where the biggest mistakes occur in the modelling process. There are two main reasons for this:

- Firstly, some models are used over a number of years, which introduces more opportunities for errors.
- Secondly, a model can be used by a number of different people, some of whom may not understand the model.

Research from the Delft Spreadsheet Lab indicates that the average Excel model is used by 12 different people over 5 years. There is a high chance some of these people are not familiar with the technical details in the model.

The best way to reduce the risk in the model use phase is in the model build phase. Once the model is built, it is often too late to target these areas.

Imagine living in a house after only ever living in the wild.

## Areas of risk

### Incorrect inputs

Invariably, when you use a model regularly, you will need to change some inputs each time it is run. This could include entering simple information such as dates, or entering significant amounts of data such as claims histories. This can be a key area where errors can occur, including mistypes or using the wrong data.

While unlikely, an example is feeding gas into an electric stove.

### Interpretation of outputs

The main user of the model will understand what the outputs mean and the limitations of the model. However, sometimes decision-makers who receive the outputs will not. This could result in an organisation making incorrect decisions or relying on a model that was intended for a different purpose.

An example is using your house as a shop to sell goods. It probably will not work well, as your house is not set up to function as a shop or in an appropriate location.

### Manual updates

Some models will require manual updates, as well as entering inputs/data. In the case of an Excel model, this could involve copying and pasting from one sheet to another or requiring a pivot table to be refreshed. Wherever there is a manual update required, there is the chance that something can go wrong, such as missing a required step.

In the house example, this could be similar to locking the door. If you forget, your house may be robbed, and it may take some time to realise what was taken.

### Transitioning

A significant area of risk is when a model transitions from being used by one person to another. This can result in a range of errors if there are parts of the model that the new

person does not understand, especially if manual updates or inputs are required. Model transition is generally unavoidable, as staff leave or become more senior.

When a house is sold to new homeowners, the new owners do not know as much about the house as the original homeowner. Operating manuals help the new owners 'operate' the new house.

### Links to other models

Where a model relies on other models, there is an increased risk of errors. In an Excel model, errors can occur through external links. In a software programme, errors can occur if another model is required to be run first. If the models that are relied upon are not updated, or are changed without updating your model, then this will likely create an error in your model. Often it goes unnoticed until later.

An example is using someone else's driveway. This places reliance on them to keep their driveway clear and in good condition, so you are able to use it.

## Reducing risk

### Reasonableness checks

A simple but effective way to mitigate the risks of mistakes when using the model is by performing reasonableness checks on outputs. The easiest way to do this is to compare the model outputs to those produced the last time it was run. Any changes should be reasonable with reference to how the inputs and data have changed. When a model is run for the first time, the model will have been reviewed and tested already, so reasonableness checks are less important. Still, the model outputs should be close to what was initially expected. Other reasonableness checks should test that the results by different segments are sensible, and they trend sensibly over time.

### Good modelling practices

There are a number of good modelling practices that can be followed when building to help model users, for example separating inputs, calculations and outputs. This makes it easier for the model user to know where inputs are located.

Ideally, an organisation will have a policy document for good modelling practices. For Excel, the FAST (flexible, appropriate, structured, transparent) standard includes a number of good modelling practices (http://www.fast-standard.org/the-fast-standard/). Building models in a consistent way and following good practices makes it easier to use models correctly, as well as to transition models to new people.

If a house is constructed using standard practices and with a sensible layout, it will be more comfortable for those who live in it. For instance, the kitchen is normally built near the dining room.

### Documentation

Documentation with any model needs to answer two key questions:

- How do you run and update the model?
- How does the model work?

This should be documented during the model building phase. Most of this information can be documented based on the model specification without much additional work. Ideally, the

documentation should be found with the model. In the case of a spreadsheet, documentation could be stored as an additional sheet of the model.

Documentation of a house consists of the plans for a house, as well as any operating manuals.

### Controls around model links

In an ideal world, models would not need to link to other models. But this is not practical in most cases. The best way to mitigate risks is to include error checks and controls around any links. For an Excel spreadsheet, a record should be made containing each link and the date it was last updated. Any spreadsheet that is linked to should also include a record of dependent spreadsheets. Then when a model is changed, the models linked to it can be updated easily.

An example is using a shared paddock with another house. In order to make effective shared use, records should be kept and plans should be made of who will do what maintenance when.

### Using the appropriate software

One of the decisions made in the specifications stage is which software should be used to build the model. This is important when it comes to using the model, as a poor software choice can make for a cumbersome model. For instance, in Excel, if you are using significant amounts of data, the model becomes difficult to use. Likewise, using programmes that are not widely understood, such as SAS, means a model can only be used by a limited number of people.

When building a house, the correct kind of building materials need to be used. For instance, a large house with multiple stories needs materials to that can support larger loads.

# 5 Model adjustment

If a model is in operation long enough, then invariably it will need adjustments or additions. These could be necessary market changes, or if someone inside the organisation decides they want to see more or different outputs. This in itself can be a key area of risk, as often insufficient time is allocated to model adjustments.

If a model needs to be adjusted, then this adjustment should be treated as if a new model is being built and should go through a specification phase. The specification phase needs to consider whether an adjustment is the appropriate way to implement the required change, or whether a new model needs to be built.

## Areas of risk

### Implementing adjustment incorrectly

As with building a model, technical mistakes can occur when adjusting a model. This could be a mistake in the original, working part of the model. Or more likely, it could be in the adjusted part of the model. Realistically, less work will be put into adjusting a model compared to building one, which means there is more room for error. However, those using the results from the model will expect the complete model to be free from error. This disconnect in accuracy and confidence in outputs creates risk.

An example is adding a second storey to the house that is not weathertight and causes leaking.

### Making a working part of the model incorrect

Before making adjustments, you hopefully have a working model. Whenever you adjust a model, there is a chance that you will change something that will make the original model dysfunctional. This could be because you need to update the original parts of the model in order to make new parts of the model work.

An example is performing some alterations to the house, but changing the wrong wire so the electricity no longer works.

### Model being laid out poorly for adjustment

When the model is originally specified, the design of the model should reflect the requirements at that time. When an adjustment needs to be made, this can mean that the model no longer has the optimal layout. If you are not rebuilding the model, then this is often unavoidable, and it can result in a model that is harder to use, inefficient to run and more prone to error.

Consider adjusting a house by adding on a new bathroom. Given available space, the only place for anew bathroom is next to the front door. The finished product has an extra bathroom, but it is not as useful as it could be elsewhere in the house.

### Version control

Where there are multiple versions of the same model which are all slightly different, there is a risk that the wrong version is used or adjusted. This is particularly important where the new version is only a slight adjustment from the previous version. Where there are significant changes, then version control will be easier.

## Reducing risk

### Documentation/understanding

When adjusting a model, one of the most common reasons for error is that the person adjusting the model does not understand it fully. There is less of a risk where the person who built the model also adjusts it, as the builder should already understand how the model works. Time should be taken for the adjuster to become familiar with the model. The understanding required here is more detailed than the understanding when transitioning to use the model. Detailed documentation that answers the question of 'how the model works' is important to bring the adjuster up to speed.

### Specification phase

As with the original model build, the specification phase is important to mitigate risk. At a simplified level, the specification should consider:

- Is an adjustment appropriate or should the model be rebuilt?
- Where should any adjustments be made?
- How is it best to build the adjustments?
- What does it mean for the existing parts of the model?

In any additions or alterations made to a house, plans will be drawn up and designs made before any construction is undertaken.

### Re-performing test/review

Just like when the model was originally built, a review should take place after an adjustment is made. Particular focus should be made on the part that was adjusted, as this is where there is most likely to be a problem. The original part of the model can be tested relatively easily by checking that the outputs are similar to the outputs before the adjustment.

The building inspector needs to sign off alterations, just like they did with the original construction.

### Rebuild model

By following a full model building process again, rather than just performing an adjustment, you are more likely to reduce the risk of error. By designing the model from scratch, the model can be made optimal for its purpose including the adjustment, rather than adjusting a design that was optimal for the original model. Rebuilding a model can be easier than building from scratch, as there are elements of the current specification/model that can be used again. Although, consideration should be given to whether the process can be improved in the rebuild.

There are some parts of models that are more difficult to adjust than others. It is also often not practical to fully rebuild a model, as it takes significantly longer than to adjust an existing model.

Where a model has been adjusted a number of times already, it may be necessary for a new model to be built instead of further adjustments. It is difficult to know where the rebuild point is. When a model is cumbersome to use and is set out poorly for its intended use, then it is definitely time for a rebuild.

On some occasions a house will be moved or destroyed and then another house built on the same site. This could include when the site is subdivided.

# 6 Conclusion

Models in some form are an important part of any organisation as they help inform key decisions. Just like in the process of constructing a house, there are a number of risks in the modelling process. These risks can occur at any stage of the process from specifying the model, through to using and adjusting the model.

Those using models as well as those using outputs from models need to be aware of potential risk areas and ways to mitigate them. This will make the whole process less prone to material errors. In the end, this results in more useful information being available to inform decisions that are made by organisations.