

Using evidence well

Joann Wilkie and Angelia Grant¹

We all rely on evidence. All decisions are based on past experiences, theories and expectations. In a policymaking context, the use of evidence can be challenging; in some cases there may be a plethora of evidence, in others a dearth; evidence may be ambiguous, sometimes partly contradictory, or not directly applicable to the policy under consideration; and there may be time constraints that restrict the gathering of evidence. Using evidence well requires an understanding of how it is produced and used in the policymaking process.

¹ The authors are from the Macroeconomic Policy Division, Australian Treasury. This article has benefited from comments and suggestions provided by Jenny Gordon, Andrew Leigh, Tony McDonald and Terry O'Brien. The views in this article are those of the authors and not necessarily those of the Australian Treasury.

Introduction

The Government has emphasised the importance it places on policy decisions being based on sound evidence. In April 2008, the Prime Minister, the Hon Kevin Rudd, emphasised the need for 'a robust, evidence-based policymaking process'.

Recognising that evidence is a vital input to the policymaking process is important. However, while using evidence can sometimes be straightforward, it can often be challenging. In some cases there may be a ready supply of good evidence. In other cases, evidence may be scant, uncertain or contested.

The purpose of this paper is to briefly examine some important issues related to producing and using evidence. A better understanding of the nature of evidence will ensure it is used well in the policymaking process.

Producing evidence

The first step in using evidence well is to understand what evidence is and how it is produced. The Macquarie Dictionary defines evidence as 'ground for belief; that which tends to prove or disprove something; proof'. That is, evidence is generated as an answer in response to a question or questions.

Having asked a question, a hypothesis is then put forward, proposing an answer. Researchers then search for evidence to prove or disprove the hypothesis. In addition to the question, which in the policymaking context is often defined by the Government, the production of evidence requires data, a theory on which to base a hypothesis that can be disproved and a methodology to test the hypothesis.

Evidence may come in all forms, including, for example, qualitative information from opinion surveys, structured interviews or focus groups. In this paper data is used to refer to the quantitative summaries of all forms of evidence. This is not to imply that only quantitative information matters, but there are some particular challenges to evaluating and testing qualitative evidence that are beyond the scope of this paper.²

The role of data

Of itself, data do not constitute evidence; data are merely numbers and can exist in abstract. For example, if we say that 'the unemployment rate in Australia was 4.3 per cent in October 2008' this is not in itself evidence of anything. If, however, we say that the Australian Bureau of Statistics' measure of the unemployment rate has

2 Those interested in issues relating to qualitative information are referred to Mays and Pope (1995).

fallen from 10 per cent in the early 1990s to 4.3 per cent in October 2008, then this may be evidence that unemployment has fallen and that it is easier to get a job in Australia.

A useful evidence base depends on good quality data (Gruen and Goldbloom 2008). Problems with data may mean that robust evidence cannot be developed. Appropriate data may not exist, or data sets may be incomplete or not available in time series. Data can also be unreliable in quality and one set of data can be incompatible with others. It may also not be possible to directly measure the concept that is to be tested, in which case a proxy may be used.

Data, in combination with theory, are used to develop evidence. Data are used to test hypotheses and therefore provide proof (or evidence) of a proposed theory. Knowing the sources and methods used to produce data is crucial when determining which data should be used to test hypotheses. Unemployment is a good example.

Unemployment is a lack of employment. It can only be measured once it has been defined (Marston and Watts 2003). To establish useable data on unemployment, statistical agencies, like the Australian Bureau of Statistics, conduct a sample survey based on a particular definition of the concept of unemployment. This definition involves identifying benchmark criteria such as age, and availability to work. It also involves making methodological assumptions, for example, about the validity of sample surveying. In assessing evidence, it is necessary to consider the definitions that have been applied in the collection of data as it may affect the conclusions drawn.

The role of theory

Data without a framework can only tell us that two sets of numbers (variables) are correlated. Theory provides a framework for how we perceive the world. It is theory that allows us to move beyond correlation to establish causation. That is, does a change in one variable cause the change in the second variable?

When data are used to test a hypothesis we can have more confidence in what the data reveal, that is, in the evidence, if the hypothesis is based in theory. As such, in using data to test hypotheses, the strength of the evidence is dependent not only on the quality of the data and the statistical methods used, but also on the robustness of the theory being used.

The theory being used will be robust when it has generated many hypotheses that have been rigorously tested over time with comprehensive and good quality data. In other cases, a lack of appropriate data or lack of time may mean that theories are less robust because they have not been subject to rigorous testing. It may also be the case that a theory has been subject to some testing, but has not yet been shown to be robust in all possible situations. For example, testing may show the theory to be good when

Using evidence well

applied to European countries, but may not yet have been tested for Australia. If the situation in Australia mimics the European situation, then the theory may be considered to be more robust than where only limited testing has taken place.

Identifying a correlation in the data can lead to the exposition of a theory to explain the relationship between the data. For example, the Cobb-Douglas function, widely used by economists to represent aggregate production of goods and services in an economy was developed when Cobb and Douglas noticed a relationship in the data between labour, capital and shares of total output (Cobb and Douglas 1928).

Theory can also drive the production of data. For example, analysis of tax reform proposals often involves detailed, disaggregated analysis of the reform's distributional impacts (Carnahan 1998). While the conceptual framework for such analysis is reasonably straightforward, the data that is available may not be appropriate for testing the hypothesis of interest to the policymaker. Theories can provide a guide for the collection of more appropriate data.

The role of methodology

The quality of evidence is also highly dependent on the methodology used to collect data and test hypotheses. A great deal of judgement goes into the manner in which data is collected and hypotheses tested.

The quality of evidence will depend on the methodology used to test hypotheses (Leigh 2009). The most important issue for researchers to determine is how to treat the counterfactual (that is, what would happen in the absence of any policy action). Good evidence should be able to distinguish the change resulting from the policy intervention relative to what would have otherwise happened. Approaches that can be used by researchers include randomised trials, natural experiments, before and after studies, econometric analysis and correlation analysis.

The methodology selected by researchers should also take non-target variables into account. Theory plays a critical role here as it helps ensure that the outcome is caused by the policy intervention and not an unaccounted for factor. If non-target variables are not taken into account, policies can lead to unintended consequences. Methodologies that address these issues include economy-wide modelling, complex systems analysis and factor analysis.

Another important issue that can affect the quality of evidence is the variables researchers use to test their hypotheses. For example, the augmented Solow-Swan model assumes gross domestic product to be a function of physical and human capital, labour and the level of technology. In testing this theory it is difficult to determine how to measure the amount of physical and human capital, labour and technology within

an economy. It is also hard to capture the interaction of these variables. For example, technology affects the productivity of physical capital and labour.

It is also often difficult to measure outcomes or the factors that may be affecting them. For example, the health status of the population is dependent on a large number of factors, including bio-medical, lifestyle and socio-economic factors. Researchers have also tried to determine whether the level of resources in the health sector affects health. That is, do countries that spend more on health have better health outcomes? For policy purposes the interest would be in testing hypotheses about what types of expenditure make a difference to population health outcomes.

Such research has to use a number of proxies. For example, life expectancy is often used a measure of health status even though it only measures length of life and does not take into account the quality of life. Although there are some measures, such as the World Health Organisation's Health-Adjusted Life Expectancy, which attempt to capture quality of life, these indicators are not widely available in time series, which limits their use. As a result, many studies fall back on life expectancy, even though they know that it is not the best measure (Wilkie and Young 2009).

Using evidence

Establishing that a problem exists is the first task for evidence in the development of policy. Evidence is then used to determine whether the proposed remedy (or which of the proposed remedies) will be effective in addressing the problem without unintended consequences.

The stock of data available to policymakers in some areas is very large, and may in part be apparently contradictory. The key question is to determine whether this constitutes evidence. Using evidence well in the policymaking process requires an understanding of the limitations of evidence and how these influence the application of evidence in solving policy problems.

Framing the question

How a policy question is framed can affect the use of evidence. In the policymaking process, the policy objective often determines the question. However, in many cases, policy objectives may not be clear or they may have changed over time.

Framing the policy question is particularly important as it can broaden or narrow the evidence that is considered relevant. For example, illicit drug taking can be framed as a crime and justice problem, a health problem, or a social problem. By framing the issue in terms of crime and justice, relevant evidence and policy solutions will naturally focus on drug-related crime, enforcement, detention and rehabilitation. On the other

Using evidence well

hand, where it is viewed as a health problem, evidence and policy solutions will focus more on health conditions that may predispose individuals to take drugs and on strategies to prevent and cure addiction. Alternatively, from a social perspective, evidence and policy solutions will focus on social conditions that may predispose individuals to take drugs and how to address disadvantage and social exclusion.

Another example of how framing a policy problem affects the evidence considered relevant is climate change. It can be framed as either an environmental or economic problem. By framing climate change as an environmental problem, the scientific evidence is the most relevant evidence. On the other hand, when Nicholas Stern framed climate change as an economic challenge, he took the scientific evidence as given and focused on the economic implications of climate change (Stern 2006).

Identifying relevant evidence

Identifying relevant evidence of high quality can be particularly important where there is a large body of evidence. Theory is the best guide to what evidence is relevant. A systematic approach may be required to select a manageable set of high-quality, reputable, relevant evidence (Leigh 2009).

One method of selecting evidence may be to use a hierarchy of research methods. Under this approach, studies are ranked according to the research methods used, for example studies using randomised trials are preferred to those using natural experiments. Further, studies using natural experiments are preferred to those using a before-after study.

There are a number of other factors that may be relevant when collecting or selecting evidence (Leigh 2009). Studies will be preferred if they are published in peer-reviewed journals. Other things being equal, studies that use Australian evidence may be more relevant than those using international evidence. More recently published studies will also be preferred. Studies that can be replicated will be preferred over those that cannot. And studies that directly address the policy question under consideration will also be given preference.

Being clear on what the evidence says

Most social science evidence is historically contingent, that is, it has been developed using quantitative and qualitative data collected about the past. As a result, data and evidence only records how variables have related in the past and how people have behaved in the past. There is no guarantee that things will be the same in the future.

The historical contingency of evidence also means that it is highly dependent on the circumstances that existed when the evidence was collected. That is, the evidence

reflects the institutional, policy and economic circumstances that prevailed over the period to which the evidence refers. Changes in circumstance can affect the interpretation of and weight placed on evidence. The fact that evidence primarily refers to the past, and that circumstances continually change, means that evidence needs to be regularly updated and reviewed. It also emphasises the importance of theory, which can be used to test whether underlying assumptions about behaviour have not changed over time.

The contingency of evidence on institutional, policy and economic circumstances is particularly relevant when assessing evidence drawn from overseas. For example, according to the OECD PISA data, Finland's students, on average, perform better than Australian students (OECD 2006). This could reflect superior education policy or other factors. Given the differences between Finland and Australia, can we be sure that Australian students would fare better if we adopted some of Finland's education policies, for example, starting school at age seven?

Care needs to be taken in interpreting evidence. In areas where evidence is contested, it is very easy to apply evidence where it is irrelevant, misinterpreted or overstated. For example, it is important to distinguish between whether there is no evidence of danger or evidence of no danger. If researchers cannot link an activity to a problem, then restricting the activity may still be appropriate for managing risk. On the other hand, if researchers prove that an activity does not cause a problem, then no policy response may be required.

Assessing the risks for policymaking

Policy decisions, like all decisions, require an assessment of risks. The evidence on which decisions are based is often ambiguous and not comprehensive. Evidence is about the past and policy is about the future: extrapolating what may happen in the future on the basis of evidence that describes the past is both difficult and uncertain.

As such, when assessing evidence it is necessary to determine its limits. What is the range for error? Is it possible that a policy will result in either large or small benefits? Or are there large costs if the evidence is misleading?

In determining the best way forward, it might be appropriate to commission an expert opinion on the evidence. However, in some cases, policy decisions may need to be made before there is time to improve the evidence base. In such cases, it is appropriate to use theory to guide policy decisions. Under these circumstances, policymakers may decide to manage the risks associated with inadequate evidence by adopting a 'no regrets' or 'no harm' policy stance.

Conclusion: how to use evidence well

Using evidence well to inform and influence policy decisions requires a good understanding of how framing a policy issue can affect the use of evidence. If a problem is framed too narrowly it can significantly affect the evidence that is considered relevant. Some problems, especially those that deal with complex systems (for example, the environment) or attitudinal and behavioural issues (for example, illegal drug taking or obesity) require a broad framework that allows for multiple causes and policy responses. Taking a broad or holistic perspective may also reduce the possibility of the implemented policy having unintended consequences. This highlights the importance of using theory to identify relevant evidence.

Knowing where evidence comes from is also particularly important in using evidence well. Understanding the theoretical basis and source and methodology of evidence can help us determine the weight that can be placed on evidence and where it can be appropriately used. It is also important to be clear about what the evidence does and does not tell us. For example, is there no evidence of danger or evidence of no danger?

In areas where there are evidence gaps or there is no consensus among available evidence it is appropriate to fall back on well-tested theory that has had widespread successful applications and partial indicators. Where evidence is sparse, it is particularly important to try to use multiple sources of evidence. The degree to which evidence is uncertain may determine the number or range of policy options proposed. In such cases, it may be appropriate to manage risks by advocating 'no regrets' or 'no harm' policy responses.

References

- Banks, G 2009, 'Evidence-Based Policy-Making: What is it? How do we get it?', ANZSOG/ANU Public Lecture Series, 4 February, Canberra.
- Carnahan, M 1998, 'Does Demand Create Poor Quality Supply: A Critique of Alternative Distributional Analysis', *Economic Roundup*, Spring, Treasury, Canberra.
- Cobb, C and Douglas, P 1928, 'A Theory of Production', *American Economic Review*, vol 18, supplement, pp 139-165.
- Leigh, A 2009, 'What evidence should social policymakers use?', *Economic Roundup*, Issue 1, Treasury, Canberra.
- Lundberg, E 1984, 'Richard Stone Presentation Speech' The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel.
- Gruen, D and Goldbloom, A 2008, 'The Smarter Use of Data', Plenary address to the NatStats08 Conference, Melbourne, 20 November.
- Marston, G and Watts, R 2003, 'Tampering with the Evidence: A Critical Appraisal of Evidence-Based Policy Making', *The Drawing Board: An Australian Review of Public Affairs*, vol 3, no. 3, pp 143-163.
- Mays, N and Pope, C 1995, 'Qualitative research: rigour and qualitative research', *BMJ*, 311:109-112.
- Mulgan, G 2003, 'Government, Knowledge and the Business of Policy-Making', keynote address at the national Institute for Governance Conference, *Facing the Future: Engaging Stakeholders and Citizens in Developing Public Policy*, Canberra, 23-24 April.
- OECD 2006, *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.
- Rudd, K 2008, 'Address to Heads of Agencies and Members of Senior Executive Service', Great Hall, Parliament House, Canberra, 30 April.
- Stern, N 2006, *Stern Review: The Economics of Climate Change*, HM Treasury and UK Cabinet Office, London.
- Wilkie, J and Young, A 2009, 'Why health matters for economic performance', *Economic Roundup*, Issue 1, Treasury, Canberra.

