In data we trust: Challenges of compiling input-output tables as official statistics

Andrew Cadogan-Cowper

Paper for the 23rd International Input-Output Association Conference

June 2015, MEXICO CITY

Abstract:
This paper examines the challenges for National Statistical Organisations (NSOs) in compiling Input-Output tables (IOTs). NSOs such as the Australian Bureau of Statistics (ABS) take the responsibility of producing 'official' statistics extremely seriously, and this ultimately requires them to be more risk averse, and place a higher threshold on data quality than many other statistical providers. Using the ABS compilation process as an example, this paper looks at some of the challenges faced by official statisticians, and will serve as an important reminder to those constructing and interpreting I-O models to carefully consider the inherent strengths and weaknesses of the original data.

Introduction

National Statistical Organisations (NSOs), such as the Australian Bureau of Statistics (ABS), are generally considered to be the official statistical providers for their country. As we will explore later, they abide by a range of fundamental principles and unique community expectations. One particular expectation is that the statistics they produce can be trusted. Trusted to be objective, and compiled on strictly professional and scientific grounds.

NSOs know that trust is hard won and easily lost. Therefore, and unsurprisingly, NSOs are risk averse when it comes to producing statistics in fields where compilation or interpretation issues may impact on their standing in the community. So assuming firstly that a policy case exists, what criteria should be used by an NSO in assessing whether to produce a particular set of official statistics? This is a perennial issue for NSOs, and the answer of course, depends on a number of factors.

One of those factors is undoubtedly the *reliability* of the statistics – can the concept of interest be sufficiently measured given the purpose for which the statistics will ultimately be used? With these thoughts squarely in mind, this paper will examine the compilation and subsequent uses of ABS input-output tables (IOTs)[1]. It will look at the range of principles that underlie the concept of official statistics, in the process revealing some of the challenges that IOT compilation presents for NSOs. The reader will see that IOTs are a hugely valuable component of the statistical landscape, but also prone to misuse.

Finally, we'll observe that statisticians have an important role to play in reminding those who construct I-O models to carefully consider the inherent strengths and weaknesses of the original data; and for those who interpret I-O modelling results, the suitability of their model to accurately answer particular questions.

A brief history of the ABS I-O program

IOTs provide a framework for recording the supply and use of products in an economy, and observing the structure of interrelationships between industries at a highly disaggregated level. They also provide detailed information on the composition of transaction values, including that part represented by taxes, transport, wholesale and retail margins. The rich detail of the dataset make IOTs the starting point for many analytical exercises.

With the release of the 2012-13 tables in 2015, the ABS will have published 26 IOTs for Australia. Previous tables were for 1958–59, 1962–63, 1968–69, 1974–75, 1977–78 to 1983–84, 1986–87, 1989–90, 1992–93 to 1994–95, 1996–97, 1998–99, 2001–02, and 2004–05 to 2009–10. Before the ABS adopted this field of work, Cameron published 3 sets of Australian I-O tables, for reference years 1946-47, 1953-54 and 1955-56.[2]

[1] The expression 'I-O tables' can incorporate a range of tables, a number of which are available under the ABS catalogue numbers 5209.0 and 5215.0. Within this paper, unless explicitly mentioned, the term is used to refer to rectangular tables, which are primarily a more detailed version of supply use tables. Various users, especially outside of national statistical offices, often use the term to refer to 'square' or 'symmetric' (industry by industry, or product by product) I-O tables that combine supply and use into the one table.

[2] See Cameron (1957), (1958) and (1960).

During their history in the ABS work program, IOTs have been produced intermittently with delays of up to 10yrs between reference year and publication year (e.g. the 1962-63 tables were not published until 1973), as well as data gaps of up to 5 reference years existing between tables (e.g. the absence of any tables for the years 1963-64 to 1967-68, and 1978-79 to 1982-83). However, since the release of the 2004-05 reference year, IOTs have been compiled annually and published approximately 3.5 years after the reference year in question (for instance the 2004-05 tables were published at the end of the 2008 calendar year). In 2015 this will be reduced to a lag of approximately 2 years between the end of the reference year and publication.

Compilation of IOTs

Producing a high quality and detailed IOT is a time and data intensive undertaking. At the ABS, the process begins with the compilation of 'Supply-Use' tables (SUTs). The SUTs actually serve dual purposes; firstly as national accounts benchmarks, and secondly as a stepping stone to production of IOTs.[3]

Compilation of the SUT begins with the receipt of data from our annual economic activity survey (EAS). Receipt of this data occurs approximately 9 months after the end of the financial year to which it relates – for instance initial data for the financial year 1 July 2013 to 30 June 2014 was received in March 2015.

This data is combined with an extensive range of other data sources that become available at various intervals – such as trade, household expenditure, and government finance – to produce a SUT. ABS SUTs are balanced at the level of 67 industries and 301 products. ABS IOTs are balanced at the level of 115 industries and 1296 products.

Now, it's important to note that the major data source on production, EAS, is designed at the level of 115 industries, but only edited at the level of 86 industries. So from an *industry* perspective, there is good coverage for SUT industry compilation. However from a *product* perspective, a compiler now needs to spread the output of 86 industries across 301 S-U products and, later, 1296 I-O products.

This is done by using relevant weights to assign industry output to products that are understood to be produced by that industry. Those weights are developed over time using industry intelligence and a rolling supplementary survey program that asks for production detail from respondents. Despite the supplementary survey program, the lack of more regular product detail data in domestic supply is a significant limitation.

With respect to the intermediate use matrix, data is even scarcer. The EAS survey is able to provide an aggregate expense for purchases of *'materials, components, containers, packaging materials, electricity, fuels and water'*, as well as transport and insurance expenses. From here, compilers again split industry data across 301 SU products and 1296 IO products, based on industry intelligence, historic weights and other available data (e.g. energy and water resource use from environmental accounts). So it should be noted that getting the initial product detail correct in both the supply and intermediate use matrices is highly challenging because of the paucity of data.

---

[3] It should also be noted that the SUTs are maintained as a time series, and are therefore subject to revisions. The IOTs are not revised – they are compiled as a snapshot in time.

Now at this point it is essential to understand that the SU and IO tables are created within a very robust conceptual framework[4] that imposes several fundamental identities. These are:

1 – For each product, supply must equal use;

2 - For each industry, the sum of outputs produced must equal the sum of inputs used; and

3 – GDP as measured by the income, expenditure and production methods must be equal.

Naturally, after the initial population of the SUT, these 3 fundamental identities do not align. This is because the inputs of survey and administrative data are of varying levels of quality due to differences in scope, concept, method, and process. As a result, compilers now need to investigate the available data and determine the best estimate. In making this determination, the compiler must consider both the micro (e.g. supply chains) and macro (e.g. industry contributions to GDP) story that is evolving, and confront this against expectations.

In forming an expectation on industry performance, the (earlier) published quarterly national accounts estimates (based on economic indicator surveys and other data sources) are an excellent confrontation tool. Significant discrepancies between the estimates are carefully reviewed before a final decision on the benchmark is made.

Another excellent confrontation tool is the environmental accounts. For example, confronting monetary estimates of energy use by industry against physical flows of energy by industry allows an analyst to look at implied prices per unit of energy.[5] If the implied price is nonsensical, an investigation into the relative accuracy and quality of the competing data sources takes place. Similarly, applying detailed product deflators (price indices) allows us to produce an SUT in prices of the previous year to remove price effects. If the resultant implicit price deflators (IPDs) don't match economic reality, a review of the current price estimate and the associated price deflator is necessary to find a resolution.

As a result of these confrontations and industry analysis, balancing decisions are made. The challenge of making adjustments in a fully integrated system is that every adjustment will have an effect somewhere else in the system. If you push down in one spot, it pops up elsewhere.

After several iterations of manually balancing the columns (industries) and rows (products), a 'RAS' procedure is used to remove the remaining imbalances. This procedure begins with columns balanced, and rows out of balance. The imbalance in services rows is assigned to the intermediate use of using industries (which throws columns out of balance). Then the imbalance in the industry columns is assigned to goods rows (throwing rows out again). Finally the rows are balanced by putting the residual into inventories. The ABS is now looking into the possibility of using 'constrained optimisation'[6] as a more sophisticated replacement for the RAS process.

---

[4] The *2008 System of National Accounts*.

[5] See for instance the ABS Energy Accounts, Cat. No. 4604.0.

[6] 'Constrained optimisation' (CO) is a term for a class of mathematical/operations research problems. It requires the definition of *decision variables* (e.g. cell values for a balanced SUT) and *constraints* (e.g. Supply equals use, cells that can't be negative). A mathematical program, or 'solver', is then used to *optimise* something, e.g. minimise overall table adjustments.

Once the SUT benchmarks are established, the focus shifts to disaggregating the SUT to form an IOT. As mentioned earlier, the formation of the IOT requires that the SUT (67 industries and 301 products) now be disaggregated into 115 industries and 1296 products using weights and industry intelligence. A limited departure from the SNA08 occurs here - whereby all transport margins, regardless of whether they are separately invoiced, are removed from the basic price of goods.

The ABS publishes the full detailed rectangular IOT in ABS Catalogue No. 5215.0. We also produce a range of 'derived' outputs, including a symmetric / square 'modelling' table in ABS Catalogue No. 5209.0. As the symmetric IOT is such an integral feature for IO modelling efforts, it is also important to understand the assumptions that are used in its creation.

Converting the basic IOT to a 'symmetric' IOT (SIOT)

A basic IOT can be converted into a SIOT of either product × product (PP), or industry × industry (II) form. To produce a PP SIOT, a technology assumption must be used. A choice can be made between (i) the assumption that each product has its own specific production function, irrespective of the producing industry, or (ii) each industry has its own specific production function, irrespective of the products it produces.

To produce an II SIOT, a fixed sales structure assumption must be used. That is, either (i) each industry has its own specific sales structure, irrespective of the products it produces, or (ii) each product has its own specific sales structure, irrespective of the producing industry. Because data is collected on an industry basis, not a product basis, the assumptions underpinning an II SIOT are possibly more acceptable from an NSO perspective - hence the ABS publishes an II SIOT, but not a PP SIOT.

Regional IOTs

Although the ABS only produces IOTs at the national level, others have compiled the tables for sub-national and regional areas. In fact IOTs have been prepared for all states and territories in Australia. Compilation methods can vary, but one technique is to apply regional economic indicators and survey information to national estimates. This allows state governments and economists to develop state or regional I-O and CGE models.

The attraction of regional IOTs is undeniable – real-world policy and investment decisions are made at varying geographical scales; and so for some purposes a national IOT may be a relatively blunt instrument. So with a clear policy demand, the fundamental issue for the ABS is whether the data supports such estimates as official statistics, and whether there is something unique about the ABS' position that would enable it to produce those estimates at a better quality than other institutions, researchers or economists. One historical shift that reduced the ABS capacity to support regional IOTs, was a change of the production unit that occurred in 2002.

Prior to 2002, the ABS economic unit model was based on a production unit known as an 'establishment'. Establishments were location-based units. The business register from which survey frames were taken was maintained by the ABS, using ABS data to update the information relating to businesses, and information from the Australian Taxation Office (ATO) to identify new and deceased businesses.

With the introduction of a new tax system in Australia in 2000, the ATO created a whole-of-government register of businesses, the Australian Business Register (ABR). When the ABS adopted the ABR and coupled it with an ABS maintained component, the statistical unit's model changed to align better with taxation reporting requirements. As a result, the production unit became a 'type of activity unit' (TAU), defined according to an activity concept, in contrast to the former location based 'establishment' model.

Although many TAUs will still be performing activity within a single I-O industry and at a single location, for regional IOTs, the loss of a geographic dimension in the statistical unit has been unhelpful. More recently the ATO has begun to collect and compile location based information for its business register. The ABS is investigating the potential for this to aid in the development of regional IOTs.

Uses and abuses of IOTs

IOTs have many important uses. Within an NSO like the ABS, they are used in the creation of national accounts benchmarks, producer price indexes, productivity statistics, environmental accounts and tourism (and other) satellite accounts. Externally, the tables are heavily used in economic and environmental analysis. Sometimes the use is appropriate, and at other times it is quite inappropriate. The most contentious area of use externally is within economic modelling. Within that realm, most models fall into one of two broad categories. The first, and most simplistic, are *I-O multipliers*.

Relationships between industries and between products can be expressed as fixed coefficients which, for example, allow analysts to study the various resources and components embodied within other components and products. Using mathematical techniques such as matrix inversion, these coefficients can be turned into multipliers. In turn, these multipliers can be used to derive estimates of industrial and consumption-induced responses to particular economic policies or events.

Use of IOTs to derive multipliers can give very misleading results, and this type of use is heavily cautioned by the ABS (2013), Productivity Commission (2013) and others. Briefly, this is because multiplier construction makes unrealistic assumptions - including the 'static' or 'fixed' nature of the coefficients (unrealistically implying that industrial structures and marginal responses by industries remain constant in the face of economic events); and the lack of supply side constraints (for instance the assumption that there are unlimited 'ghost' workforces with appropriate skills available for employment) which mean these applications fail to consider the opportunity cost of both spending measures and resource allocation. As a result, this type of analysis generally results in a misleading overstatement of the impacts on employment and value added. A more comprehensive list of the limitations of using multipliers for economic impact analysis is available in appendix 2.

An excellent expression of these types of concerns was made in a speech by former Australian Treasury Secretary, Dr. Ken Henry, to his staff during a period of nearly full national employment:

'Consider, for example, recent commentary in the press which argues that the government should support a nuclear power sector because jobs would be created. Where will the nuclear scientists

and technicians come from? Is it seriously being suggested that they will come from the dole queue or from indigenous community development employment projects?'

'The next time any of you get an opportunity to write a co-ordination comment on a cabinet submission that proposes a taxpayer-funded handout for some stunning new investment proposition - and I predict that some of you won't have to wait very long for such an opportunity - I suggest you draw attention to the submission's failure to identify the businesses that will lose labour and be forced to reduce output if the proposal is agreed to.'[7]

Where the inter-industry relationships, expressed as multipliers or coefficients, are more appropriately used, are in circumstances where a 'static' analysis is sufficient – i.e. where understanding an existing relationship is itself the goal, rather than understanding how the relationship may change in response to a shock.

Examples of this type of work are numerous, and can include analysis of Trade in Value Added (TiVA), and Consumption-based natural resource accounting (for instance CO2 emissions footprinting, water footprinting, energy footprinting, etc.). In 2012 the ABS itself constructed an environmentally-extended IOT to analyse GHG emissions on a consumption-basis for Australia (ABS 2012).

The second major use of IOTs in economic modelling is their use as the core for *Computable General Equilibrium (CGE) models*. These models have the capacity to overcome many of the concerns associated with multipliers – for instance coefficients no longer need to be held 'static', and supply side constraints are introduced.

One natural drawback of CGE modelling is that the results are heavily reliant on each model's particular specifications, and these can be quite opaque to all but the model compilers themselves. The specifications (or assumptions) must also be kept in mind when interpreting results. For instance if a model assumes that a small cut to company tax will result in greater business investment, then the model will inevitably 'find' the assumption to be true - i.e. converting an assumption into a conclusion.

That said, when a 'dynamic' analysis (meaning a 'shock' is applied to a system that is in equilibrium, for instance to determine the economic impacts of a policy change) is being undertaken, CGE models are typically far superior to I-O multipliers, and are used extensively by the Productivity Commission and the Australian Treasury.

The concept of 'official' statistics

The word 'official' is defined by the Oxford dictionary as *'relating to an authority or public body…'* and this corresponds well with the widely held notion that official statistics are those statistics published by Governments[8]. In Australia, the ABS is regarded as the nation's chief statistician, and is supported by legislation including:

---

[7] Henry (2007)

[8] Intergovernmental organisations such as the UN and IMF would typically also be included in this definition.

- the *Australian Bureau of Statistics Act 1975* – which establishes the ABS as an independent statutory authority, and describes the terms under which the Australian Statistician can be appointed to, and removed from, office; and
- the *Census and Statistics Act 1905* - which provides the Australian Statistician with the authority to conduct statistical collections, including the Census of Population and Housing, and, when necessary, to direct a person to provide statistical information. The Act requires the ABS to publish and disseminate compilations and analyses of statistical information and to maintain the confidentiality of information collected under the Act.

In particular, the independence of the ABS from the rest of Australian government helps to establish trust in the user community that ABS statistics are impartial and free from political interference. In the provider community, the legislated requirement to maintain confidentiality underpins the trust of providers.

But beyond impartiality and confidentiality, what other elements constitute official statistics? Well, at the United Nation's General Assembly in January 2014 the UN's *'Fundamental Principles of Official Statistics'* were endorsed[9]. A full articulation of the principles can be found in appendix 1 of this paper, but here is a summary:

1. Relevance, Impartiality, and Equal Access
2. Professional Standards and Ethics
3. Accountability and Transparency
4. Prevention of Misuse
5. Cost-Effectiveness
6. Confidentiality
7. Public Legislation
8. National Coordination
9. International Standards
10. International Cooperation

In many respects, the principles are a foundation from which to establish trust. For the data user - trust that any given set of statistical outputs will exhibit objectivity, professionalism, transparency and scientific rigour. For the data provider - trust that the imposition on them will be minimised, the value of their data maximised, and their privacy maintained and supported by law.

The challenge of compiling IOTs as official statistics

Examining the construction of IOTs in light of the above, UN principles 2 and 4 in particular are worth further discussion. In full, principle 2 reads:

'To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.'

---

[9] They had earlier been adopted by the United Nation's Statistical Commission in 1994; substantively the same as they are today. See: United Nations (2014).

As we saw in detail earlier, construction of IOTs is a complex, data intensive operation. Inevitably, there are data gaps, and judgements are needed in order to complete the process. For the ABS, the compilation challenges are broadly:

1. Stretching *industry* survey data across *products* in the supply matrix.
2. Stretching *industry* survey data across *products* in the intermediate use matrix.
3. Aggregate balancing decisions.
4. The fixed sales structure assumption used to create the II SIOT.

Overcoming these challenges requires a significant number of assumptions and imputations. While the need to fill data gaps is not ideal, that is also the stark reality of producing statistics. It's easily forgotten that statistical data sets are only estimates. The key question is how good the estimate needs to be to be worthy of publication - and that answer depends on how the estimate will be used – is the estimate fit for purpose?

Obviously NSOs can't control how their published data is used, but having an understanding of the primary purposes for which data is being applied is essential. As noted earlier, in the case of IOTs the ABS is aware of many highly credible and important uses where the estimates are fit for purpose. In other cases however, the estimates are either not fit for purpose, or are used to construct multipliers which are then misapplied.

The frequent misuse of IOTs has been deeply concerning to the ABS and many of its users. In 2001, the ABS ceased publishing I-O multipliers, and also began to include commentary explicitly highlighting their inherent shortcomings within its Concepts, Sources and Methods (CSM) publication (ABS 2013). While UN principle 4 states that *'Statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics',* the ABS has not sought to comment on individual cases of misuse, preferring instead to draw user's attention to the CSM instead.

Conclusion

This paper has given an overview of IOT compilation at the ABS, then set out two broad areas of concern with respect to compiling IOTs as official statistics. The first is the volume of assumptions that are needed to cover gaps in the data. The second is the frequent misuse of the final estimates. These concerns must be balanced against the huge value IOTs create for users when appropriately applied in the various ways described above.

Ultimately the question of whether any particular estimate should be compiled as an official statistic is unavoidably philosophical, and all NSOs will draw their own conclusions. In the ABS, as in perhaps most NSOs, it is deemed that the power of the IOTs to inform - and the benefit that brings – justifies their creation. The significant misuse is best dealt with by educating the user community with clear guidance on concepts, sources, methods and interpretation.

Appendix 1: UN Fundamental Principles of Official Statistics

1. Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.

2. To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

3. To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.

4. The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

5. Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.

6. Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

7. The laws, regulations and measures under which the statistical systems operate are to be made public.

8. Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

9. The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.

10. Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.

Appendix 2: Inherent shortcomings and limitations of multipliers for economic impact analysis

- ***Lack of supply-side constraints***
  The most significant limitation of economic impact analysis using multipliers is the implicit assumption that the economy has no supply-side constraints. That is, it is assumed that extra output can be produced in one area without taking away resources from other activities, thus overstating economic impacts. The actual impact is likely to be dependent on the extent to which the economy is operating at or near capacity.

- ***Fixed prices***
  Constraints on the availability of inputs, such as skilled labour, require prices to act as a rationing device. In assessments using multipliers, where factors of production are assumed to be limitless, this rationing response is assumed not to occur. Prices are assumed to be unaffected by policy and any crowding out effects are not captured.

- ***Fixed ratios for intermediate inputs and production***
  Economic impact analysis using multipliers implicitly assumes that there is a fixed input structure in each industry and fixed ratios for production. As such, impact analysis using multipliers can be seen to describe average effects, not marginal effects. For example, increased demand for a product is assumed to imply an equal increase in production for that product. In reality, however, it may be more efficient to increase imports or divert some exports to local consumption rather than increasing local production by the full amount.

- ***No allowance for purchasers' marginal responses to change***
  Economic impact analysis using multipliers assumes that households consume goods and services in exact proportions to their initial budget shares. For example, the household budget share of some goods might increase as household income increases. This equally applies to industrial consumption of intermediate inputs and factors of production.

- ***Absence of budget constraints***
  Assessments of economic impacts using multipliers that consider consumption induced effects (type two multipliers) implicitly assume that household and government consumption is not subject to budget constraints.

- ***Not applicable for small regions***
  Multipliers that have been calculated from the national I-O tables are not appropriate for use in economic impact analysis of projects in small regions. This is because small region multipliers tend to be smaller than national multipliers since their inter-industry linkages are normally relatively shallow. Inter-industry linkages tend to be shallow in small regions since they usually don't have the capacity to produce the wide range of goods used for inputs and consumption, instead importing a large proportion of these goods from other regions.

References

ABS 2013, *Australian System of National Accounts: Concepts, Sources and Methods Australia*. Cat. No. 5216.0, ABS, Canberra

ABS 2012, Hao, S., Legoff, G., Mahadeva, K., Chia, N. and Cadogan-Cowper, A. *Australian Greenhouse Gas Emissions: Experimental Consumption-Based Estimates Via Input-Output Analysis*, 20th International Input-Output Association conference (IIOA) 2012, Bratislava, 24-29 June 2012.

ABS 1995, *Information Paper: Australian National Accounts: Introduction to Input-Output Multipliers, 1989-90,* Cat. No. 5246.0, ABS, Canberra

Cameron, B, 1957, '*The 1946-47 transactions table'*, Economic Record, vol. 33, pp. 353-60.

Cameron, B, 1958, '*New aspects of Australia's industry structure'*, Economic Record, vol. 34, pp. 362-74.

Cameron, B, 1960, 'Inter-sector accounts 1955-56', Economic Record, vol. 36, pp. 269-71.

Denniss, R. 2012, *The use and abuse of economic modelling in Australia,* The Australia Institute

Henry, K. 2007, *Treasury's effectiveness in the current environment*, Address to staff at the Hyatt Canberra, 14 March 2007.

PC 2013, Gretton, P, *On input-output tables: uses and abuses*, Productivity Commission Staff Research Note, September 2013.

United Nations 2014, *Fundamental Principles of Official Statistics,* Available at: http://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf (accessed 27-03-15)